Annex B Workshop Presentations

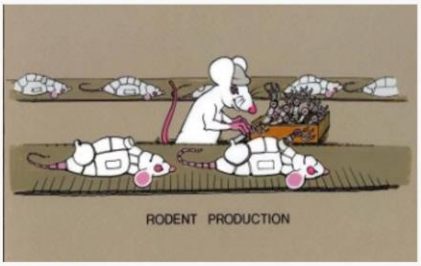
Table of contents

Presentations Day 1

1. Standardisation and Biosecurity in Laboratory Rodent Breeding

(Jutta Davidson, Urte Jäh, Charles River Laboratories)

STANDARDIZATION AND BIOSECURITYIN LABORATORY RODENT BREEDING



Jutta Davidson

Research Models and Services Specialist
Charles River Laboratories, GmbH

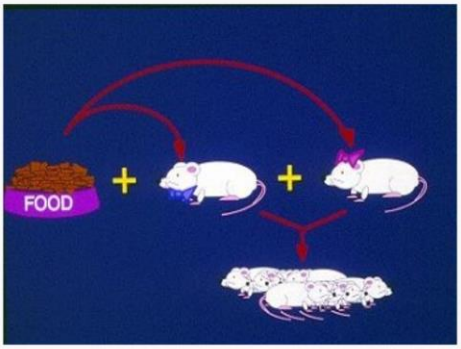
EVERY STEP OF THE WAY

charles river

Laboratory animals are used as "biological sensors" and are thereby essential for the reproduceabilityof scientific results in research!

... Therefore.....

Breeding of research animal models is far more than simply putting "a male and a female animal togetherand to hope for the best"!



EVERY STEP OF THE WAY

charles river

What do we have to do to provide you with healthy, genetically defined and suitable animals for your

Standardization

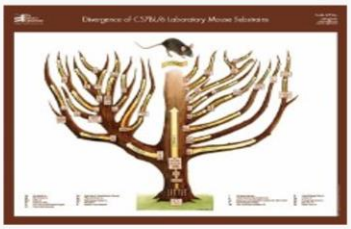
+

Biosecurity

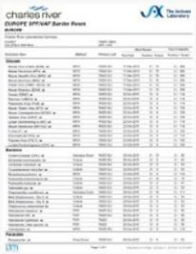
EVERY STEP OF THE WAY

charles river

STANDARDIZATION



Genetics



Health

Environment

Handling

EVERY STEP OF THE WAY

charles river

GENETIC STANDARDIZATION BY GENETIC STABILITY PROGRAMS

Goal: preserving integrity of a given mouse or rat strain

- Over time (Generations)
- over geographical distance (multiple breeding locations)

5 | EVERY STEP OF THE WAY

charles river

FACTORS IMPACTING GENETIC IDENTITY

Selection
Mutation
Genetic Drift

6 | EVERY STEP OF THE WAY

charles river

SELECTION

Natural Selection

- Plays a limited role in laboratory populations, especially if the rearing practices are similar in all subpopulations.

Unconscious Selection

- Docile temperament
- Litter size
- Short interval between successive litters
- Rapid growth
- Parental care of mother
- Traits that pose economical or other advantages

7 | EVERY STEP OF THE WAY

charles river

MUTATION

A mutation is a permanent alteration in the nucleotide sequence of one or more genes or in the number or structure of one or more chromosomes (Merriam Webster Dictionary)

Mutations can result from

- DNA copying mistakes
- exposure to ionizing radiation or mutagens,
- infection by viruses

8 | EVERY STEP OF THE WAY

charles river

GENETIC DRIFT

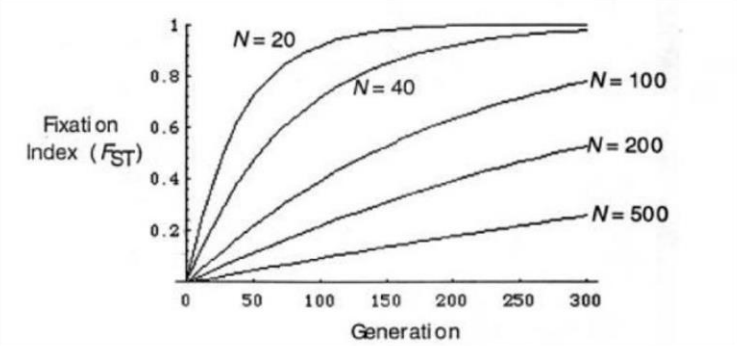
"The constant tendency of genes to evolve even in the absence of selective forces. Genetic drift is fueled by spontaneous neutral mutations that disappear or become fixed in a population at random"
(Lee Silver, "Mouse Genetics" Oxford University Press, 1995)

Leads to genetic divergence of two sub-populations

9 | EVERY STEP OF THE WAY

charles river

DYNAMICS OF GENETIC DRIFT



Interpretation by Wright (1978)

- F_{ST} in the range 0 to 0.05 indicates little genetic divergence
- F_{ST} in the range 0.05 to 0.15 indicates moderate genetic divergence
- F_{ST} in the range 0.15 to 0.25 indicates great genetic divergence
- F_{ST} above 0.25 indicates very great genetic divergence

10 | EVERY STEP OF THE WAY

charles river

MINIMIZING GENETIC DRIFT - OUTBREDS

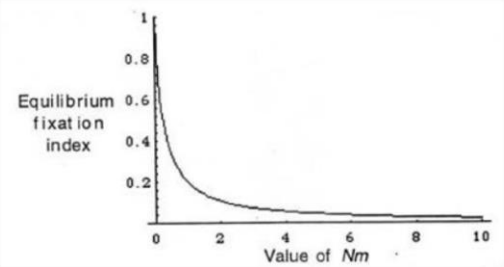
- Large Population Size
- Avoidance of Inbreeding
- **Migration between sub-populations**

11 | EVERY STEP OF THE WAY

charles river

MIGRATION BETWEEN SUBPOPULATIONS

White and Lee: „Migration can be viewed as a form of genetic glue that holds colonies together."



migration and random drift eventually reach anequilibrium stage at which the additional divergence attributable to random genetic drift in any generation is exactly offset by the homogenizing effects of the migration.

Nm = population size * migration rate

12 | EVERY STEP OF THE WAY

charles river

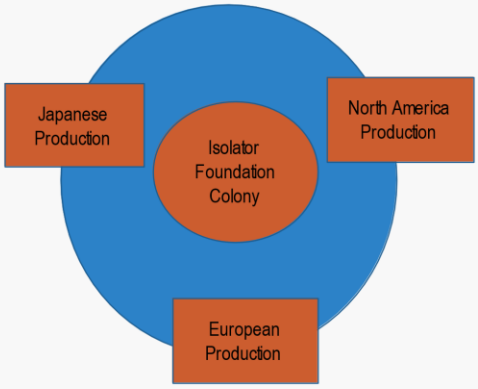
OUTBRED FOUNDATION AND PRODUCTION BREEDING

International Genetic Standardization (IGS) - Outbreds

Maintain diversity by preventing inbreeding, by standardising production colonies that are geographically separated such that each colony has the same range of genetic variation

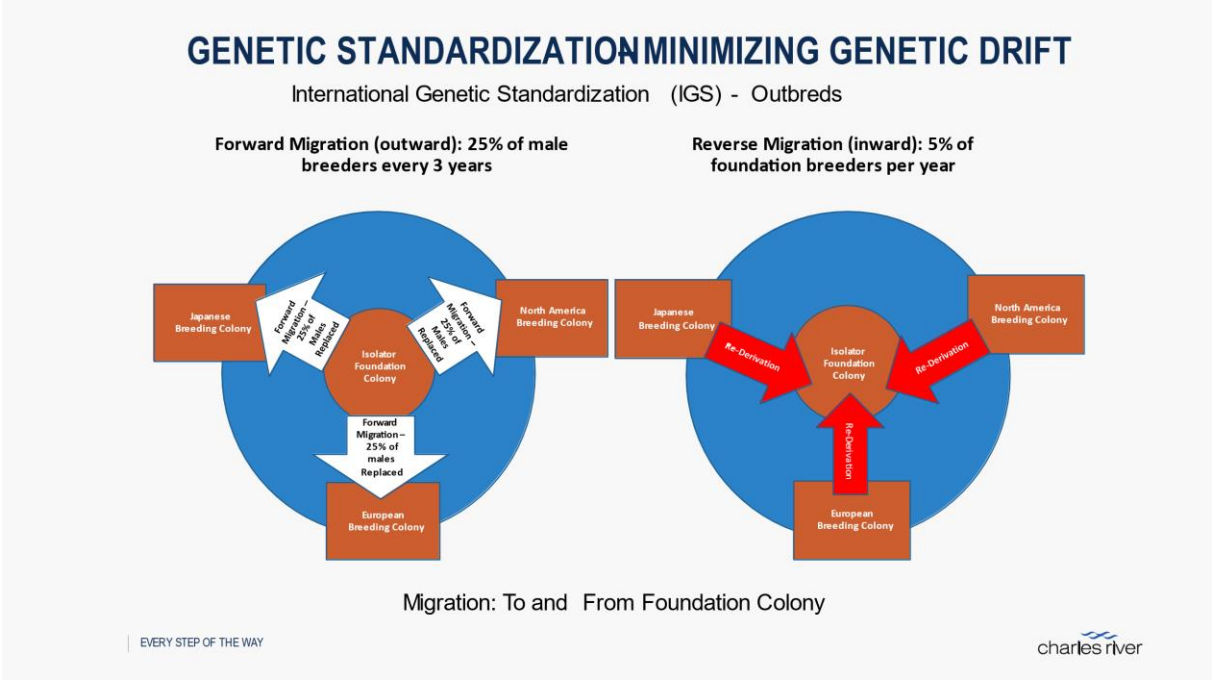
Outbred stocks: maintained using a foundation colony and 'satellite' production breeding colonies.

IGS is achieved by using
-Migration
-Rotational Breeding
-Quality Control



13 | EVERY STEP OF THE WAY

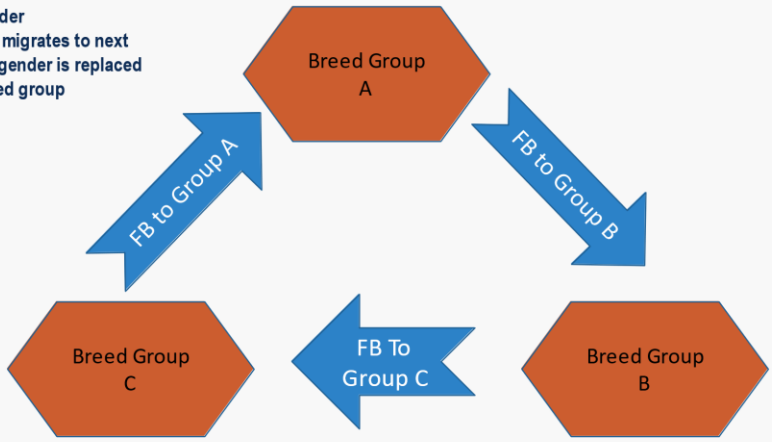
charles river


GENETIC STANDARDIZATION MINIMIZING GENETIC DRIFT
International Genetic Standardization (IGS) - Outbreds
Forward Migration (outward): 25% of male breeders every 3 years
Reverse Migration (inward): 5% of foundation breeders per year
Japanese Breeding Colony
Forward Migration – 25% of Males Replaced
Isolator Foundation Colony
Forward Migration – 25% of Males Replaced
North America Breeding Colony
Forward Migration – 25% of males Replaced
European Breeding Colony
Japanese Breeding Colony
Re-Derivation
Isolator Foundation Colony
Re-Derivation
North America Breeding Colony
Re-Derivation
European Breeding Colony
Migration: To and From Foundation Colony
EVERY STEP OF THE WAY
charles river

OUTBRED FOUNDATION AND PRODUCTION BREEDING

IGS - Outbreds- Production Breeding

FB: Future Breeder
only one gender migrates to next block, the other gender is replaced within same breed group



15 | EVERY STEP OF THE WAY

charles river

CHARLES RIVER'S GENETIC STANDARDIZATION PROGRAM

International Genetic Standardization (IGS) - Outbreds

IGS is a genetic management system using

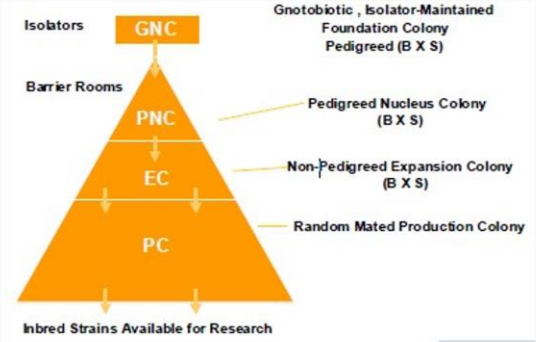
- pedigreed gnotobiotic foundation colonies
- Large Population Size
- Equal Distribution of Genders
- Aviodance of inbreeding in production colonies
- **Program of regular breed stock migration between subpopulations**

16 | EVERY STEP OF THE WAY

charles river

GENETIC STANDARDIZATION – MINIMIZING GENETIC DRIFT

International Genetic Standardization (IGS) - Inbreds

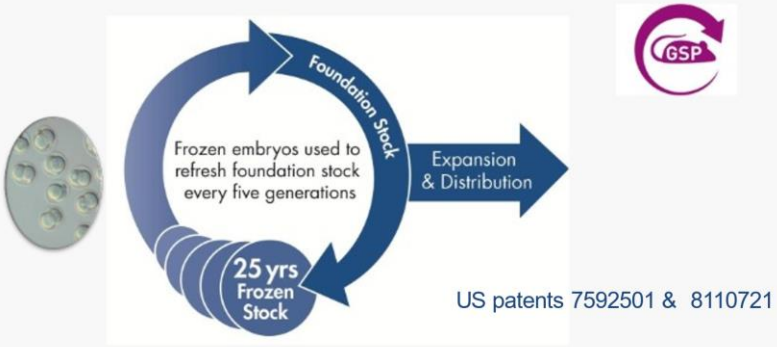


EVERY STEP OF THE WAY

charles river

GENETIC STANDARDIZATION – MINIMIZING GENETIC DRIFT

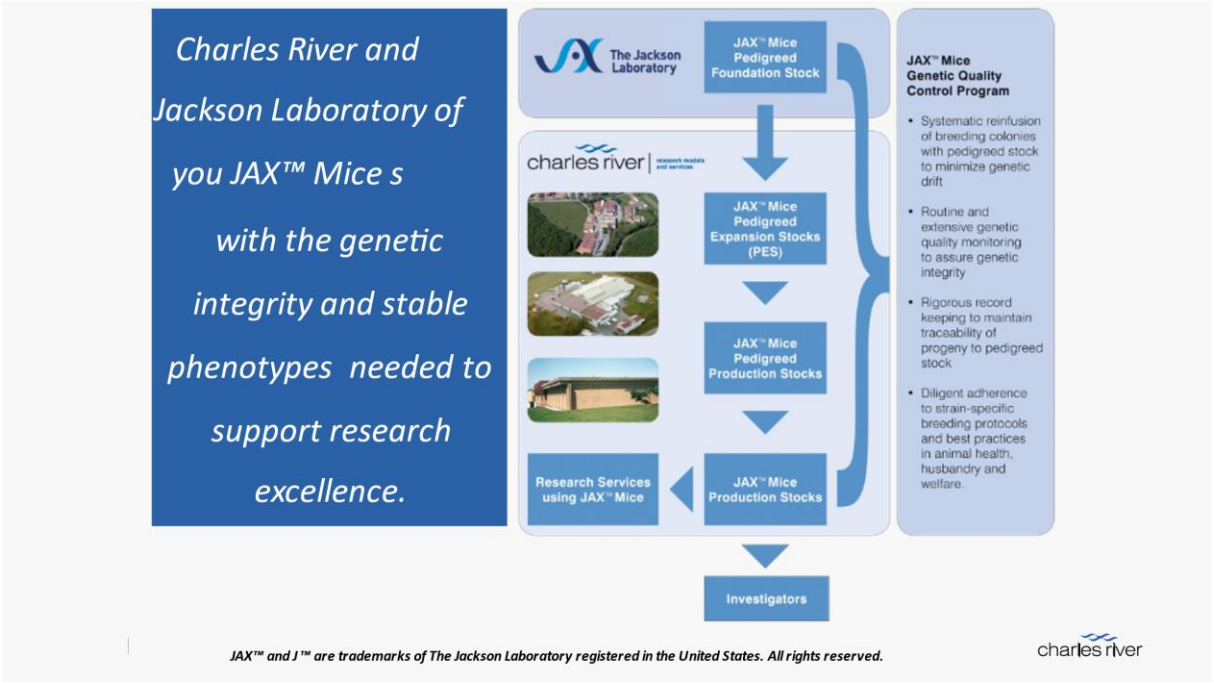
The Jackson Laboratroy Genetic Standardization Program (GSP)

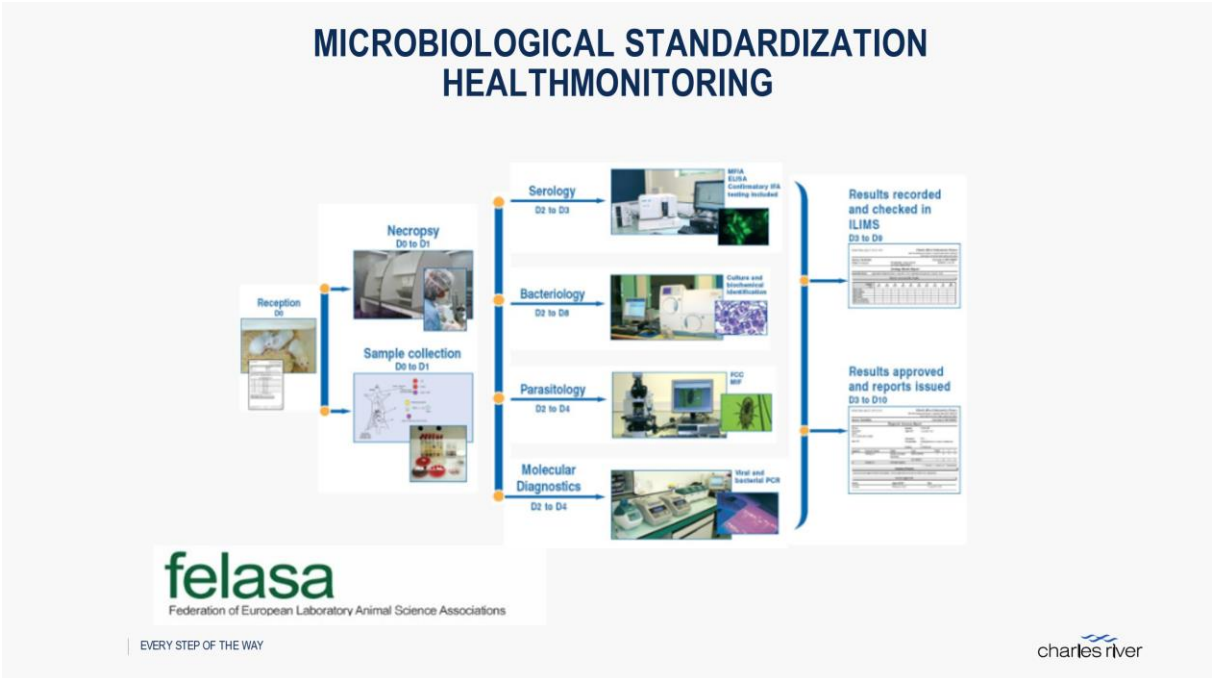


Replace the Foundation Stocks Breeders using cryopreserved embryos at frequent intervals (Genetic Stability Program–GSP)

EVERY STEP OF THE WAY

charles river


Charles River and Jackson Laboratory of you JAX™ Mice s with the genetic integrity and stable phenotypes needed to support research excellence.
The Jackson Laboratory
JAX™ Mice Pedigreed Foundation Stock
charles river
JAX™ Mice Pedigreed Expansion Stocks (PES)
JAX™ Mice Pedigreed Production Stocks
Research Services using JAX™ Mice
JAX™ Mice Production Stocks
Investigators
JAX™ Mice Genetic Quality Control Program
Systematic reinfusion of breeding colonies with pedigreed stock to minimize genetic drift
Routine and extensive genetic quality monitoring to assure genetic integrity
Rigorous record keeping to maintain traceability of progeny to pedigreed stock
Diligent adherence to strain-specific breeding protocols and best practices in animal health, husbandry and welfare.
JAX™ and J ™ are trademarks of The Jackson Laboratory registered in the United States. All rights reserved.
charles river


MICROBIOLOGICAL STANDARDIZATION HEALTHMONITORING
Reception
Necropsy
D0 to D1
Sample collection
D0 to D1
Serology
D2 to D3
Bacteriology
D2 to D8
Parasitology
D2 to D4
Molecular Diagnostics
D2 to D4
Results recorded and checked in ILIMS
D3 to D9
Results approved and reports issued
D3 to D10
felasa
Federation of European Laboratory Animal Science Associations
EVERY STEP OF THE WAY
charles river

MICROBIOLOGICAL STANDARDIZATION HEALTH MONITORING

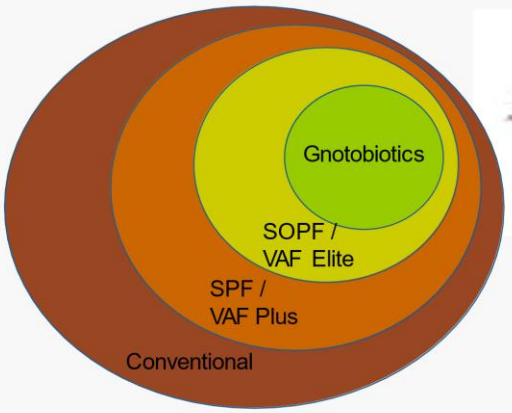
- **Creation of Microbiological Standard**
  - Biosecurity, Housing System
- **Maintenance of Microbiological Standard**
  - Surveillance by Health Monitoring
- **Emergency Plan** = COLONY POLICY FOR POSITIVE RESULT

EVERY STEP OF THE WAY

charles river

MICROBIOLOGICALSTANDARDIZATION

Hygenic Status



SPF = Specified Pathogen Free (VAF in the US)

SOPF = Specified Pathogen and Opportunistic Free (ELITE in the US)

Negative Definitions → definition which germs are ***absent***

EVERY STEP OF THE WAY

charles river

MICROBIOLOGICAL STANDARDIZATION TWO BREEDINGSYSTEMS

barrier rooms
SPF Health status

isolator stations
SOPF Health status
Contract breeding and housing

EVERY STEP OF THE WAY

charles river

EMERGENCY PLAN = COLONY POLICY FOR POSITIVE RESULT

FOR INFORMATION, CRL BREEDING COLONY POLICY FOR POSITIVE RESULT
a = immediate termination
b = planned future recycled of the colony
c = no action except if presence of clinical sign and / or lesion

TESTING SCHEDULE
d = screened quarterly (barrier room) ; quarterly on 1/4 of isolators (isolators)
e = screened quarterly (barrier room) ; quarterly or semi-annually (isolators)
f = screened annually

Germ of Category A:
Immediate stop of breeding / sales→ recycling

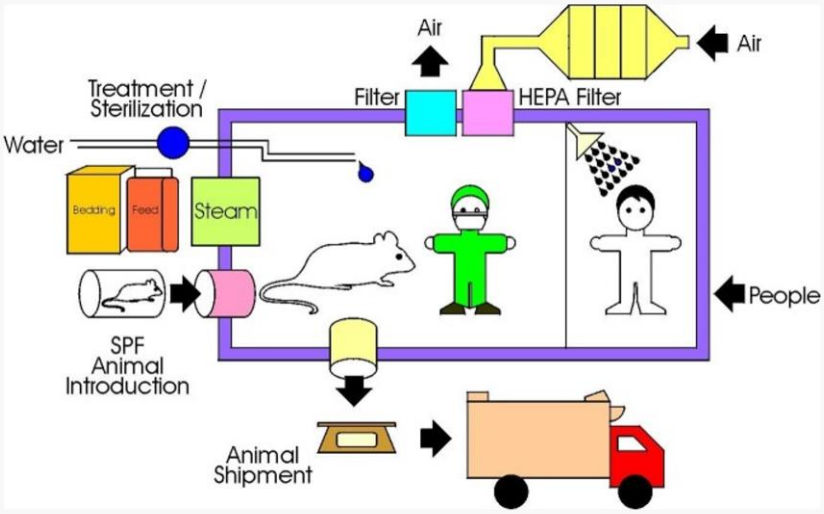
Germ of Category B:
Planned future recycling

Germ of Category C:
No action except if presence of clinical signs and / or lesions

EVERY STEP OF THE WAY

charles river

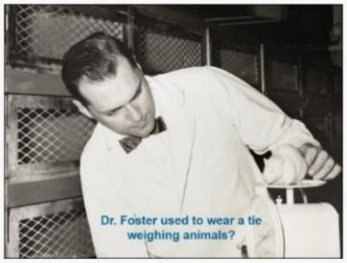
ENVIRONMENTALSTANDARDIZATION BIOSECURITY



EVERY STEP OF THE WAY

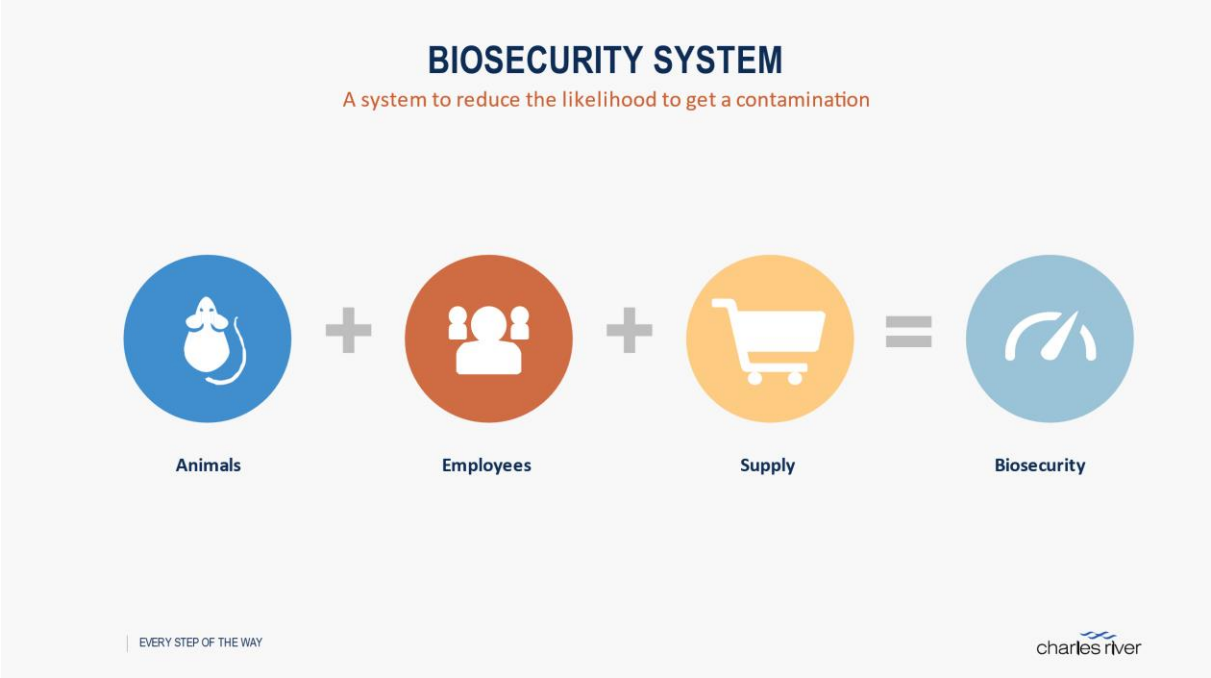
charles river

BIOSECURITY– WORKINGWITH ANIMALSIN THE PAST



EVERY STEP OF THE WAY

charles river


BIOSECURITY SYSTEM
A system to reduce the likelihood to get a contamination
Animals
Employees
Supply
Biosecurity
EVERY STEP OF THE WAY
charles river

ENVIRONMENTAL STANDARDIZATION

Air, Water, Light, Noise,...

- Temperature and Humidity
- Air exchange rates / -Circulation
- HEPA-filtered air
- Drinking water treatment (acidification, chlorination, filtration, UV-treatment)
- Desinfektion / Sterilisation (autoclaving, irradiating, H_2O_2 treatment)
- Choice of cages
- Light cycle (incl. Light intensity)
- Noise level
- Dedicated animal caretakers per barrier

EVERY STEP OF THE WAY

charles river

ENVIRONMENTAL STANDARDIZATION

Diet and Bedding

- Weender analysis for each diet batch
- Quarterly check for toxicological agents (external)
- Certificate of analysis for each bedding batch
- Autoclaving
- **Controlled storage** in Clean-Room, with Stainless steel walls and positive pressure
- **Distribution** through closed high-pressure pipeline system

EVERY STEP OF THE WAY

charles river

ENVIRONMENTAL STANDARDIZATION-PERSONNEL

dedicated Personnel per breeding area / strain

- Regular Health check of the staff
- Private pets policy (no rodents / no animals using rodents as feed (e.g. snakes)
- Entry rules to barriers and isolator rooms (cloth changes / air showers)
- Peronalized helmets (filtered air)
- Annual Biosecurity any Personal Hygiene trainings
- Hygiene-memo
- Entry rules also for visitors (no contact to relevant species within the last 24 hours)

EVERY STEP OF THE WAY

charles river

ENVIRONMENTAL STANDARDIZATION

Enrichment - A balance between animal and researcher needs

Breeders (Permant matings inbred mice):
Paper house and paper nestlets

Breeding females, mothers with litters, pregnant females (intermittend breeding, outbred mice):
Paper nestlets

Animals for sale:
Gnawing sticks

Make Caring Happen

Enrichment Team established on Site in 2015 following Corporate Guideline: Environmental Enrichment in RMS. Team perfoms tests with different enrichment tools on site, consists of members with scientific background and caretakers

EVERY STEP OF THE WAY

charles river

HANDLING STANDARDIZATION

Typical Routines

- Cages changes at least once a week
- Daily controls of all cages
- Daily count of all litters in colonies > 100 breeding females
- 3 times per week count of all litters in colonies < 100 breeding females, and monogamous matings
- Pool weaning at ~ 3 weeks of age, according to age and sex
- Males stay in stable groups

Orders filled – listed by age

EVERY STEP OF THE WAY

charles river

….. WHICH MEANS: SHIPMENT AND CONSEQUENTLY ENVIRONMENTAL CHANGES OCCUR DURING ADOLESCENCE

"(…) exposure to stressors during the juvenile period can exert long-term effects on the brain and behaviour and that these effects differ depending on whether the animals are tested during adolescence or adulthood."

Peleg-Raibstein & Feldon, Psychopharmacology, 2011

"Experiences throughout adolescence also have profound effects on behaviour later in life. Mice exposed to chronic social stress between PND 28-77 showed increased anxiety-like behaviour in adulthood"

Brust et al., Frontiers in Zoology, 2015

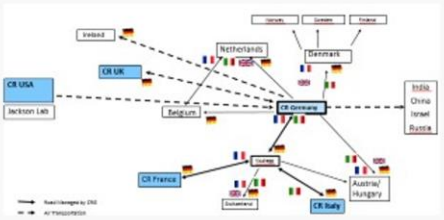
"(…) results suggested thatsecond postnatal week may be the critical period for establishing proper behavioral responses to emotional stress in the adolescent mouse."

Nishio et al., Int J Dev Neurosci., 2006

EVERY STEP OF THE WAY

charles river

TRANSPORT



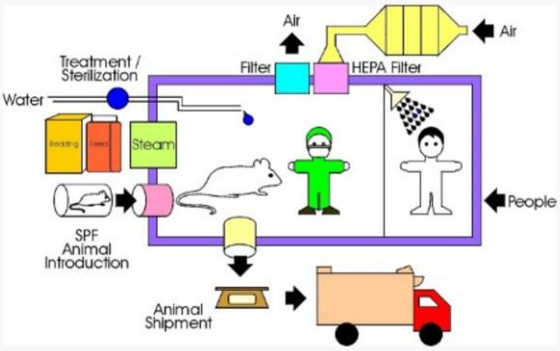
EVERY STEP OF THE WAY

charles river

SUMMARY

Vendors deal with innumerable variables during the ~5% of a rodent's lifespan prior to shipping

Many of the variables are actually very well standardized and controlled, and are unlikely sources of short-term variation



EVERY STEP OF THE WAY

charles river

SUMMARY

Many other potential variables are rarely considered:

- Did I choose the right population model ? (inbred vs. outbred)
- Did I order a given strain always from the same vendor – no two vendors breed same substrains!
- Did I receive animals from same breeding area – Microbiome ? Remember: health monitoring is based on EXCLUSION lists
- Do I need BIOLOGICAL litters or "just" young animals of same age?
- Is my acclimatization period adjusted to my research focus ?

-> Always discuss your specific needs with your vendor. Audit when possible !

EVERY STEP OF THE WAY

charles river

CHARLES RIVER ANIMAL WELFARE: THE HUMANE CARE INITIATIVE

Charles River is aaalac INTERNATIONAL accredited.

AAALAC International is a private, nonprofit organization that promotes the humane treatment of animals in science trough voluntary accreditation and assessment programs.

EVERY STEP OF THE WAY

charles river

CONTACT US

Jutta Davidson, Dipl. Biol.
Research Models Specialist

Charles River Research Models and Services GmbH
Sandhofer Weg 7
D-97633 Sulzfeld

Email:
jutta.davidson@crl.com

Website:
www.criver.com

Phone:
+49 (0)9761 406 -49

charles river

2. National Toxicology Program's Perspective and Use of Historical Control Data

(Chad Blystone, NTP)

NTP National Toxicology Program

National Toxicology Program's Perspective and Use of Historical Control Data

Chad Blystone
Division of National Toxicology Program
National Institute of Environmental Health Sciences

May 3, 2022

International workshop on how to report, use and interpret historical control data in (eco)toxicity studies

Overview

- NTP Background and Use of Historical Control Database (HCD) in animal studies
- Provide examples of Changes and Challenges to our HCD
  - Animal Model
  - Exposure Paradigm
  - Pathology
  - Reduction in animal studies

NTP Historical Control Use

- NTP makes "confidence calls" on evidence of neoplastic and now Developmental and Reproductive Toxicity (DART) responses in rat and mouse studies:
  - Clear Evidence
  - Some Evidence
  - Equivocal Evidence

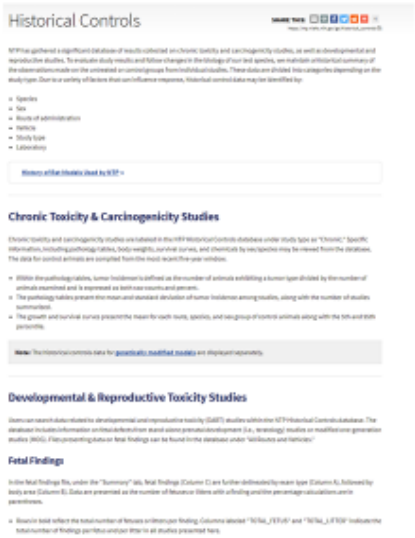
  } Historical Control influence likely the greatest here
  - No Evidence
- An animal model historical control will be influenced by genetics, route of exposure, nomenclature, study design, diet, etc. We attempt to keep these consistent as possible.

NTP Historical Control Use

- The concurrent control is considered the most important for evaluation of a study's findings
- The historical control database provides context and can increase or decrease confidence in a call
  - Increased confidence when identifying a response in a rare neoplasm
  - Lower confidence in variable high background rate neoplasm (e.g. when concurrent control lower than historical range)

NTP Historical Control Background

- The NTP has reported over 600 cancer bioassays (2-yr studies) in model species of rats and mice
- Prenatal toxicity studies initiated in the 2010's have also been added in the last 5 years
- Historical control data is collected and published on NTP website:
  - https://ntp.niehs.nih.gov/data/controls/index.html


Historical Controls

Chronic Toxicity & Carcinogenicity Studies

Developmental & Reproductive Toxicity Studies

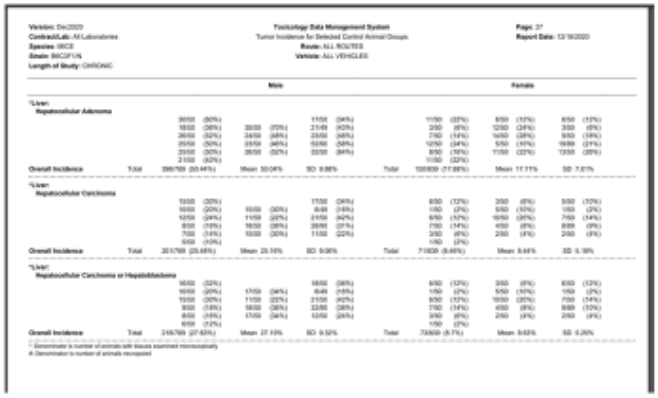
Fetal Findings

NTP Historical Background

- Historical data are collated and published in a rolling 5-year window based on study start date
  - Incidence rates of neoplasms, growth, and survival data
  - Data are provided by
    - Species, strain/stock, and sex
    - Route of exposure
    - Diet

NTP Historical Control

- Neoplastic findings are provided by organ system
  - Individual study findings, total, mean, and standard deviation are reported
  - Combinations of neoplasms are also provided (e.g. adenomas and/or carcinomas)



NTP Historical Control

Standardization

- NTP Pathology Review of 2-Year Chronic Studies
  - NTP Pathology conducts a review of each carcinogenicity study
  - QA pathologist review and report
  - Inconsistencies (e.g. nomenclature, diagnosis) are resolved, via a Pathology Working Group
- Together this strengthens the NTP Historical Control database

Research Summary

The Pathology Evaluation and Peer Review (PEPR) Group provides support to NIEHS, including studies conducted by the Division of the National Toxicology Program (DNTP). The group participates in and oversees pathology-related issues for rodent toxicology studies and other studies, including study design and management, data analysis, peer review, and reporting.

National Toxicology Program

The PEPR Group provides pathology expertise to evaluate findings, interpret those findings, and works with other staff to determine the toxicologic potential of substances studied by the DNTP. Staff within the Data Coordination Unit (DCU) receive, track, and assist with finalizing the pathology data. Many of these studies are conducted on behalf of the National Toxicology Program (NTP) and include studies to evaluate the effects of environmental substances on general toxicity, carcinogenicity, neurodevelopment, reproduction and development, and the immune, nervous systems, and cardiovascular systems. These studies are reported in peer reviewed NTP technical reports and journal manuscripts.

The PEPR Group also works with NIEHS investigators to provide pathology expertise regarding study design and special techniques; data evaluation, management, and interpretation; and interpretation and reporting of findings. These include studies using rodent models of disease (such as genetically modified models) and/or in vitro studies to identify environmental hazards or investigate the mechanisms involved in human diseases, toxicity, or carcinogenesis. This research has resulted in numerous publications that elucidated the mechanisms of environmental carcinogenesis.

https://www.niehs.nih.gov/research/atniehs/labs/lep/ntp-path/index.cfm

Use Example

2,3 Butanedione (TR-593) – Wistar Han Rats

Nose	Chamber Control	12.5 ppm	25 ppm	50 ppm
Squamous Cell Papilloma	0/50	0/50	0/50	1/50
Squamous Cell Carcinoma	0/50	0/50	0/50	3/50
Papilloma or Carcinoma	0/50*	0/50	0/50	4/50*
Inhalation HC	0/200			
All Routes HC	0/349			

Less than 0.5% in F344 and Wistar rats

Some Evidence of Carcinogenic Activity

Use Example

Mouse Hepatocellular Neoplasms

- Well known high background rate for the mouse model B6C3F1/N

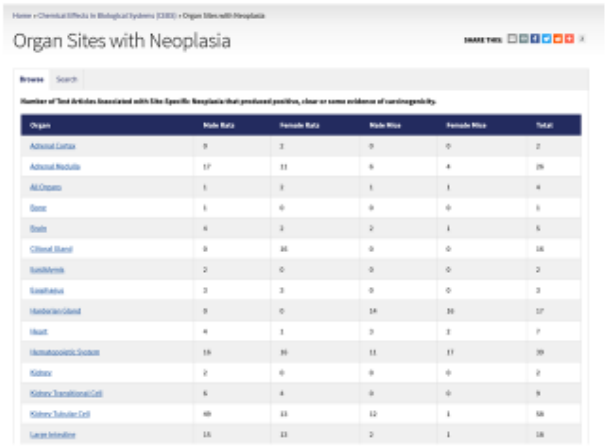
2-Butoxyethanol	Chamber Control	62.5 ppm	125 ppm	250 ppm
Adenomas	22/50	18/50	18/50	17/50
Carcinomas	10/50	11/50	16/50	21/50**
Adenoma or Carcinoma	30/50	24/50	31/50	30/50

	Avg	Low	High
Carcinomas	13/50	6/50	24/50
Adenoma or Carcinoma	26/50	10/50	43/50

Equivocal Evidence of Carcinogenic Activity

Sites and Associations

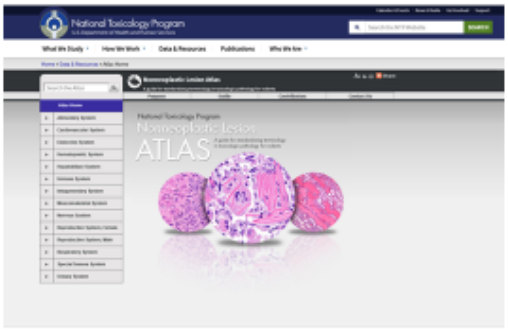
- Beyond the historical control database, NTP provides a database for the neoplastic responses observed in studies
- The NTP organ site database collates responses within tissue across the 600+ chemicals that have been evaluated



https://cebs.niehs.nih.gov/organsites/

Non-neoplastic lesions

- NTP does not provide a historical control for non-neoplastic lesions
- However, NTP provides a nonneoplastic lesion atlas
- The Atlas provides diagnostic guidelines for microscopic nonneoplastic lesions in rats and mice
- Overall goal is to ensure the diagnostic consistency of findings in NTP studies and provide photomicrograph examples and descriptions of findings that other pathologists/organizations may use as a reference



https://ntp.niehs.nih.gov/nnl/index.htm

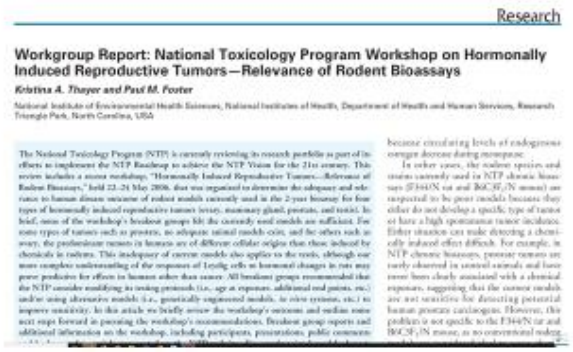
Developmental Toxicity

- NTP initiated a historical control database for DART studies
- Currently consists of fetal morphology findings and litter data from teratology studies
- Plan to expand with more teratology studies and other DART endpoints

Note: Reported by lab

NTP Changes and Challenges

- Rat stock changes
  - NTP evaluated the ability to detect hormone induced reproductive tumors
  - Due to various concerns of the F344/N rat, NTP switched rat stocks
    - Wistar Han (2008-2012 = 7 studies)
    - HSD:SD Rat (2014-present = 8+ studies)
  - Impacted historical control database:
    - Starting from scratch in rats



Research

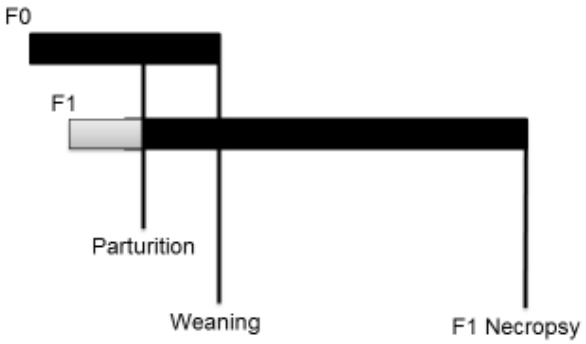
Workgroup Report: National Toxicology Program Workshop on Hormonally Induced Reproductive Tumors—Relevance of Rodent Bioassays

Kristina A. Thayer and Paul M. Foster

https://pubmed.ncbi.nlm.nih.gov/17805427/

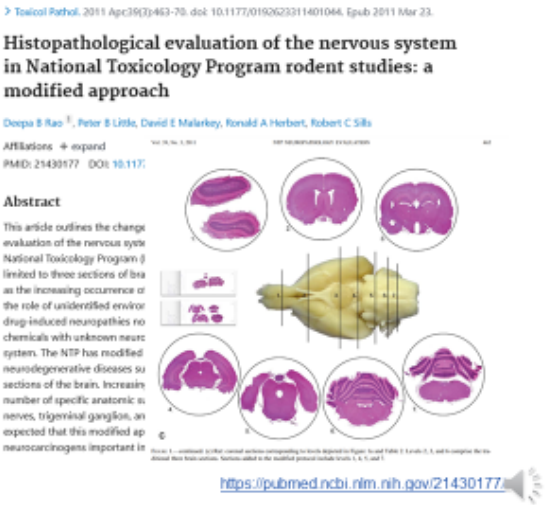
NTP Changes and Challenges

- Change in Exposure Paradigm
  - In utero and lactational exposure often not included in traditional carcinogenicity studies
  - Concern arose that not including potentially sensitive life stage results in lower detection rate of certain neoplasms or chronic toxicity
  - NTP rat studies now often include (ex. Inhalation studies) exposure during this period
  - In practice NTP does not differentiate exposure paradigms (i.e. perinatal vs nonperinatal) with Historical Control use



NTP Changes and Challenges

- Nervous System
  - NTP evaluation of the nervous system changed circa 2011
  - Four additional sections of the brain included = 3 three sections now 7 sections
  - Increases probability of detecting low incidence findings
  - Splintered the NTP historical control database:
    - Studies with older evaluation separated from studies with newer evaluation


Toxicol Pathol. 2011 Apr;39(3):463-70. doi: 10.1177/0192623311401046. Epub 2011 Mar 23.

Histopathological evaluation of the nervous system in National Toxicology Program rodent studies: a modified approach

Deepa B Rao [1], Peter B Little, David E Malarkey, Ronald A Herbert, Robert C Sills

Affiliations + expand

PMID: 21430177 DOI: 10.117

Abstract

This article outlines the change evaluation of the nervous syste National Toxicology Program (limited to three sections of bra as the increasing occurrence o the role of unidentified environ drug-induced neuropathies no chemicals with unknown neuro system. The NTP has modified neurodegenerative diseases su sections of the brain. Increasin number of specific anatomic u nerves, trigeminal ganglion, an expected that this modified ap neurocarcinogens important in

https://pubmed.ncbi.nlm.nih.gov/21430177/

NTP Changes and Challenges

- Uterine evaluation
  - Longitudinal sectioning of uterine horns now standard practice for carcinogenicity studies
  - Increases sensitivity for detecting neoplastic and nonneoplastic lesions

Toxicol Pathol. 2020 Aug;48(6):747-755. doi: 10.1177/0192623320948843. Epub 2020 Aug 20.

The Assessment of Longitudinal Sections of Rat Female Reproductive Tissues for NTP 2-Year Toxicity and Carcinogenicity Studies

Susan A Elmore [1], Chad Blystone [2], Beth A Lubeck [1], Shawn P Harris [3], Crystal L Johnson [4]

Affiliations + expand

PMID: 32815403 PMCID: PMC7486810 DOI: 10.1177/0192623320948843

Free PMC article

Abstract

The National Toxicology Program (NTP) now uses an extended longitudinal sectioning protocol for the uterus to better evaluate female rodent reproductive tract toxicity for all developmental and reproductive toxicology and 2-year toxicity and carcinogenicity bioassays. The previous protocol for toxicity/carcinogenicity studies involved 1 cross section midway through each uterine horn and collection of uterine cervix and vagina only if gross lesions were present. Here we compare the histological findings of the original cross sections with the additional longitudinal sections of residual uterine tissues of 7 chronic NTP rat bioassays. The goal of this study was to determine whether there might be any advantages to examining additional uterine tissue. The longitudinal protocol allowed for 10 to 20 times more uterine tissue for evaluation. Results indicate that the potential advantages of a more complete evaluation of female reproductive tract tissue include increased detection of reproductive targets, increased detection of neoplastic and nonneoplastic lesions, improved detection of tissue origin of neoplasms, less reliance on gross identification of lesions, improved accuracy in the application of severity grades, and increased detection of preneoplastic or subtle lesions.

Keywords: extended evaluation; female; longitudinal section; reproductive tract; rodent; trimming protocol; uterus.

NTP Changes and Challenges

- Retrospective evaluation of change in methodology found
  - Increase in detection of neoplasms and non-neoplastic lesions, often at lower exposures
  - Atypical hyperplasia detected at much greater rates
  - Influence on confidence of response varied from study to study

Figure 1. Example of cervix and vagina embedded as 1 piece and sectioned longitudinally. Hematoxylin and eosin.

Table 2. Uterine Findings From Original and Extended Evaluations.[a]

Chemical	Sex	Adenoma	Adenocarcinoma	MMMT	Stromal tumors combined[b]	Atypical hyperplasia	Cystic endometrial hyperplasia
DE-71	Mice	1, 0, 1, 0[c]	1, 1, 3, 2	1, 0, 0, 0	8, 7, 10, 6	3, 1, 2, 0	15, 9, 17, 14
		(1, 1, 3, 1)[d]	(2, 1, 4, 4)	(1, 0, 0, 0)	(5, 13, 14, 10)	(5, 6, 7, 3)	(24, 40, 34, 26)
GTE	Mice	3, 0, 0, 1	3, 0, 5, 4	0, 0, 0, 0	2, 1, 5, 4	3, 0, 3, 0	12, 14, 15, 13
		(3, 1, 1, 2)	(8, 13, 10, 8)	(0, 0, 0, 0)	(8, 2, 10, 5)	(2, 4, 4, 2)	(26, 36, 32, 17)
TBBPA	Mice	0, 0, 2, 4	2, 2, 8, 9	0, 4, 0, 2	2, 7, 2, 2	3, 0, 3, 0	8, 13, 11, 18
		(3, 2, 4, 4)	(4, 12, 11, 16)	(0, 4, 0, 2)	(5, 12, 9, 9)	(2, 12, 11, 13)	(24, 31, 30, 32)
CHETAN 3800	Mice	3, 0, 0, 0	1, 1, 1, 2	0, 0, 0, 1	4, 4, 4, 5	3, 0, 0, 0	11, 4, 9, 5
		(3, 0, 0, 0)	(1, 4, 5, 5)	(0, 0, 0, 1)	(11, 10, 8, 6)	(13, 10, 8, 12)	(40, 42, 41, 43)
Diocetyl	Mice	3, 0, 0, 0	0, 2, 0, 2	0, 0, 0, 0	6, 8, 7, 9	3, 0, 0, 0	11, 5, 14, 17
		(3, 0, 0, 2)	(2, 4, 2, 3)	(0, 0, 0, 0)	(9, 15, 12, 15)	(3, 1, 0, 0)	(24, 16, 17, 25)
DIC	SD	0, 0, 0, 0	0, 1, 4, 4	0, 0, 0, 0	12, 10, 15, 9	3, 0, 0, 0	29, 28, 29, 21
		(0, 2, 1, 1)	(5, 4, 13, 10)	(0, 0, 0, 0)	(12, 11, 16, 13)	(15, 14, 21, 16)	(34, 37, 34, 32)
PFOA NP[e]	SD	2, 1, 3	0, 1, 4	0, 0, 0	6, 4, 6	3, 0, 3	28, 22, 18
		(1, 1, 3)	(1, 3, 8)	(0, 0, 0)	(16, 11, 13)	(3, 4, 3)	(38, 37, 37)
PFOA P	SD	3, 0, 3	0, 2, 1	0, 0, 0	6, 4, 3	3, 0, 3	32, 31, 28
		(3, 0, 3)	(1, 3, 5)	(0, 0, 0)	(16, 13, 17)	(3, 7, 3)	(38, 36, 37)

Abbreviations: CHETAN 3800, manufacturing fluid; dioctyl, 2,3-butanediol; DE-71, pentabromodiphenyl oxide mixture; GTE, green tea extract; DIC, indole-3-carbinol; MMMT, malignant mixed Müllerian tumor; PFOA, perfluorooctanoic acid; SD, Hsd:Sprague Dawley SD; TBBPA, tetrabromobisphenol A; WH, Wistar Han (Crl: WI (Han)).

[a] 50 control animals per group.

[b] Stromal tumors (includes multiples) = combined stromal polyp, stromal sarcoma, leiomyosarcoma, or malignant schwannoma.

[c] x, x, x, x = incidences from original evaluation (control, low-, mid-, and high-dose groups).

[d] (x, x, x, x) = incidences from combined original and extended evaluation groups.

[e] The PFOA study includes both PFOA NP (nonperinatal exposure) and PFOA P (perinatal exposure). The 2 studies shared the same control group with high- and low-exposure groups (no mid-exposure group).

NTP Changes and Challenges

- Statistical Analysis
  - In an attempt to use the large NTP historical database, formal statistics were generated to test a study's response against the HCD.
  - This could strengthen the confidence in a call for rare tumors and provide higher confidence when background rate is variable
  - Mixed response regarding its use.
    - Concerns for over reliance
    - Positive results hard to believe at times
  - No longer published on stats tables

J Am Stat Assoc. 2007 Jan 1;102(480):1212-1220. doi: 10.1198/016214506000001356.

Incorporating Historical Control Data When Comparing Tumor Incidence Rates

Shyamal D Peddada, Gregg E Dinse, Grace E Kissling

PMID: 20396669 PMCID: PMC2853781 DOI: 10.1198/016214506000001356
Free PMC article

Abstract

Animal carcinogenicity studies, such as those conducted by the U.S. National Toxicology Program (NTP), focus on detecting trends in tumor rates across dose groups. Over time, the NTP has compiled vast amounts of data on tumors in control animals. Currently, this information is used informally, without the benefit of statistical tests for carcinogenicity that directly incorporate historical data on control animals. This article proposes a survival-adjusted test for detecting dose-related trends in tumor incidence rates, which incorporates data on historical control rates and formally accounts for variation in these rates among studies. An extensive simulation, based on a wide range of realistic situations, demonstrates that the proposed test performs well in comparison to the current NTP test, which does not incorporate historical control data. In particular, our test can aid in interpreting the occurrence of a few tumors in treated animals that are rarely seen in controls. One such example, which motivates our work, concerns the analysis of histiocytic sarcoma in the NTP's 2-year cancer bioassay of benzophenone. Whereas the occurrence of three histiocytic sarcomas in female rats was not significant according to the current NTP testing procedure ($P = 0.074$), it was highly significant ($P = 0.004$) when control data from six recent historical studies were included and our test was applied to the combined data.

https://pubmed.ncbi.nlm.nih.gov/20396669/

NTP Changes and Challenges

- Fewer animal studies being conducted
  - Decreases historical control size and utility
  - Are there viable work arounds?
    - Expanding 5-year window
    - More reliance on non-NTP databases
    - Increasing number of control groups within a study

Progression and Standardization

- Historical control an invaluable resource for evaluating current datasets
  - Requires consistency in methodology, nomenclature, and use
- Progression in methods to increase sensitivity and standardization to increase consistency leads to fracturing of historical control database
  - Changes need to be planned ahead of time
- Declining use of animal studies presents a challenge for maintaining a healthy historical control database
  - Method changes and declining use creates a conflict

3. Use of HCD on Pharmaceutical Toxicology Studies at Charles River Edinburgh

(Aidan McGuire, Charles River Laboratories)

Use of HCD data on Pharmaceutical Toxicology Studies at Charles River Edinburgh

Aidan McGuire Director Toxicology Scientific Support

Unlike agrochemical (crops and chemicals) and TAS (Target Animal Safety) studies there is no requirement to add HCD ranges to pharmaceutical toxicology reports. Therefore, Charles River does not add these to their toxicology studies as standard.

1 | EVERY STEP OF THE WAY

GENERATION OF HCD DATA AT CHARLES RIVER EDINBURGH, UK

Current Standard HCD Data Streams and Format

Clinical Pathology Data (Electronic, Provantis)
Histopathology Incidence Data (Electronic, Provantis)
Macroscopic Incidence Data (Electronic, Provantis, theoretically)
Organ Weights (Electronic, Provantis)
Survival rates (Electronic, Provantis, by study duration, species, strain, route of administration, sex, Sponsor and study type)
Reproductive Parameters (Manual)
Neurotoxicity Observations (Manual)
Safety Pharmacological data (Manual)

Compliance of Electronic Provantis Data

- Non-Regulatory and GLP mixed but can be separated for GLP only data. Pharmaceutical Tox does not require GLP HCD data, only to be of a suitable reproducible quality.

2 | EVERY STEP OF THE WAY

charles river

Truth is we rarely use HCD data in Pharmaceutical Toxicology

Reasons

All comparisons are made with contemporaneous control group as per current guidance

Statistical analysis is always only performed against the control group for this reason

A secondary comparison is available for non rodent species with pretrial data for Clinical Pathology

Only really use HCD data to explain high individual values, usually in the control animals that cause artificial statistical significance.

Normally also have other correlations to confirm a change was chance and not test item related.

Exceptions, Carcinogenicity incidence data, Incidence data specific tumours.

3 | EVERY STEP OF THE WAY

Rational for HCD Electronic Clinical Pathology Data Collection

Let's take Clinical Pathology Data as an example

Data taken from Control animals only.

Data must be collected at the same site as the data it is compared with (housing)

Same Equipment (analytical machines)

Same kits for the analysis

Same substrate (Plasma Chem v Serum Chem)

Data must be contemporaneous (searchable by date and study number)

Fasting status should be known

Use of anaesthesia must be recorded

Data can be broken down by age and animal supplier

Only Finalised study data is used

4 | EVERY STEP OF THE WAY

Output

How we report our Clin Path Data

Supplier (if requested)
Strain
Age range
Parameter quoted

N	95% confidence limits
Mean	Min
Median	Max
SD	

We usually only quote or use the 95% confidence limits in our reports if required and keep a printout in the raw data as when the system is updated it alters the values.

5 | EVERY STEP OF THE WAY

Dog data age ranges

Age break down set by our Clinical pathologists

Less than 6 months
6-7 months
8-9 months
10-12 months
More than 12 months

6 | EVERY STEP OF THE WAY

Dog Example

6-7 month dogs Day -12 Pretrial Males Beagles

	Statistic	ALP (iu/L)
Control Group 1	N	4
	Mean	78.8
High Dose Group 4	N	4
	Mean	117.7*
	* = Statistical significance	≤ 0.05

	Group 1 ALP (iu/L)	Group 4 ALP (iu/L)
Individual values		
1	72	118
2	82	101
3	74	150
4	89	101

Group 4 values clearly higher than Control
Pretrial so clearly not treatment related
Also observed in Week -1 bleed
HCD n=294 95% confidence limits = 64-180 iu/L

Clearly as noted during pretrial we would not normally use HCD data but if we did all values were within the normal range. So these animals were considered normal and this change was just due to chance.
Change no longer evident when on study.

7 | EVERY STEP OF THE WAY

charles river

SUMMARY OF RANDOMLY SELECTED DOG CLINICAL PATHOLOGY PARAMETERS

Males Beagles only

Supplier	Statistic	Glu (mmol/L)	Trig (mmol/L)	ALT (iu/L)	AST (iu/L)	CPK (iu/L)	Creatinine (umol/L)	Phosphorus (mmol/L)	Potassium (mmol/L)
All Dogs	N	294	294	294	294	174	294	294	294
	Mean	5.78	0.4	31.2	34	306	48.6	1.82	4.36
	95% Confidence	4.85-6.59	0.23-0.62	19-50	23-49	178-691	34-65	1.41-2.29	3.8-4.9
Most Popular	N	140	140	140	140	98	140	140	140
	Mean	5.74	0.41	30.1	33.8	307.3	49	1.81	4.41
	95% Confidence	4.72-6.54	0.23-0.65	19-42	24-52	181-703	34-67	1.42-2.31	3.9-5

6-7 Month all dogs
6-7 Month most popular supplier at Edi

Conclusion: Supplier makes little difference to dog HCD data

8 | EVERY STEP OF THE WAY

charles river

Housing

Reducing stress

Work done in house by CRL Edinburgh has shown that gang housing of NHPs in European housing significantly modifies the absolute number of lymphocyte subsets when even compared with linked cage housed cynomolgus monkeys.

9 | EVERY STEP OF THE WAY

SUMMARY OF CLINICAL PATHOLOGY PARAMETERS OF FASTED V UNFASTED RATS AT CHARLES RIVER EDINBURGH, UK

Clinical Chemistry Sprague-Dawley ca 10 Weeks old

Yellow Finding correlated with Comparison of HCD data

Red Statistical Significance

Food Status	Statistic	Glu (mmol/L)	Trig (mmol/L)	ALT (iu/L)	AST (iu/L)	CPK (iu/L)	Creatinine (umol/L)	Phosphorus (mmol/L)	Potassium (mmol/L)
Fasted Plasma Chemistry	N	20	19	19	18	18	19	17	18
	Adjusted Mean	8.65	0.75	60	86	238	28	2.54	4.2
	Std Error (Mean)	0.28	0.09	2	3	-	1	0.05	0.1
Not Fasted Plasma Chemistry	N	20	20	20	19	19	20	19	19
	Adjusted Mean	11.55	1.69	79	88	227	25	2.01	4.6
	Std Error (Mean)	0.28	0.08	2	3	-	1	0.4	0.1
	Prob	0.001	0.001	0.001	0.61	0.54	0.001	0.001	0.001
Serum Chemistry	xFold in nonFasted	1.33	2.25	1.32	1.02	0.95	0.89	0.79	1.09

Fasting rats causes expected changes to glucose and triglyceride levels and minor changes to liver enzymes (blood flow) and electrolytes (reduced water intake in fasted rats)

CPK difference likely to be due to substrate (serum v Plasma chemistry)

No evidence to suggest the data from fasted animals was less variable than that from non-fasted animals

10 | EVERY STEP OF THE WAY

charles river

SUMMARY OF GLUCOSE LEVELS SAMPLED WITH OR WITHOUT ISOFLURANE AT CHARLES RIVER EDINBURGH, UK

Clinical Chemistry Han Wistar

Food Status	Statistic	Glu (mmol/L)
No Isoflurane	N	20
	Mean	6.609
With Isoflurane	N	20
	Mean	10.623
	xFold in nonFasted	1.61

Isoflurane anaesthesia has a greater impact on glucose levels than fasting

11 | EVERY STEP OF THE WAY

charles river

4. Challenges in Using the HCD: a methodological perspective

(Laura Martino, EFSA)

Workshop on how to report, use and interpret historical control data

3rd – 5th May 2022

Challenges and opportunities in using the HCD:

a methodological perspective

Laura Martino

Senior statistician

MESE Unit

Trusted science for safe food

efsa European Food Safety Authority

HIGHLIGHTs

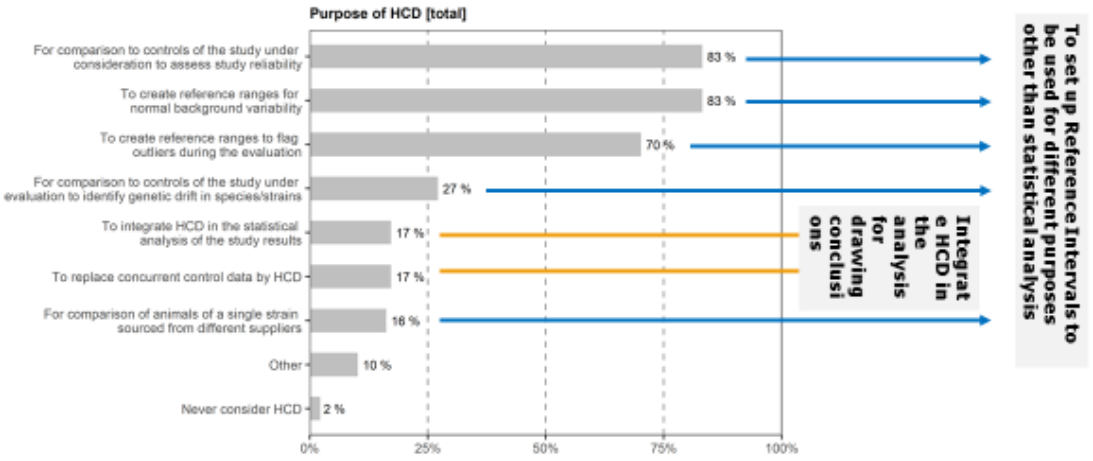
efsa European Food Safety Authority

- Possible uses of HCD (based on the EFSA survey)
  - HCD to establish Reference Intervals
  - Integrate HCD in the statistical analysis to draw conclusions
- Minimal requirements for using HCD
- Importance of planning upfront: need for a protocol
- Open access historical control database

2

Results of the survey: purpose of the HCD

efsa European Food Safety Authority



3

efsa European Food Safety Authority

Establishing Reference Intervals from HCD

4

(ASVCP) GD for de novo reference intervals in veterinary species

efsa European Food Safety Authority

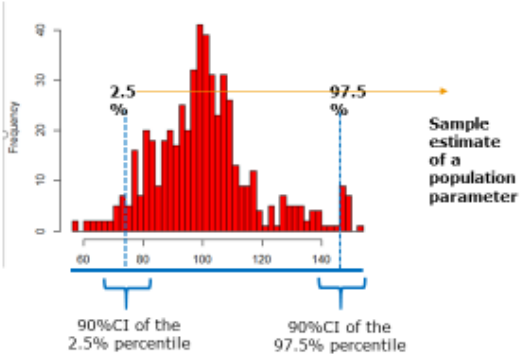
Establishing HCD-based Reference Intervals:

Interval reflecting the intra- and inter-individual variability of the endpoint measurements in individuals belonging to a healthy population

Typical way to derive Reference Intervals is to set a range covering 95% of the values measured on individuals representing a random sample from a healthy reference population. Since extremes of the 95% range are based on a random sample, a CI reflecting the sampling uncertainty on the extreme centile estimates is recommended

Steps

- Reference individuals
- Reference values
- Reference distribution
- Reference limits (95% coverage)
- Reference intervals



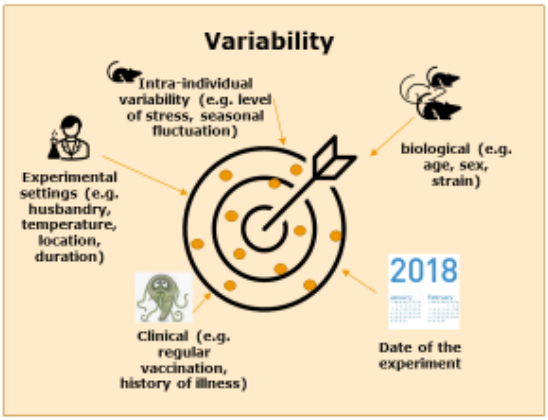
5

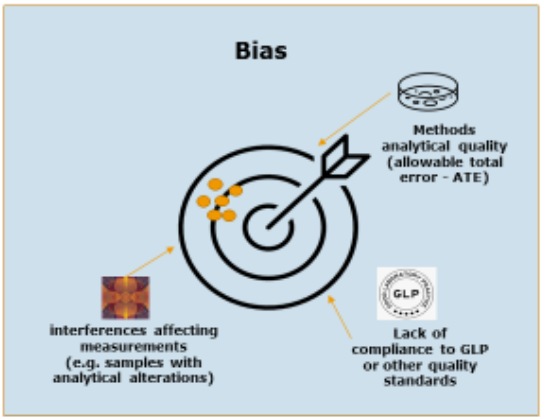
Challenges for the set of Reference Intervals

efsa European Food Safety Authority

Sources of:

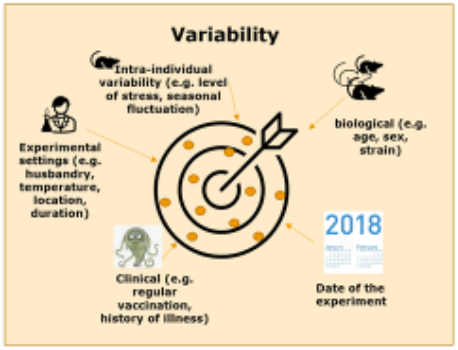
- ***Variability*** -> It includes intra/inter-individual variability, variability across studies
- ***Bias*** -> It is related to the experimental setting and reliability of the analytical measures





Variability

efsa European Food Safety Authority



Factors introducing variability can affect comparability of the HCD to the concurrent control
Need to account for these factors

Possible strategies:

- Subgrouping RI according to factors that are sources of variability
- Sensitivity analysis (assess influence of the individuals/studies or subgroups on RIs)
- Use of weights (0-1) reflecting similarity of the individuals from HCD and the concurrent controls

7

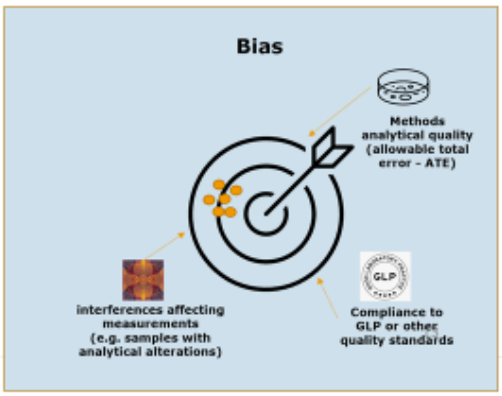
Bias

efsa European Food Safety Authority

Factors introducing bias can affect RIs accuracy & lead to wrong conclusions

Possible strategies:

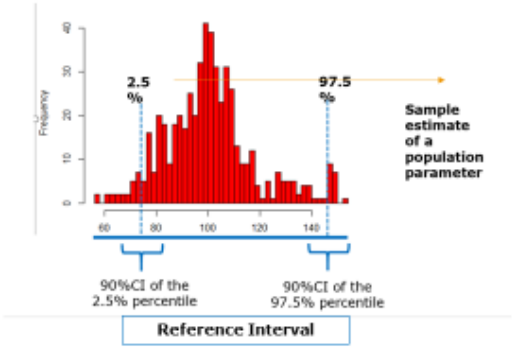
- Sensitivity analysis (assess influence of the individuals or subgroups on RIs)
- Careful removal of the individuals potentially affected from bias



8

How many individuals are necessary

efsa European Food Safety Authority



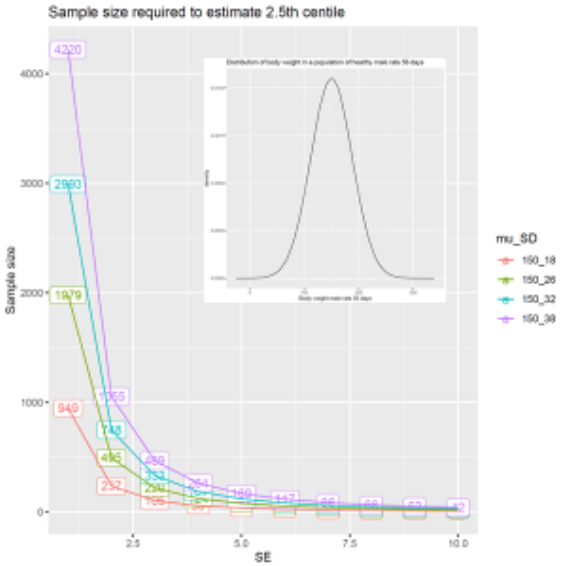
- Estimate of extreme centiles is a crucial issue
- Sample size required to estimate extreme centiles with a pre-defined level of precision (SE) is greater than the one necessary to estimate median (imprecision increases with the centile's distance from the median)

Important to get sufficient precision for the estimate of the extreme centiles to reduce sampling uncertainty to an acceptable level

9

2.5th centile estimate: required sampe size – illustrative example

efsa European Food Safety Authority



CI width: 3.5-45

SE	n	mu_SD P2.5=115	n	mu_SD P2.5=99	n	mu_SD P2.5=87	n	mu_SD P2.5=75
1	949	150_18	1977	150_26	2993	150_32	4220	150_38
2	237	150_18	494	150_26	748	150_32	1056	150_38
3	105	150_18	220	150_26	333	150_32	469	150_38
4	59	150_18	124	150_26	187	150_32	264	150_38
5	38	150_18	79	150_26	120	150_32	169	150_38
6	26	150_18	55	150_26	83	150_32	117	150_38
7	19	150_18	40	150_26	61	150_32	86	150_38
8	15	150_18	31	150_26	47	150_32	66	150_38
9	12	150_18	24	150_26	37	150_32	52	150_38
10	9	150_18	20	150_26	30	150_32	42	150_38

No one solution fits all. Requested sample size is affected by:
- Variability of the outcome in the healthy population
- Desired precision

Cole (2021) Sample size and sample composition for constructing growth reference centiles. Statistical Methods in Medical Research 2021, Vol. 30(2) 488–507

Using Reference Intervals from HCD

efsa European Food Safety Authority

✓Compare the control value to RI to establish whether experiment is reliable

✓Use the HCD distribution to identify potential outliers

Critical issues

- comparability of the control individuals to the HCD (HCD and controls belong to the same population)
- Availability of HCD at individual level or at least reporting quantiles of the distribution

11

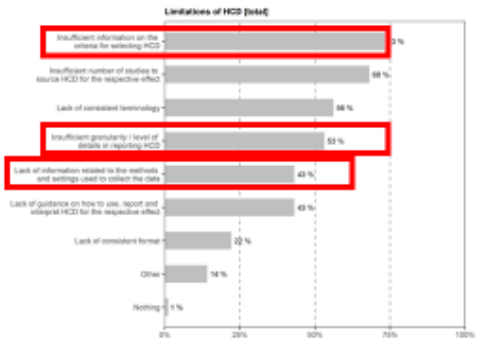
Which descriptive statistics are the minimum requirements

efsa European Food Safety Authority

Ideally HCD should be provided at the level of individual units complemented by metadata to allow assessing influence of possible sources of variability and bias

If individual data are not reported, minimum requirements could be:

- Mean, standard deviation, minimum and maximum of the HCD distribution overall and by subgroups (e.g. age, sex, strain)
- Median, quartiles and extreme centiles (e.g. 2.5% and 97.5%) of the HCD



Why these statistics are necessary

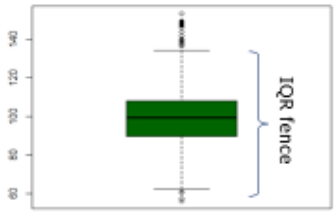
- to assess distribution of the HCD and comparability with concurrent controls
- to identify outliers in the concurrent controls (methods of detection frequently rely on specific centiles of the distribution e.g. Horn's algorithm use the Tukey's interquartile fences (Q1-1.5*IQR); (Q3+1.5*IQR))

12

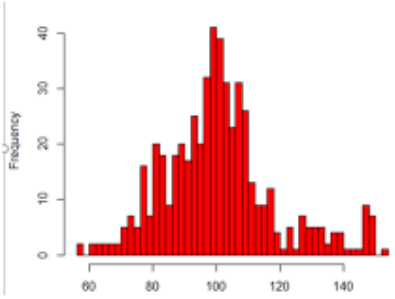
Graphical representation of the HCD

efsa European Food Safety Authority

- Graphical presentation of the HCD are useful to
  - describe the **distribution of the data**
  - identify **potential outliers i.e.** individuals that are not representative of the underlying population (i.e. healthy individuals) and can jeopardise the comparability with concurrent control, introducing a bias in the conclusion



Body weight

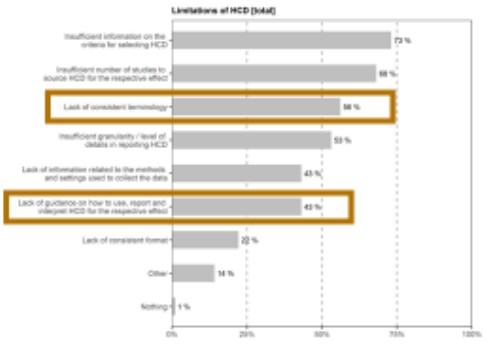


Centiles								
0	0.05	0.1	0.25	0.5	0.75	0.9	0.95	1
56.3	75.18	80.1	89.6	99.5	108	124.5	136.8	153.1

Need for a more harmonised approach for HCD

efsa European Food Safety Authority

- GD is needed on the use, report and interpretation of HCD
- Harmonised terminology is needed to make HCD more usable
- Large set of studies to source HCD is needed



14

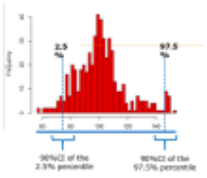
Integrate HCD in the statistical analysis to draw conclusions

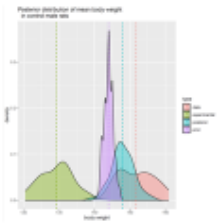
15

Possible uses

efsa European Food Safety Authority

- To derive equivalence limits from HCD to **test for equivalence** between the outcome measures in the treated and control group
- Integrate HCD in the results in the form of a **prior distribution** (reduce n. animals required for control or increase the power keeping the sample size)

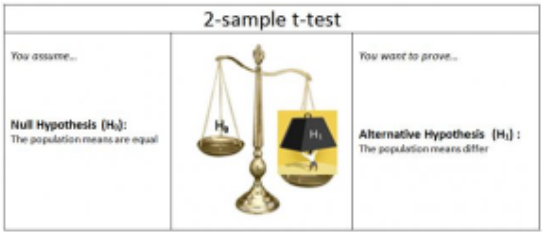


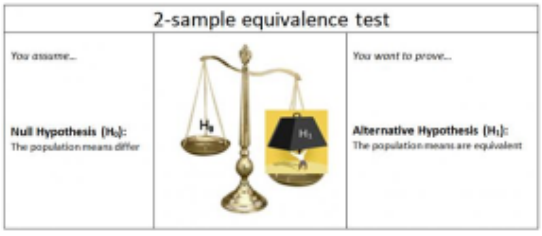


16

Safety concerns: testing for difference or for equivalence?

efsa European Food Safety Authority





In testing Hps priority is given to controlling alpha error (false positive), beta error (false negative) is considered less important and typically is at least 4 times larger than alpha error.
Say alpha=0.05, beta=0.20 (power=0.80)

Example: efficacy of a treatment
From the regulatory perspective, objective is to minimise the risk to introduce on the marker treatment that doesn't work (alpha=5%)

$$\begin{cases} H_0: & \theta_T = \theta_C \\ H_1: & \theta_T \neq \theta_C \end{cases}$$

Example: safety of a substance
From the regulatory perspective objective is to minimise the risk of introducing or keeping in the market a substance that is unsafe = minimise the risk that we conclude towards lack of difference when it is not true (false negative – beta error). Generally the beta error is not so small (above 20%).

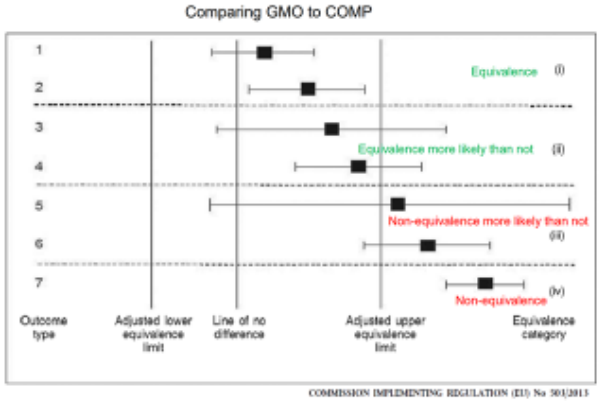
Equivalence testing case represents a more appropriate alternative for addressing safety concerns

$$\begin{cases} H_0: & \theta_T \leq \theta_{LB,EI} \text{ or } \theta_T \geq \theta_{UB,EI} \\ H_1: & \theta_{LB,EI} \leq \theta_T \leq \theta_{UB,EI} \end{cases}$$

EI=Equivalence Interval

From HCD based-RI to equivalence limits

efsa European Food Safety Authority



Commission implementing Regulation 503/2013 on applications for authorisation of genetically modified food and feed

More discussion needed on this topic!
How to derive Equivalence Limits from HCD
- ✓ sufficiently conservative
- ✓ reflecting the uncertainty in their estimate

Too large equivalence limits can inflate conclusion of equivalence (safety) when it is not true (false positive)

18

Bayesian priors using HCD

efsa European Food Safety Authority

Assumption: Concurrent control and HCD come from the same population

HCD are used to derive priors on the parameters of interest (e.g. mean of the body weight)

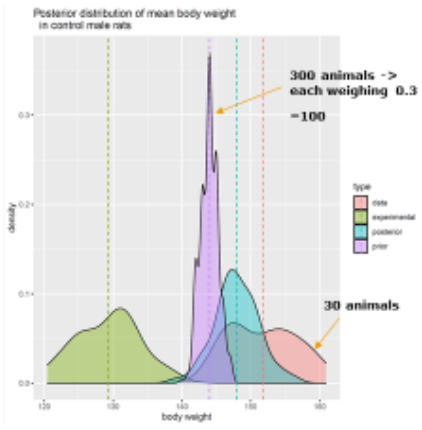
-> Informative prior is used only for the control group

Informative prior from HCD and data from control group are combined according to the bayes theorem

$$\text{posterior} \sim \text{likelihood} \times \text{prior}.$$

Priors can be used:
- To reduce the number of animals in the control group
- To increase the power of the study keeping the sample size for the control

Similarity between concurrent control and HCD can be used to derive a factor (between 0 and 1) to scale down the contribution of 'less similar' studies or individuals to the total control group sample size



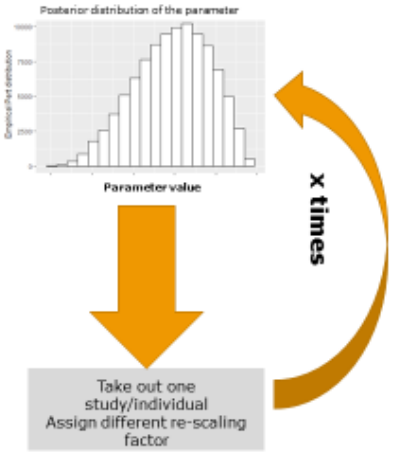
V. Bonapersona, H. Hoijtink, RELACS Consortium*, R. A. Sarabdjitsingh and M. Joëls (2021): Increasing the statistical power of animal experiments with historical control data. Nature NeurOscience, 24, **pp.**470–477. https://doi.org/10.1038/s41593-020-00792-3

19

Role of the sensitivity analysis

efsa European Food Safety Authority

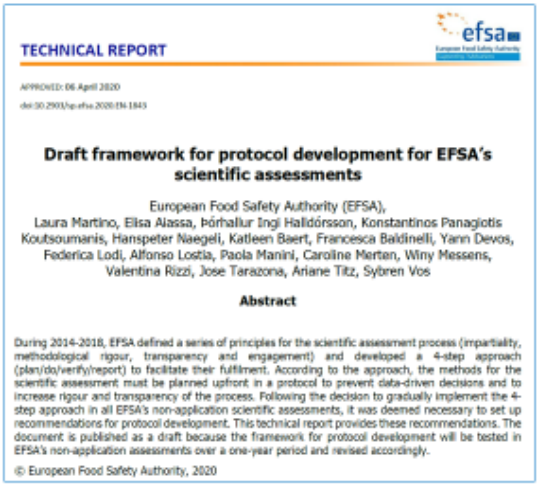
- To assess the influence of each studies/individuals on the final results
- To determine the influence of the factor applied to scale down the sample size contributed from HCD



20

Planning upfront – need for a shared protocol

efsa European Food Safety Authority



TECHNICAL REPORT

efsa

APPROVED: 06 April 2020

doi:10.2903/sp.efsa.2020.EN-1843

Draft framework for protocol development for EFSA's scientific assessments

European Food Safety Authority (EFSA), Laura Martino, Elisa Aiassa, Þórhallur Ingi Halldórsson, Konstantinos Panagiotis Koutsoumanis, Hanspeter Naegeli, Katleen Baert, Francesca Baldinelli, Yann Devos, Federica Lodi, Alfonso Lostia, Paola Manini, Caroline Merten, Winy Messens, Valentina Rizzi, Jose Tarazona, Ariane Titz, Sybren Vos

Abstract

During 2014-2018, EFSA defined a series of principles for the scientific assessment process (impartiality, methodological rigour, transparency and engagement) and developed a 4-step approach (plan/do/verify/report) to facilitate their fulfilment. According to the approach, the methods for the scientific assessment must be planned upfront in a protocol to prevent data-driven decisions and to increase rigour and transparency of the process. Following the decision to gradually implement the 4-step approach in all EFSA's non-application scientific assessments, it was deemed necessary to set up recommendations for protocol development. This technical report provides these recommendations. The document is published as a draft because the framework for protocol development will be tested in EFSA's non-application assessments over a one-year period and revised accordingly.

© European Food Safety Authority, 2020

Report can be downloaded here

Adavntages of planning upfront

- Limiting methodological flaws like Hypothesizing After the Results are Known (HARKing) or data-contingent analysis decisions (P-hacking), by requiring assessors to articulate analytical decisions prior to acquiring knowledge about (and possibly be influenced by) the available results (Munafo et al., 2017).
- Safeguarding against arbitrary decision making during the assessment process.
- Protecting from cognitive biases (Munafo et al., 2017; Shamseer et al., 2015) such as the confirmation bias, i.e., the tendency to focus on evidence that is in line with expectations or favoured explanation (Kerr, 1998).

Overall, these aspects contribute to improve the integrity and defensibility of results.

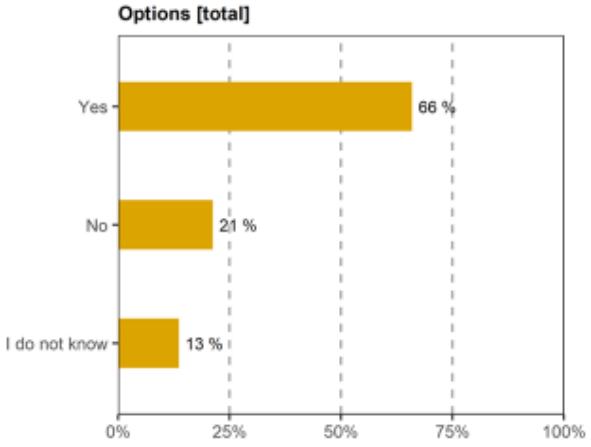
Use of HCD should be planned a-priori in a protocol

21

Open access

efsa European Food Safety Authority

Would a global historical control database and open access to this be beneficial for your purpose?



Do we need a cultural shift to move toward an OPEN ACCESS historical control database ?

22

Presentations Day 2

1. **Historical Control Data in Pathology: Meaningful Use and Limitations**
   (Sibylle Gröters, on behalf of RITA initiative)

Historical Control Data in Pathology: Meaningful Use and Limitations

International Workshop on How to report, use and interpret historical control data in (eco)toxicity studies
3rd – 5th May 2022, virtual event

Dr. Sibylle Gröters, PhD, DVM, Dipl. ECVP, FTA Pathology
Head of Pathology, Vice President
BASF SE, Ludwigshafen, Germany

Co-Chair of the RITA Group

Introduction

- **Generation of pathology data**
  - Who – training and education of Pathologists
  - How – generation of data and diagnostic criteria
  - When – which study types
  - Pathology Peer Review and Pathology Working Group (PWG)
- **RITA**
- **Types of HCD in pathology**
  - Gross lesions, histopathology
  - Organ weights, morphometry, DOFC, cell proliferation
- **Use of HCD in pathology**
  - Sources and examples

May 03-05, 2022 International HCD Workshop, Dr. Sibylle Groeters

RITA Registry of Industrial Toxicology Animal-data

Who – training and education of Pathologists

- Studies of **Veterinary Medicine**
- **Postgraduate** education/**Residency** in Veterinary Pathology
- Board exam
  - International:
    - Board-exam as **Diplomate** of the *American* or *European* or *Japanese College of Veterinary Pathologists* (DACVP, DECVP, DJCVP)
  - National examination in different European countries (France, Germany, UK, Netherlands…)

RITA Registry of Industrial Toxicology Animal-data

May 03-05, 2022 International HCD Workshop, Dr. Sibylle Groeters

Who – training and education of Pathologists

- **First…Diagnostic Pathology**
  - Domestic, wildlife and livestock animals
  - Spontaneous and infectious diseases, primarily
  - University, diagnostic laboratories…
- **Then…Toxicologic Pathology**
  - Laboratory animals (rodent, non-human primates, dog, minipig…)
  - Induced lesions, primarily
  - Spontaneous diseases and infections
  - Industry, CRO's (contract research organization), University

→ **Mixed knowledge necessary**

RITA Registry of Industrial Toxicology Animal-data

May 03-05, 2022 International HCD Workshop, Dr. Sibylle Groeters

How – diagnostic criteria and generation of data

- All histotechnical processing is done according to **inhouse SOPs** (Standard Operating Procedure)
- Trimming of organs is done according to international harmonized guidance documents "**Revised guides for organ sampling and trimming in rats and mice**" (https://reni.item.fraunhofer.de/reni/trimming/)
- Standardization and comparability of HE-stained slides must be given
  - Within one company over the years
  - Between companies worldwide

RITA Registry of Industrial Toxicology Animal-data

May 03-05, 2022 International HCD Workshop, Dr. Sibylle Groeters

How – diagnostic criteria and generation of data

- Diagnostic criteria and terms must be used **standardized** worldwide
- "**INHAND** - the **In**ternational **Ha**rmonization of **N**omenclature and **D**iagnostic criteria" is the standard reference for nomenclature and diagnostic criteria in toxicologic pathology
- Initiative was founded by international ToxPath Societies in **2005**
- INHAND organ working groups have up to 15 international **recognized pathologists** from all relevant ToxPath Societies (BSTP, ESTP, JSTP, STP…)
- INHAND fascicles are published by the official journals of ToxPath Societies
  - Up to now, 20 rodent and non-rodent publications available
- **goRENI** (https://www.goreni.org/) (**g**lobal **o**pen **Re**gistry **N**omenclature **I**nformation System) is the Internet discussion platform for this global initiative
- **Access** to goRENI is available to all members of the participating **ToxPath Societies** as well as to members of **authorities**

RITA Registry of Industrial Toxicology Animal-data

May 03-05, 2022 International HCD Workshop, Dr. Sibylle Groeters

How – diagnostic criteria and generation of data

Toxicologic Pathology, 40: 78-138, 2012
Copyright © 2012 by The Author(s)
ISSN: 0192-6233 print / 1533-1601 online
DOI: 10.1177/0192623312439125

Inte
Nomen

Peter C. Ma
Takan

Toxicologic Pathology, 40: 878-1578, 2012
Copyright © 2012 by The Author(s)
ISSN: 0192-6233 print / 1533-1601 online
DOI: 10.1177/0192623312439125

**Proliferative and N
Mouse Central**

Wolfgang Kaufmann[1], Brad Bolon[2], Catherine George[6], Sibylle Gröters[7], G
Ma

Invited Review

International Harmonization of Nomenclature and Diagnostic Criteria (INHAND): Nonproliferative and Proliferative Lesions of the Dog

Toxicologic Pathology
2021, Vol. 49(1) 5-109
© The Author(s) 2021
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0192623320948161
journals.sagepub.com/home/tpx
SAGE

Jochen Woicke[1], Muthafar M. Al-Haddawi[2], Jean-Guy Bienvenu[3], Jessica M. Caverly Rae[4], Franck J. Chanut[5], Karyn Colman[6], John M. Cullen[7], Wendell Davis[8], Ryo Fukuda[9], Maike Huisinga[10], Ursula Junker Walker[11], Kiyonori Kai[12], Ramesh C. Kovi[13,14], Nicholas P. Macri[15], Heike-Antje Marxfeld[10], Kristen J. Nikula[16], Ingrid D. Pardo[17], Thomas J. Rosol[18], Alok K. Sharma[19], Bhanu P. Singh[20], Kazutoshi Tamura[21], Michael S. Thibodeau[22], Enrico Vezzali[23], Justin D. Vidal[24], and Emily K. Meseck (GESC Liaison)[25]

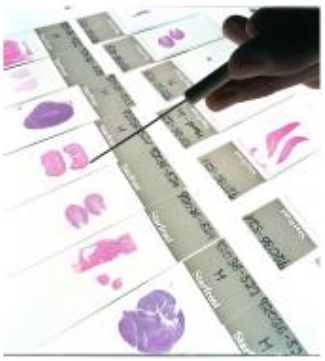
May 03-05, 2022

Interna

When – which study types

Scope of examination (rat)

28d (OECD 407)	**900** tissue samples
90d (OECD 408)	**2,500** tissue samples
2-generation (OECD 416)	**3,500** tissue samples
Ext. 1-Gen (OECD 443)	**15,000** tissue samples
24-month (OECD 453)	**20,000** tissue samples



RITA
Registry of Industrial Toxicology Animal-data

May 03-05, 2022

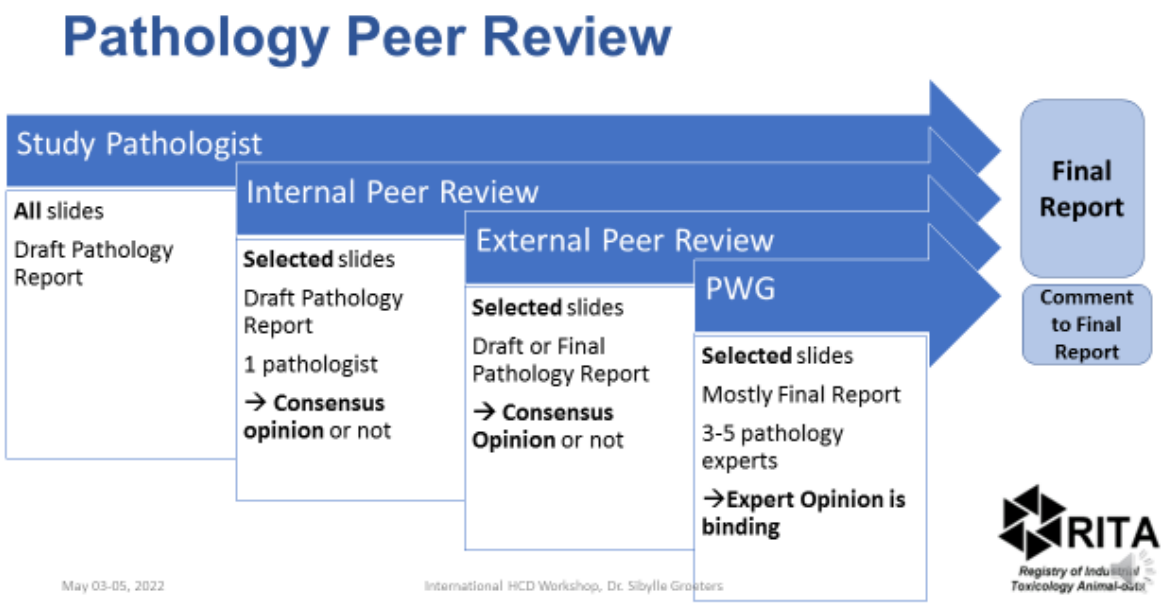
International HCD Workshop, Dr. Sibylle Groeters

Pathology Peer Review

- Pathology results can undergo an internal and/or external **Pathology Peer Review** (second/third opinion)
- In case of questionable findings, a so-called **PWG** (Pathology Working Group) with international specialists might be helpful
- Procedures are **well-established** and described in inhouse SOP's and publications (see literature slide)
- Consensus opinion should be reached

May 03-05, 2022 International HCD Workshop, Dr. Sibylle Groeters

RITA
Registry of Industrial Toxicology Animal-data


Pathology Peer Review
Study Pathologist
All slides
Draft Pathology Report
Internal Peer Review
Selected slides
Draft Pathology Report
1 pathologist
→ Consensus opinion or not
External Peer Review
Selected slides
Draft or Final Pathology Report
→ Consensus Opinion or not
PWG
Selected slides
Mostly Final Report
3-5 pathology experts
→Expert Opinion is binding
Final Report
Comment to Final Report
RITA
Registry of Industrial Toxicology Animal-data
May 03-05, 2022
International HCD Workshop, Dr. Sibylle Grooters


RITA – Registry of Industrial Toxicology Animal-data
Founded in 1988 by German and Swiss (agro)chemical and pharmaceutical companies (industry-sponsored project)
Today 8 companies from Europe and USA
Main activities:
Generation and maintenance of database for historical control data on tumors and tumor precursor lesions in rodents
Setting of standards in conduct and interpretation of rodent carcinogenicity studies
Training of young toxicologic pathologists in diagnostics of carcinogenicity studies
Homepage RITA - Registry of Industrial Toxicology Animal-data (fraunhofer.de)
RITA
Registry of Industrial Toxicology Animal-data
May 03-05, 2022
International HCD Workshop, Dr. Sibylle Grooters

RITA – Registry of Industrial Toxicology Animal-data

- **Industry partners:**
  - Pharmaceutical and (agro)chemical companies with IP rights on carcinogenicity studies
    - Provide data and slides from carcinogenicity studies
    - Participate in scientific activities (panel meetings, publications, …)
    - Receive historical control data from service provider
- **Service provider and scientific support**:
  - Fraunhofer Institute for Toxicology and Experimental Medicine (ITEM), Hannover, Germany
    - Preparation for and host of the RITA Panel Meetings
    - Maintenance of the RITA database

RITA Registry of Industrial Toxicology Animal-data

May 03-05, 2022 International HCD Workshop, Dr. Sibylle Groeters

RITA – Registry of Industrial Toxicology Animal-data

- **Purpose** is
  - to centralize the collection of HCD in rodent carcinogenicity studies from different laboratories in a consistent manner
  - → **harmonized data** from these studies in a comprehensive database
- Cross organizational **review** of studies → histopathological assessment of tumors meets highest standards for **reliability**, **robustness** and **quality**

RITA Status: **15-Mar-2022**

Species	Status	Studies	Animals	Primary Tumors (including Polyps)	Pre-neoplastic Lesions	Total Cases	Open Cases
rat	finalized	200	20,656	31,732	36,746	78,763	8
mouse	finalized	102	10,312	9,332	6,183	39,370	12
mouse (transgenic)	finalized	16	788	160	120	293	7
mouse (transgenic: pos. contr.)	finalized	12	407	761	96	2,164	1
hamster	finalized	5	500	909	1,636	2,921	0

RITA Registry of Industrial Toxicology Animal-data

May 03-05, 2022 International HCD Workshop, Dr. Sibylle Groeters

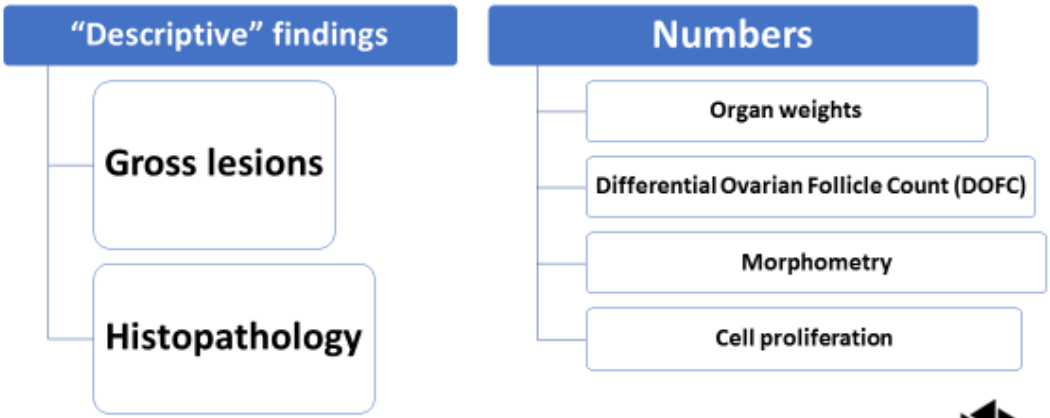
RITA – Registry of Industrial Toxicology Animal-data

- Panel meetings are valuable tool for
  - Continuing education of experienced pathologists
  - Training for less experienced pathologists
  - Scientific exchange and discussion on borderline cases
  - Accepted by authorities as "ring trial" for pathologists (cross evaluation of findings)

RITA Registry of Industrial Toxicology Animal-data

May 03-05, 2022 International HCD Workshop, Dr. Sibylle Groeters

Types of HCD in Pathology



RITA Registry of Industrial Toxicology Animal-data

May 03-05, 2022 International HCD Workshop, Dr. Sibylle Groeters

Types of HCD in Pathology

- **Gross** lesions
  - Macroscopic findings → must be investigated by light microscopy
  - Solely descriptive, mostly no diagnostic value
    - Mass, lesion, focus, discoloration…
  - Can be collected for HCD data base, but only of limited value
  - Incidence of gross lesions might correlate to different histopathological diagnoses

Gross lesion (liver)	Histopathology (liver)
Focus	Inflammation (abscess)
Focus	Adenoma, hepatocellular
Focus	Carcinoma, hepatocellular
Focus	Focus, tigroid
Focus	Focus, eosinophilic
Focus	….

May 03-05, 2022

RITA
Registry of Industrial Toxicology Animal-data

Histopathology

- **Histopathological** findings
  - Non-neoplastic lesions
    - Vacuolation, hypertrophy, necrosis, inflammation…
    - In most facilities not collected on a regular basis for HCD data base
  - Pre-neoplastic lesions
    - Focal hyperplasia, atypical hyperplasia…
    - regularly collected for HCD database
  - Neoplastic lesions
    - Benign tumors
    - Malignant tumors
    - regularly collected for HCD data base

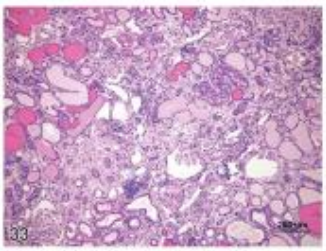
May 03-05, 2022 International HCD Workshop, Dr. Sibylle Grooters

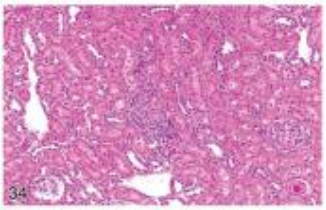
RITA Registry of Industrial Toxicology Animal-data

Histopathology

- **Non-neoplastic** lesions:
- lesions might be diagnosed differently in different study types (407/408 versus 453) → CAVE HCD
  - Kidney, tubules basophilic
    - Might be treatment-related in younger animals → collected as single diagnosis
    - Part of common background lesion "CPN" (chronic progressive nephropathy) → not collected as single diagnosis

FIGURE 33.—Chronic progressive nephropathy, kidney, rat, (advanced), characterized by tubular atrophy, interstitial fibrosis, tubular dilation, casts, hyperplasia, dilated bowman's space, glomerulosclerosis, glomerular atrophy, casts and interstitial inflammatory infiltrates. FIGURE 34.—Chronic progressive nephropathy kidney, rat (early) characterized by focal tubular basophilia, nuclear crowding, and thickened basement membranes. FIGURE 35.—Casts, hyaline, kidney, mouse. FIGURE 36.—Casts, granular, kidney, rat.





Frazier et al (2012) Toxicol Pathol 40:145-865

RITA Registry of Industrial Toxicology Animal-data

Histopathology

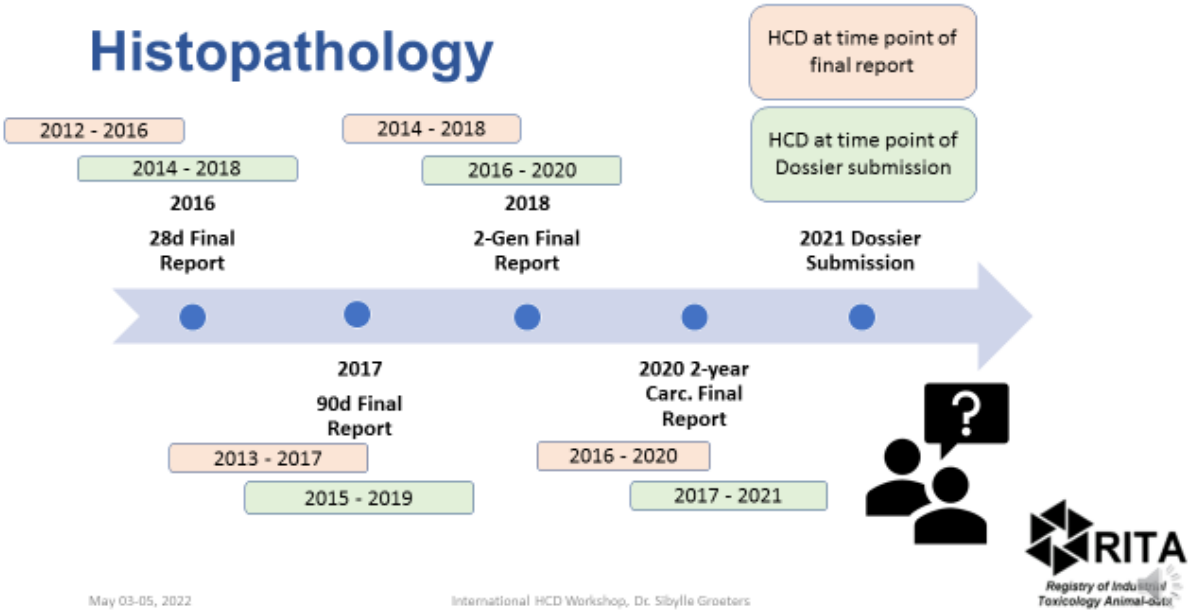
- **Pre-neoplastic/neoplastic** lesions
- Regularly collected
- **Difficulties** for HCD:
  - Low incidence of tumor – more studies needed
  - Low number of studies available in required time frame (5 years)
  - In a 5-year time frame more short-term studies can be performed compared to long-term studies
  - 5-year time frame can change over the time

    Final report ←→ Dossier submission

RITA Registry of Industrial Toxicology Animal-data

May 03-05, 2022 International HCD Workshop, Dr. Sibylle Groeters

Histopathology



RITA Registry of Industrial Toxicology Animal-data

May 03-05, 2022 International HCD Workshop, Dr. Sibylle Groeters

Histopathology

- Example
  - comb. Chronic/Carcinogenicity study (OECD 453), July 2019 – July 2021 (in-life)

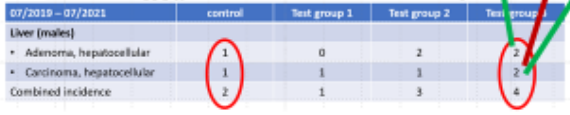
07/2019 – 07/2021	control	Test group 1	Test group 2	Test group 3
Liver (males)				
• Adenoma, hepatocellular	1	0	2	2
• Carcinoma, hepatocellular	1	1	1	2
Combined incidence	2	1	3	4

- Question:
  - Dose-dependency? → could be in TG 2 and 3
  - Statistically significant? → NO
  - HCD? → see next slide
  - Treatment-related?

RITA Registry of Industrial Toxicology Animal-data

May 03-05, 2022 International HCD Workshop, Dr. Sibylle Groeters

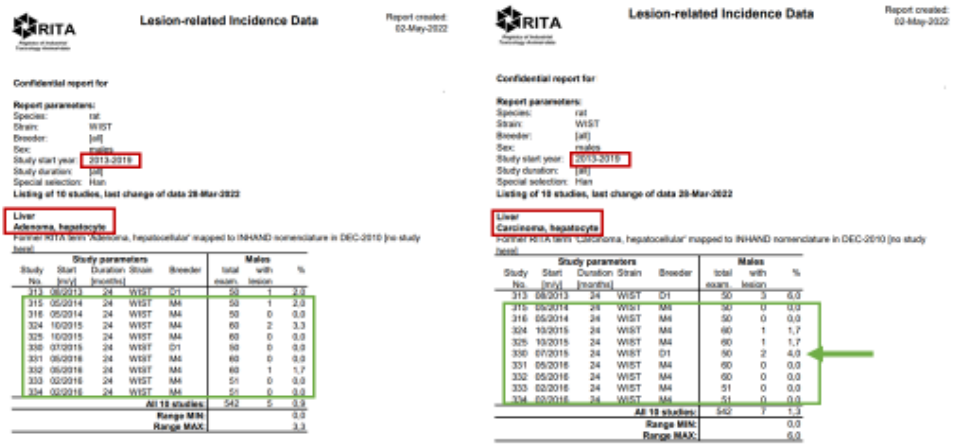
	Time	Route	No of animals	Adenoma number	Adenoma %	Carcinoma number	Carcinoma %	Combined number	Combined %
6y	02.13-02.15	feeding	50	3	6	3	6	6	12
	03.13-03.15	feeding	50	3	6	3	6	6	12
	05.13-05.15	feeding	50	0	0	2	4	2	4
	07.13-07.15	feeding	50	1	2	2	4	3	6
	10.13-10.15	feeding	50	4	8	0	0	4	8
5y	04.15-04.17	feeding	50	2	4	1	2	3	6
	07.15-07.17	feeding	50	0	0	1	2	1	2
	07.17-07.19	feeding	50	2	4	1	2	3	6
	SUM		400						
	Number of studies		8	☑	☑	☒			
	Mean			1.875	3.75	2	4	3.875	7.75
	MIN			0	0	0	0	2	4
					8	3	6	6	12



RITA Registry of Industrial Toxicology Animal-data

shop, Dr. Sibylle Groeters

Histopathology



RITA Lesion-related Incidence Data — Report created: 02-May-2022

Confidential report for

Report parameters:
Species: rat
Strain: WIST
Breeder: [all]
Sex: males
Study start year: 2013-2018
Study duration: [all]
Special selection: Han
Listing of 10 studies, last change of data 28-Mar-2022

Liver
Adenoma, hepatocyte

Former RITA term 'Adenoma, hepatocellular' mapped to INHAND nomenclature in DEC-2010 [no study found]

Study No.	Start [m/y]	Duration [months]	Strain	Breeder	Males: total exam.	Males: with lesion	Males: %
313	08/2013	24	WIST	D1	50	1	2.0
315	05/2014	24	WIST	M4	50	1	2.0
316	05/2014	24	WIST	M4	50	0	0.0
324	10/2015	24	WIST	M4	60	2	3.3
325	10/2015	24	WIST	M4	60	0	0.0
330	07/2015	24	WIST	D1	50	0	0.0
331	05/2016	24	WIST	M4	60	0	0.0
332	05/2016	24	WIST	M4	60	1	1.7
333	02/2016	24	WIST	M4	51	0	0.0
334	02/2016	24	WIST	M4	51	0	0.0
				All 10 studies:	542	5	0.9
				Range MIN:			0.0
				Range MAX:			3.3

RITA Lesion-related Incidence Data — Report created: 02-May-2022

Confidential report for

Report parameters:
Species: rat
Strain: WIST
Breeder: [all]
Sex: males
Study start year: 2013-2018
Study duration: [all]
Special selection: Han
Listing of 10 studies, last change of data 28-Mar-2022

Liver
Carcinoma, hepatocyte

Former RITA term 'Carcinoma, hepatocellular' mapped to INHAND nomenclature in DEC-2010 [no study found]

Study No.	Start [m/y]	Duration [months]	Strain	Breeder	Males: total exam.	Males: with lesion	Males: %
313	08/2013	24	WIST	D1	50	3	6.0
315	05/2014	24	WIST	M4	50	0	0.0
316	05/2014	24	WIST	M4	50	0	0.0
324	10/2015	24	WIST	M4	60	1	1.7
325	10/2015	24	WIST	M4	60	1	1.7
330	07/2015	24	WIST	D1	50	2	4.0
331	05/2016	24	WIST	M4	60	0	0.0
332	05/2016	24	WIST	M4	60	0	0.0
333	02/2016	24	WIST	M4	51	0	0.0
334	02/2016	24	WIST	M4	51	0	0.0
				All 10 studies:	542	7	1.3
				Range MIN:			0.0
				Range MAX:			6.0

RITA Registry of Industrial Toxicology Animal-data

May 03-05, 2022 — International HCD Workshop, Dr. Sibylle Groeters

Histopathology

- Treatment-related increase in carcinoma, hepatocellular?
  - Higher incidence in TG 3 compared to concurrent control
    - But no statistical significance
    - No clear dose-dependency (only increase in TG 3)
  - HCD test facility 5-year time frame (2019 → **2014**)
    - **Outside** HCD in this time frame, but **only 3 studies** available
  - HCD test facility 6-year time frame (2019 → **2013**)
    - **Within** HCD in this time frame (8 studies)
  - HCD RITA 5-year time frame (2019 → **2014**)
    - **Within** HCD in this time frame (9 studies)
  - HCD RITA 6-year time frame (2019 → **2013**)
    - **Within** HCD in this time frame (10 studies)

→ **incidental**

RITA Registry of Industrial Toxicology Animal-data

May 03-05, 2022 — International HCD Workshop, Dr. Sibylle Groeters

Histopathology

- Important point in Carc-studies → n
  frame

RITA Lesion-related Incidence Data

Report created: 14-May-2020

Confidential report for

Report parameters:
Species: rat
Strain: WIST
Breeder: [all]
Sex: [all]
Study start year: 2015-2020
Study duration: [all]
Special selection: Han
Listing of 1 study, last change of data 20-Nov-2019

Uterus
Adenocarcinoma, endometrium

Study No.	Start [m/y]	Duration [months]	Strain	Breeder	Males: total exam	Males: with lesion	Males: %	Females: total exam	Females: with lesion	Females: %
324	10/2015	24	WIST	M4	0			60	2	3,3

[No summary required]

RITA Lesion-related Incidence Data

Report created: 14-May-2020

Confidential report for

Report parameters:
Species: rat
Strain: WIST
Breeder: [all]
Sex: [all]
Study start year: 2010-2020
Study duration: [all]
Special selection: Han
Listing of 10 studies, last change of data 20-Nov-2019

Uterus
Adenocarcinoma, endometrium

Study No.	Start [m/y]	Duration [months]	Strain	Breeder	Males: total exam	Males: with lesion	Males: %	Females: total exam	Females: with lesion	Females: %
275	02/2010	25	WIST	S	0			60	2	3,3
286	12/2010	25	WIST	D4	0			60	5	8,3
287	12/2010	25	WIST	D4	0			60	1	1,7
288	02/2012	24	WIST	D4	0			52	3	5,8
298	06/2012	25	WIST	S	0			59	1	1,7
315	05/2014	24	WIST	M4	0			50	0	0,0
316	05/2014	24	WIST	M4	0			50	2	4,0
322	06/2013	24	WIST	D1	0			50	2	4,0
323	06/2013	30	WIST	D1	0			49	7	14,3
324	10/2015	24	WIST	M4	0			60	2	3,3
				All 10 studies:	0			550	25	4,5
				Range MIN:						0,0
				Range MAX:						14,3

RITA Registry of Industrial Toxicology Animal-data

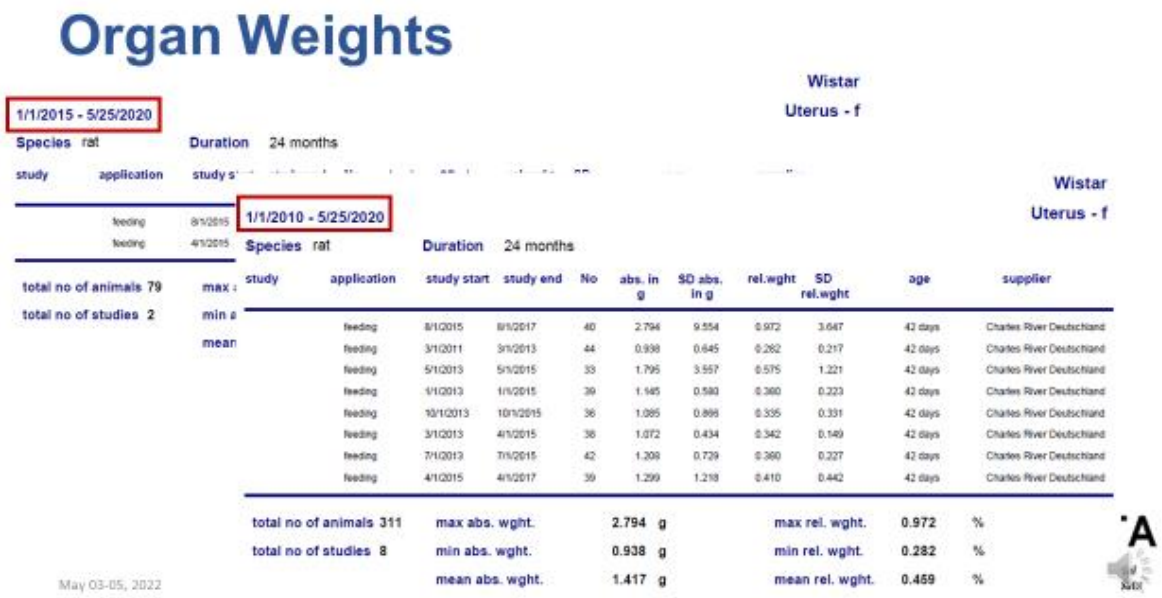
May 03-05, 2022 International HCD Workshop, Dr. Sibylle Groeters

Organ Weights

- **General**
  - Small organs have high variation (measuring scale and preparation!)
  - Low number of studies available in required time frame (5 years) for certain study types (long-term studies!)
- **Rats**
  - Most organs give robust measurement
  - Carcinogenicity study → tumors!
- **Mice**
  - Adrenal glands, spleen, thymus, seminal vesicles have high variability in 28d study (Marxfeld et al., 2019)

RITA Registry of Industrial Toxicology Animal-data

May 03-05, 2022 International HCD Workshop, Dr. Sibylle Groeters



Organ Weights

Wistar

Uterus - f

1/1/2015 - 5/25/2020

Species rat Duration 24 months

study	application	study start	study end	No	abs. in g	SD abs. in g	rel.wght	SD rel.wght	age	supplier
feeding		8/1/2015	8/1/2017	40	2.794	9.554	0.972	3.647	42 days	Charles River Deutschland
feeding		4/1/2015	4/1/2017	39	1.299	1.218	0.410	0.442	42 days	Charles River Deutschland

Wistar

Uterus - f

1/1/2010 - 5/25/2020

Species rat Duration 24 months

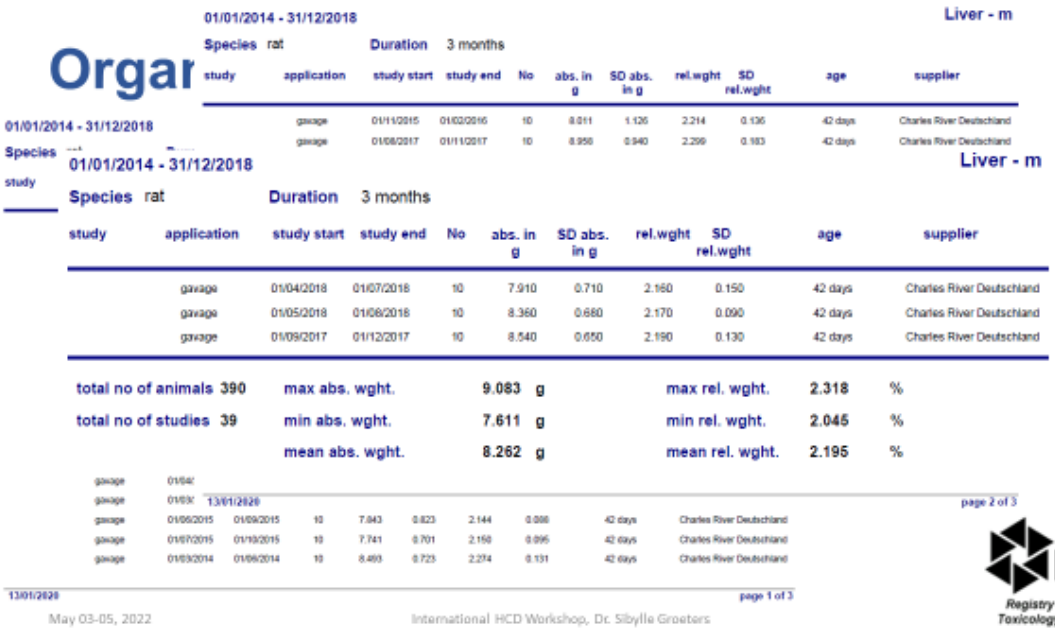
study	application	study start	study end	No	abs. in g	SD abs. in g	rel.wght	SD rel.wght	age	supplier
feeding		8/1/2015	8/1/2017	40	2.794	9.554	0.972	3.647	42 days	Charles River Deutschland
feeding		3/1/2011	3/1/2013	44	0.938	0.645	0.282	0.217	42 days	Charles River Deutschland
feeding		5/1/2013	5/1/2015	33	1.795	3.557	0.575	1.221	42 days	Charles River Deutschland
feeding		1/1/2013	1/1/2015	39	1.145	0.580	0.360	0.223	42 days	Charles River Deutschland
feeding		10/1/2013	10/1/2015	36	1.085	0.898	0.335	0.331	42 days	Charles River Deutschland
feeding		3/1/2013	4/1/2015	38	1.072	0.434	0.342	0.149	42 days	Charles River Deutschland
feeding		7/1/2013	7/1/2015	42	1.209	0.729	0.360	0.227	42 days	Charles River Deutschland
feeding		4/1/2015	4/1/2017	39	1.299	1.218	0.410	0.442	42 days	Charles River Deutschland

total no of animals 311 max abs. wght. 2.794 g max rel. wght. 0.972 %

total no of studies 8 min abs. wght. 0.938 g min rel. wght. 0.282 %

mean abs. wght. 1.417 g mean rel. wght. 0.459 %

May 03-05, 2022



Organ

01/01/2014 - 31/12/2018

Species rat Duration 3 months

Liver - m

study	application	study start	study end	No	abs. in g	SD abs. in g	rel.wght	SD rel.wght	age	supplier
	gavage	01/04/2018	01/07/2018	10	7.910	0.710	2.160	0.150	42 days	Charles River Deutschland
	gavage	01/05/2018	01/08/2018	10	8.360	0.680	2.170	0.090	42 days	Charles River Deutschland
	gavage	01/09/2017	01/12/2017	10	8.540	0.650	2.190	0.130	42 days	Charles River Deutschland

total no of animals 390 max abs. wght. 9.083 g max rel. wght. 2.318 %

total no of studies 39 min abs. wght. 7.611 g min rel. wght. 2.045 %

mean abs. wght. 8.262 g mean rel. wght. 2.195 %

13/01/2020 page 1 of 3

13/01/2020 page 2 of 3

RITA Registry of Industrial Toxicology Animal-data

Organ weights

- **Dogs**
  - Most organs give robust measurements
  - BUT high variability in 28d and 90d studies (and 1y study)
    - Reproductive organs in males
      - Prostate (!), testis, epididymis
    - Reproductive organs in females
      - Uterus (!), ovaries
  - due to
    - → different age at start of the study (4-9 months)
    - → interindividual differences in start of puberty
    - → stage of development (prepubertal, postpubertal)
    - → long estrous cycle in females

May 03-05, 2022 International HCD Workshop, Dr. Sibylle Groeters

RITA Registry of Industrial Toxicology Animal-data

Organ weights – Dog studies

	Prostate (weight in gramm)			Uterus (weight in gramm)		
	28d	90d	1y	28d	90d	1y
	2,37	6,93	9,33	14,24	15,26	11,43
	4,21	9,59	7,27	12,58	15,28	8,41
	11,99	4,73	14,43	3,54	5,71	4,06
	1,94	6,56	8,13	3,42	13,30	4,97
	5,10	2,62	9,27	1,27	18,29	3,67
	2,68			1,17		
	3,46			5,56		
	3,17			3,57		
MIN	1,94	2,62	7,27	1,17	5,71	3,67
MAX	11,99	9,59	14,43	14,24	18,29	11,43
Mean	4,37	6,09	9,69	5,67	13,57	6,51

- 28d, 90d (OECD 409) and 1y (OECD 452) in Marshall Beagles Dogs

May 03-05, 2022 International HCD Workshop, Dr. Sibylle Groeters

RITA Registry of Industrial Toxicology Animal-data

Other "numbers"

- **DOFC**
  - Can be compared within one test facility, only
  - Different SOP's for counting
  - Counting of primordial and growing follicles is done manually by light microscopy
  - First results available regarding AI and machine learning
- **Morphometry** in DNT studies (Developmental Neurotoxicity; OECD 443, 426)
  - Precise trimming, embedding and cutting necessary
  - Critical to measure symmetrical levels (Paxinos brain atlas!)
  - Technical artefacts might produce wrong measurements
- **Cell proliferation**
  - Mostly done on BrdU-stained slides
  - Counting (manually or on scanned slides with computer systems)

May 03-05, 2022 International HCD Workshop, Dr. Sibylle Groeters

RITA Registry of Industrial Toxicology Animal-data

Conclusion

- 5-year time frame might be critical
  - Depends on **study type** and **number** of available studies
  - Might also depend on **frequency** of finding
- Organ weights
  - **Not all** organs and all species can be judged on the **same** way
- Ranking of HCD
  - **Concurrent** control
  - **Inhouse** (same test facility, identical conditions, same strain, breeder, age of animals, form of application, comparable study duration…)
  - **RITA** (for tumors)
  - **Other** sources: NTP, literature, CRO's
- Necessary endpoints
  - Min, max, mean, SD (only for true numbers)
  - **Trend analysis** in some cases (tumors over the years)

May 03-05, 2022 International HCD Workshop, Dr. Sibylle Groeters

RITA Registry of Industrial Toxicology Animal-data

I would like to say **THANK YOU** to all colleagues working over the years at the RITA-project, the participating companies and especially the colleagues from Fraunhofer ITEM in Hannover!

May 03-05, 2022 International HCD Workshop, Dr. Sibylle Groeters

RITA Registry of Industrial Toxicology Animal-data

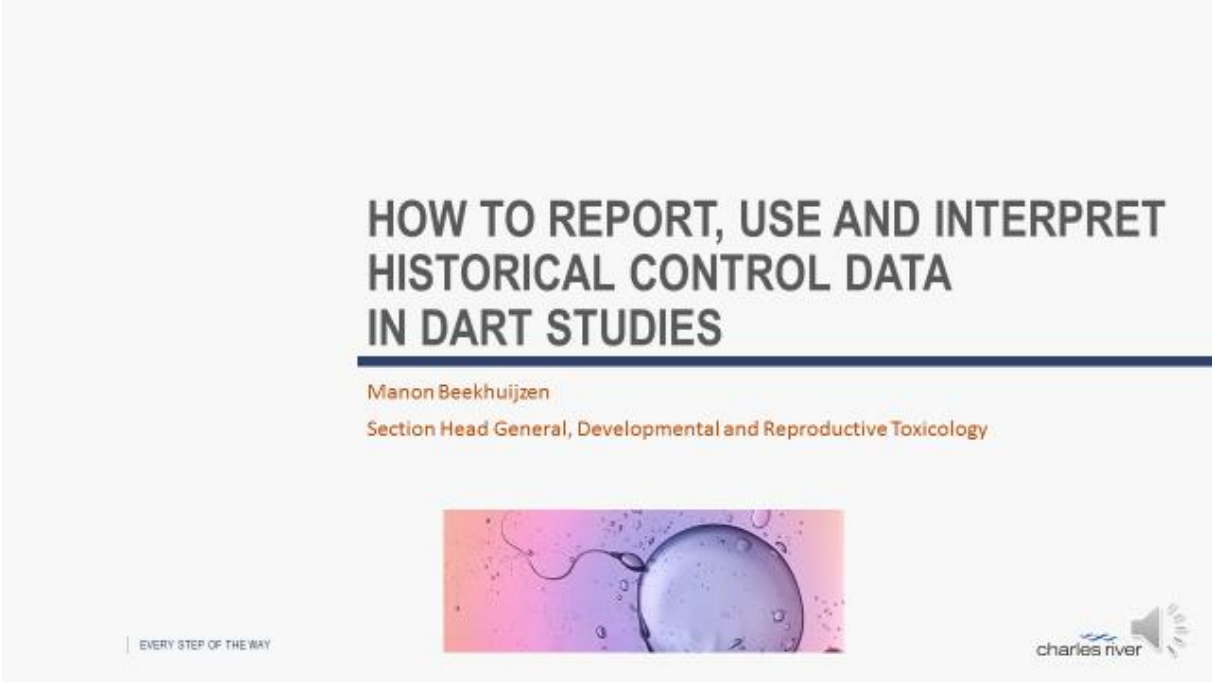
Literature

- Fikes, J. D., Patrick, D. J., Francke, S., Frazier, K. S., Reindel, J. F., Romeike, A., Spaet, R. H., Tomlinson, L., Schafer, K. A. (2015). **Scientific and Regulatory Policy Committee Review: Review of the Organisation for Economic Co-operation and Development (OECD) Guidance on the GLP Requirements for Peer Review of Histopathology.** *Toxicologic Pathology*, 43(7), 907–914. doi:10.1177/0192623315596382
- Morton, D., Sellers, R. S., Barale-Thomas, E., Bolon, B., George, C., Hardisty, J. F., Irizarry, A., McKay, J. S., Odin, M., & Teranishi, M.F. (2010). **Recommendations for Pathology Peer Review.** *Toxicologic Pathology*, 38(7), 1118–1127. doi:10.1177/0192623310383991
- STP Regulatory Affairs Committee. (1997). **Documentation of Pathology Peer Review.** *Toxicologic Pathology*, 25(6), 655–655. doi:10.1177/019262339702500624
- US FDA Pathology Peer Review in Nonclinical Toxicology Studies: Questions and Answers Guidance for Industry Pathology Peer Review in Nonclinical Toxicology Studies: Questions and Answers (fda.gov)
- Marxfeld HA, Küttler K, Dammann M, Gröters S, van Ravenzwaay B. Body and organ weight data in 28-day toxicological studies in two mouse strains. Data Brief. 2019 Oct 7;27:104632. doi: 10.1016/j.dib.2019.104632. eCollection 2019 Dec.
- Keenan, C., Elmore, S., Francke-Carroll, S., Kemp, R., Kerlin, R., Peddada, S., Pletcher, J., Rinke, M., Schmidt, S. P., Taylor, I., & Wolf, D. C. (2009). **Best Practices for Use of Historical Control Data of Proliferative Rodent Lesions.** *Toxicologic Pathology*, 37(5), 679–693. doi:10.1177/0192623309336154
- Elmore, S., & Peddada, S., (2009). **Points to Consider on the Statistical Analysis of Rodent Cancer Bioassay Data Incorporating Historical Control Data.** *Toxicologic Pathology*, 37(5), 672–676. doi:10.1177/0192623309339606

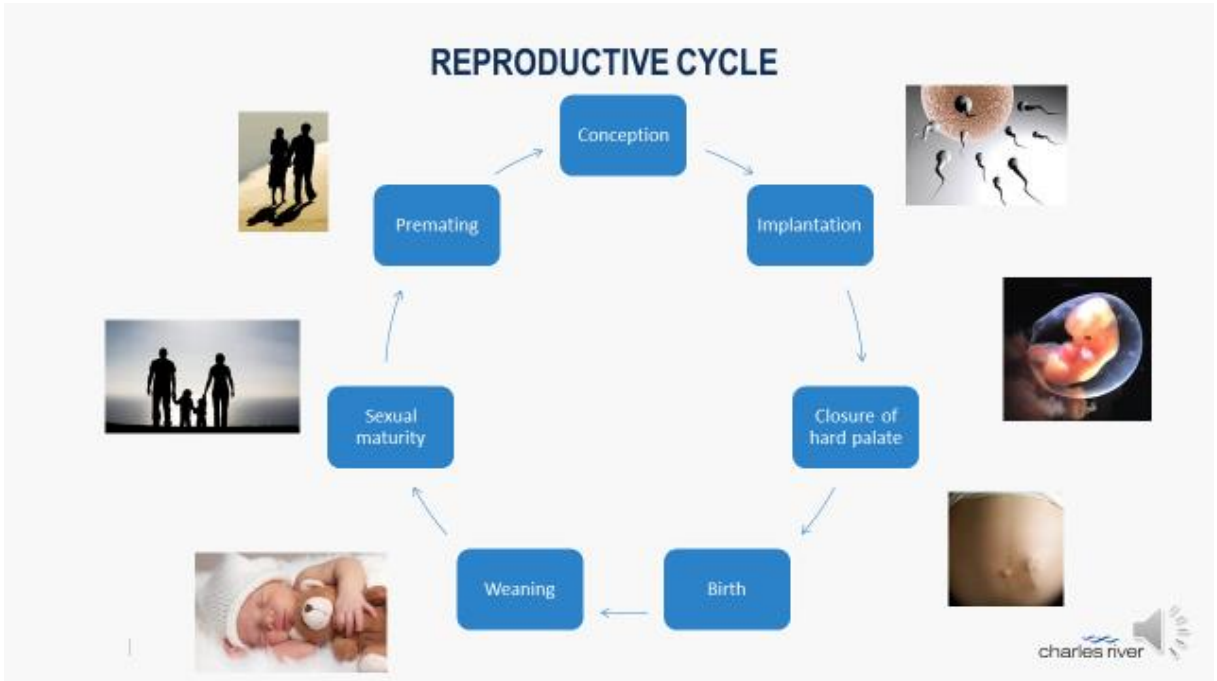
RITA
Registry of Industrial Toxicology Animal-data

2. How to report, use and interpret historical control data in DART studies

(Manon Beekhuijzeen, Charles River Laboratories)




Introduction DART
DEVELOPMENTAL AND REPRODUCTIVE TOXICOLOGY


REPRODUCTIVE CYCLE
Conception
Implantation
Closure of hard palate
Birth
Weaning
Sexual maturity
Premating
charles river

DART STUDY TYPES FOR CHEMICALS

efsa

TYPE OF DEVELOPMENTAL STUDY	OECD guideline
Prenatal developmental toxicity study in the rat	414
Prenatal developmental toxicity study in the rabbit	414

TYPE OF REPRODUCTION STUDY	OECD guideline
Screening reproductive/developmental test (rat)	421 or 422
Two-generation reproductive toxicity study (rat)	416
Extended one-generation reproductive toxicity study (EOGRTS) (rat)	443

4 | EVERY STEP OF THE WAY

charles river

PRENATAL DEVELOPMENTAL TOXICITY

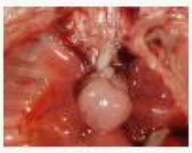
OECD 414

Study design:

- Mated female rats or rabbits (22/group, 4 groups) are exposed to the test item during pregnancy
- Additional ED endpoints for rats (not rabbits)

Specific measurements:

- External examination of the fetuses
- Visceral examination of the fetuses (i.e. organs, blood vessels, etc)
- Skeletal examination of the fetuses (i.e. bone and sometimes cartilage)



5 | EVERY STEP OF THE WAY

charles river

REPRODUCTION TOXICOLOGY STUDIES

Chemicals

- Reproduction/developmental toxicity screening test (OECD 421/422)
- Two-generation reproduction toxicology study (OECD 416)
- Extended One-generation reproduction toxicology study (EOGRTS; OECD 443)

6 | EVERY STEP OF THE WAY

charles river

SCREENING REPRODUCTIVE/DEVELOPMENTAL TEST OECD 421/422

Number of animals

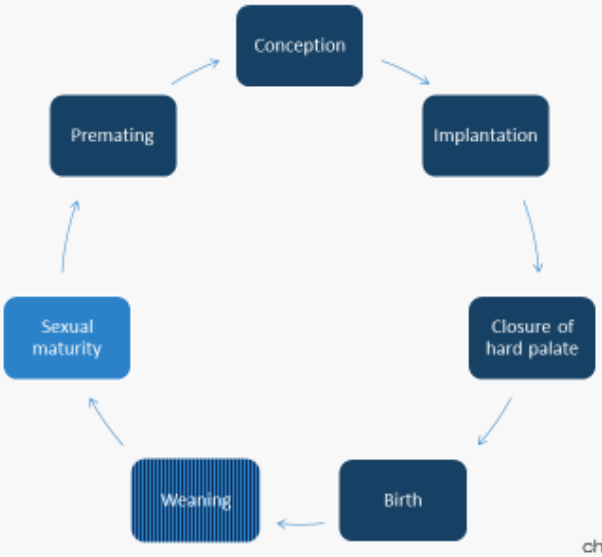
10 animals/sex/group

Dosing

2 weeks premating

↓

Day 14 of lactation



7 | EVERY STEP OF THE WAY

charles river

SCREENING REPRODUCTIVE/DEVELOPMENTAL TEST OECD 421/422

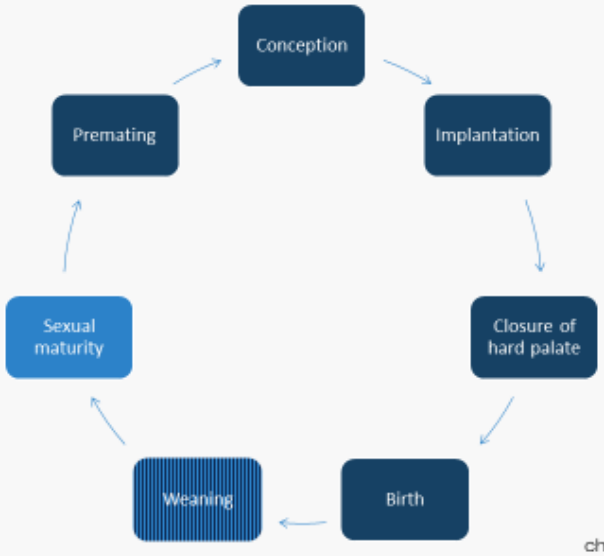
OECD 421

Only DART parameters

OECD 422

DART + 28d parameters

(functional observations, clinical pathology, full necropsy, full histopathology)



8 | EVERY STEP OF THE WAY

charles river

TWO-GENERATION STUDY OECD 416

Number of animals

25 animals/sex/group

Dosing

10 weeks premating

↓

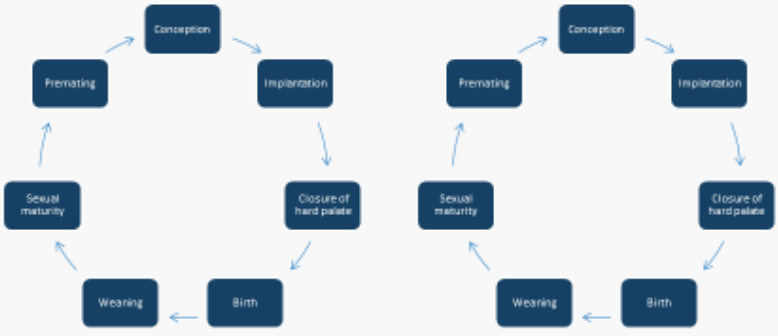
Day 21 of lactation

↓

10 weeks premating

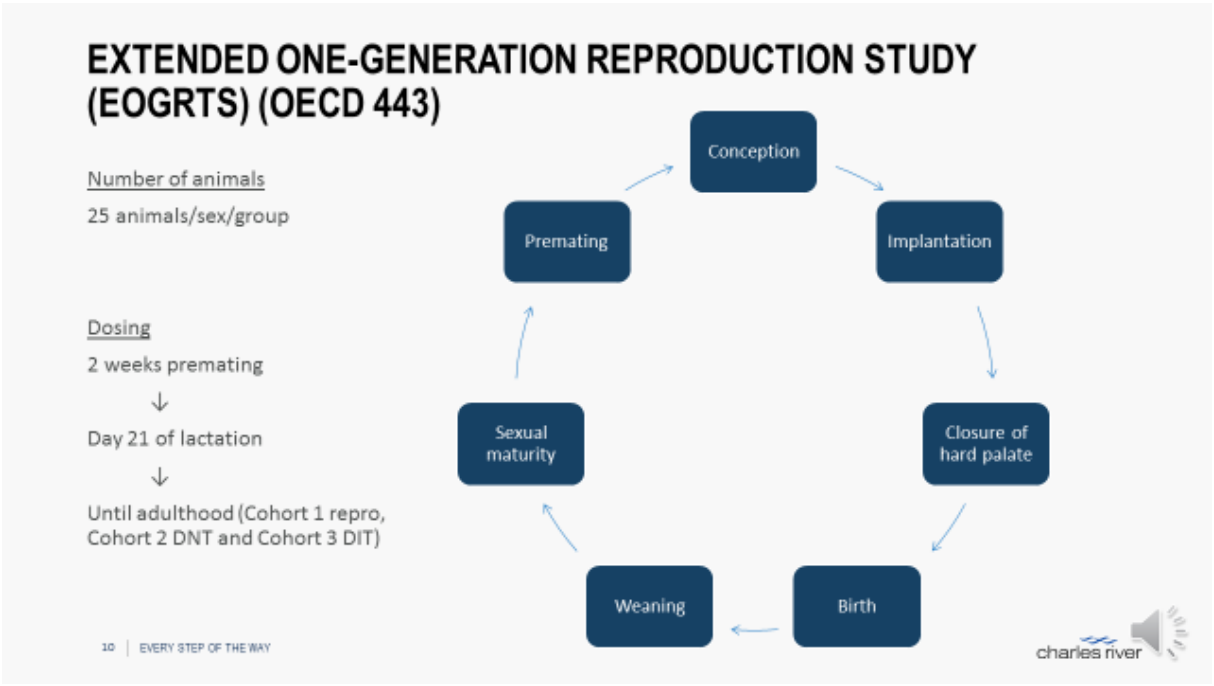
↓

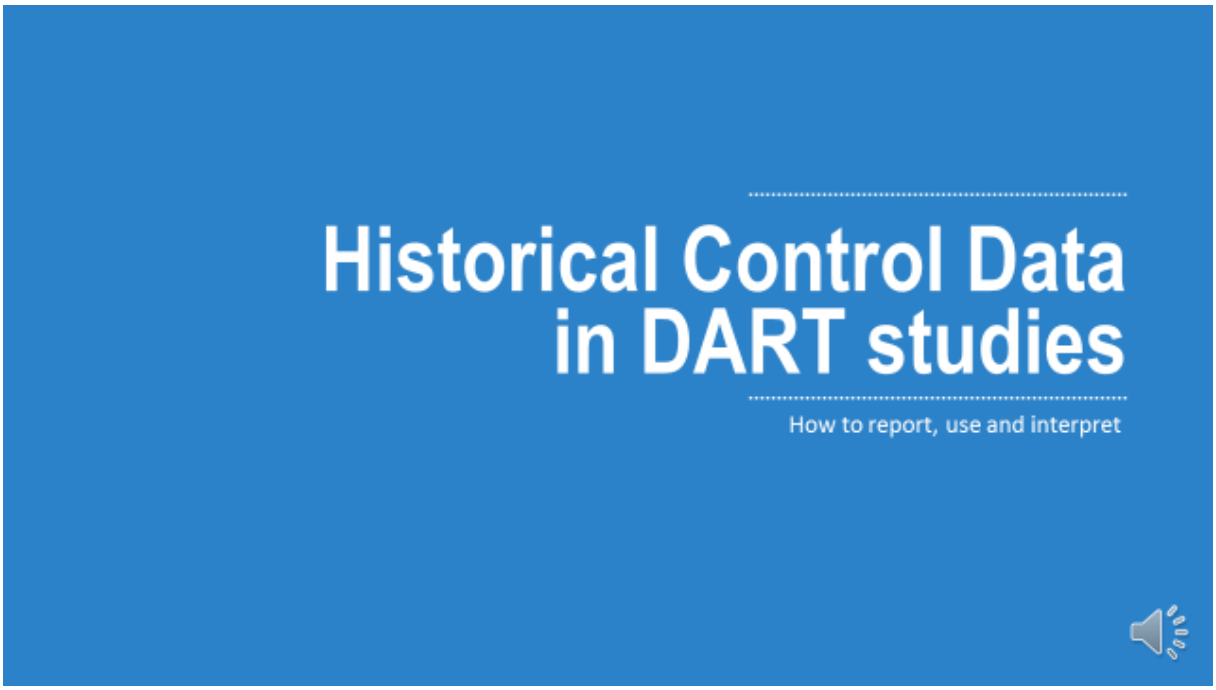
Day 21 of lactation



9 | EVERY STEP OF THE WAY

charles river


EXTENDED ONE-GENERATION REPRODUCTION STUDY (EOGRTS) (OECD 443)
Number of animals
25 animals/sex/group
Dosing
2 weeks premating
Day 21 of lactation
Until adulthood (Cohort 1 repro, Cohort 2 DNT and Cohort 3 DIT)
Conception
Implantation
Closure of hard palate
Birth
Weaning
Sexual maturity
Premating
10 | EVERY STEP OF THE WAY
charles river


Historical Control Data in DART studies
How to report, use and interpret

USE OF HISTORICAL CONTROL DATA (HCD)

The use of HCD should be viewed as a tool for developing a better understanding of the events or apparent differences observed within a study.

HCD should not be used as a convenient device for discounting unwanted or 'difficult' findings.

12 | EVERY STEP OF THE WAY

charles river

USE OF HISTORICAL CONTROL DATA

General requirements of HCD

- Same strain, age, sex and animal supplier
- Same laboratory
- Same procedures
- Performed within reasonable time period (max. 5 years)

13 | EVERY STEP OF THE WAY

charles river

USE OF HISTORICAL CONTROL DATA

Specific for DART

- It is recommended to follow international harmonized terminology developed by the International Federation of Teratology Societies (IFTS)

Review Article

Terminology of Developmental Abnormalities in Common Laboratory Mammals (Version 2)

- In general, fetal findings are categorized as malformations or variations.
  - Malformation = permanent structural change that is likely to adversely affect the survival or health of the species under investigation
  - Variation = change that occurs within the normal population under investigation and is unlikely to adversely affect survival or health

14 | EVERY STEP OF THE WAY

charles river

USE OF HISTORICAL CONTROL DATA

The litter effect

The litter (not fetus/pup) should be used as the experimental unit in DART studies.

- For continuous data (e.g., fetal weight), this is achieved by calculating a mean for each litter from the values collected for each fetus.
- For binary data (e.g., malformations), the proportion of fetuses affected in the litter is calculated by dividing the number of fetuses affected by the number of fetuses examined.
  The mean of these individual litter means, and standard deviation, are then calculated for each group on the study.

15 | EVERY STEP OF THE WAY

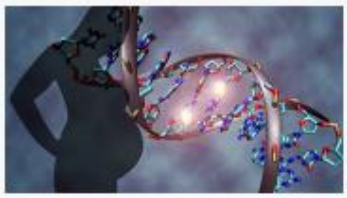
charles river

USE OF HISTORICAL CONTROL DATA

The litter effect

Important because:

- Similar findings in a single litter should be viewed of lesser concern than similar findings in isolated fetuses from several litters in a treatment group
  (as fetuses in a given litter are genetically similar and are exposed to the same maternal environment as their littermates)



- Group size would be artificially inflated if the fetus (rather than the litter) was used as the experimental unit. This could lead to invalid statistics.

16 | EVERY STEP OF THE WAY

charles river

USE OF HISTORICAL CONTROL DATA

Statistics

Results of statistical analyses alone are generally not sufficient to judge whether a study finding is a true treatment-related effect or has occurred by chance.

1. The study may not have the statistical power to detect a significant change for rare events such as malformations or highly variable data such as resorptions.
2. There is a possibility that multiple observations per study will attain statistical significance by chance alone because several hundred observations are made and analyzed in DART studies.

   At the level of statistical significance of $p \leq 0.05$, 1 of every 20 comparisons (5%) will be statistically significant by chance alone.

   Other considerations (such as the dose-dependency and HCD) should be evaluated to resolve whether or not the statistically significant finding is of biological importance.

17 | EVERY STEP OF THE WAY

charles river

USE OF HISTORICAL CONTROL DATA

Small groups

- HCD may be used more often in case of screening/pilot studies
- Important to take low N into account
- Example pregnancy rate in case of two non-pregnant rats in a group:
  - Main study (24/group), pregnancy rate is 92%
  - Screening study (10/group), pregnancy rate is 80%
  - Pilot study (6/group), pregnancy rate is 67%

18 | EVERY STEP OF THE WAY

charles river

USE OF HISTORICAL CONTROL DATA

General use

HCD may be used in **three primary ways**:

1. Identification of aberrant control values
2. Understanding relevance of low-incidence findings (e.g. malformations)
3. Understanding relevance of high-incidence findings (e.g. variations)

charles river

Identification of aberrant control values

USE OF HISTORICAL CONTROL DATA

Identification of aberrant control values

- To ensure the concurrent control group is consistent with the larger population of controls
- Important to take account of any drift in the HCD

EXAMPLE

Identification of aberrant control values

Table 3: Absolute testes weights in rats

Parameter	Control	Low dose	Mid dose (1)	Mid dose (2)	High dose
Mean absolute testes weight (g)	3.42 ± 0.18	3.02 ± 0.29*	3.06 ± 0.05**	3.10 ± 0.20*	3.17 ± 0.17
% of concurrent control	100	88.4	89.4	90.7	92.7

Number examined = 5 per group
*$p \leq 0.05$; **$p \leq 0.01$, Kruskal-Wallis, and Wilcoxon test (two sided)
HC = Historical control

- No dose response
- No statistical significant change in relative weights
- No corresponding histopathological findings

ECETOC, 2002, Technical Report No. 85. Recognition of, and differentiation between, adverse and non-adverse effects in toxicology studies

EXAMPLE

Identification of aberrant control values

Table 3: Absolute testes weights in rats

Parameter	Control	Low dose	Mid dose (1)	Mid dose (2)	High dose	HC
Mean absolute testes weight (g)	3.42 ± 0.18	3.02 ± 0.29*	3.06 ± 0.05**	3.10 ± 0.20*	3.17 ± 0.17	3.19 ± 0.15
% of concurrent control	100	88.4	89.4	90.7	92.7	93.3

Number examined = 5 per group
* $p \leq 0.05$; ** $p \leq 0.01$, Kruskal-Wallis, and Wilcoxon test (two sided)
HC = Historical control

- No dose response
- No statistical significant change in relative weights
- No corresponding histopathological findings
- Mean values of treated groups were within historical control data and concurrent control was outside this range

- Therefore considered not treatment related

23 | EVERY STEP OF THE WAY

ECETOC, 2002, Technical Report No. 85. Recognition of, and differentiation between, adverse and non-adverse effects in toxicology studies

charles river

Understanding relevance of low-incidence findings

USE OF HISTORICAL CONTROL DATA

Understanding relevance of low-incidence findings

HCD is important to evaluate rare events (e.g. malformations):

- may be spontaneous
- due to exposure to a teratogen

Developmental toxicity studies with group sizes of 16–20 litters (100–300 fetuses evaluated) do not have the statistical power to detect events that occur at frequencies of 1/1,000 to 1/10,000.

USE OF HISTORICAL CONTROL DATA

Understanding relevance of low-incidence findings

- By the nature of the low incidence, it is possible for a treated group to show a low spontaneous incidence, while the control group has lower or even no incidence.

USE OF HISTORICAL CONTROL DATA

Understanding relevance of low-incidence findings

- Omphalocele in rabbits
- Evaluation of 58 papers (4905 litters, 36.977 fetuses) → all groups
- Omphalocele was reported in 43% and was among the most common defects, occurring at a rate of 1.10% (litter) and 0.16% (fetus).

Table 3
Fetal incidence by dosage of selected classes of external malformations.

Treatment group (number of groups)	Control (71)*	Low dose (63)	Mid-dose (66)	High dose (75)
Omphalocele	18	6	16	19
Gastroschisis	2	3	1	05
Tail defect	8	5	15	45
Limb flexure	18	11	21	53
Neural tube defect	13	8	10	25
Limb defect	3	7	10	4
Edema	1	0	2	8
Anophthalmia/microphthalmia	2	4	0	0

* The number of treatment groups differs from the number of studies because of differences in study design. Some studies had additional control groups. A few studies had only one or two dose groups, lowering the number of low and mid dose groups, while a few studies had additional dose groups, which were included in the mid dose group. One study tested several compounds at a single, high dose level, and one tested a single high dose over different critical periods, adding to the numbers in the high dose group. There was also a high dose positive control compound that was added to the high dose group.

- HCD Charles River (34 studies, 668 litters, 6122 fetuses): 8 fetuses / 8 litters (mean of 0,2%)

Daston, 2018, Is omphalocele a non-specific malformation in NZW rabbits?

Understanding relevance of high-incidence findings

USE OF HISTORICAL CONTROL DATA

Understanding relevance of high-incidence findings

Important for evaluation of fetal variations in developmental toxicity studies

- In general, fetal findings are categorized as malformations or variations.
  - Malformation = permanent structural change that is likely to adversely affect the survival or health of the species under investigation
  - **Variation = change that occurs within the normal population under investigation and is unlikely to adversely affect survival or health**

USE OF HISTORICAL CONTROL DATA

Understanding relevance of high-incidence findings

Table 7: Skeletal findings (% incidence) in New Zealand White rabbits

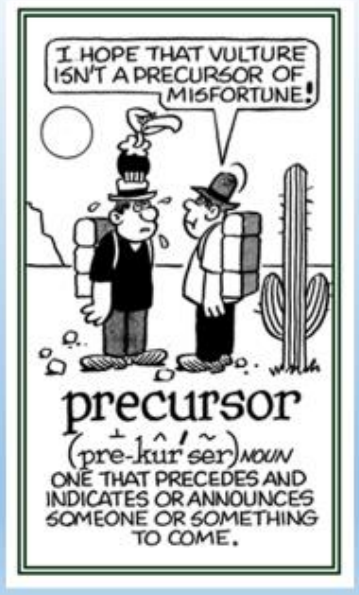
Skeletal finding	Historical Control	Control	Low dose	Mid dose	High dose
Odontoid – partially ossified	26.3-45.7	40	64.5**	61.7**	72**
Transverse process of 7th cervical vertebra partially ossified	0-6.7	6.7	0.8*	0.7*	1.0**
Transverse processes of 3rd lumbar vertebra fully ossified	2.9-13.8	8.0	0.8**	1.3*	2.5
27 Pre-sacral vertebrae	14.6-36.5	28.0	58.9**	55.1**	59.5**
Unossified 5th sternebra	2.6-13.1	12.7	3.2**	3.4**	5.2*
Partially ossified 5th sternebra	13.3-52.0	52.0	32.3	28.9**	24.6**
Partially ossified 6th sternebra	0-8.	0 8.	0 7.3	6.7	4.2
13th rib short and floating	4.3-14.0	4.0	5.6*	2.7**	5.9*
13th rib normal length	17.1-55.2	42.0	78.2**	82.6**	81.4**

*p≤ 0.05, ** p ≤ 0.01, Student's t test

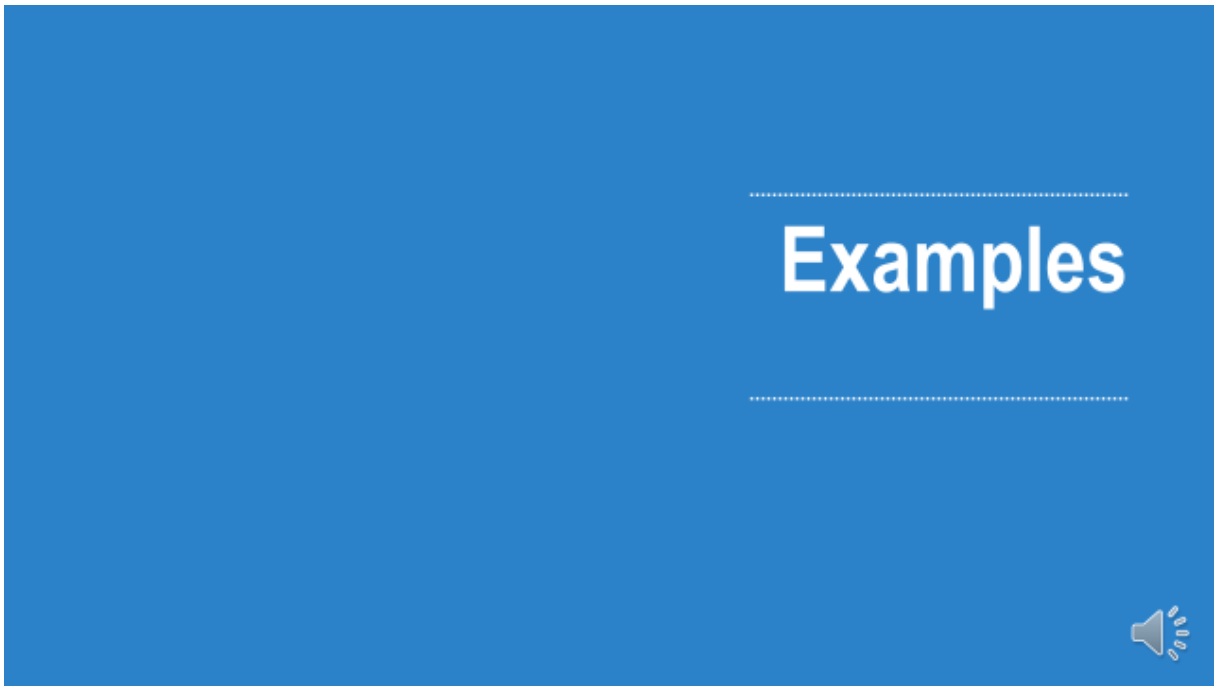
The shaded data are those dismissed as effects of treatment as they fall within the range expected for untreated animals.

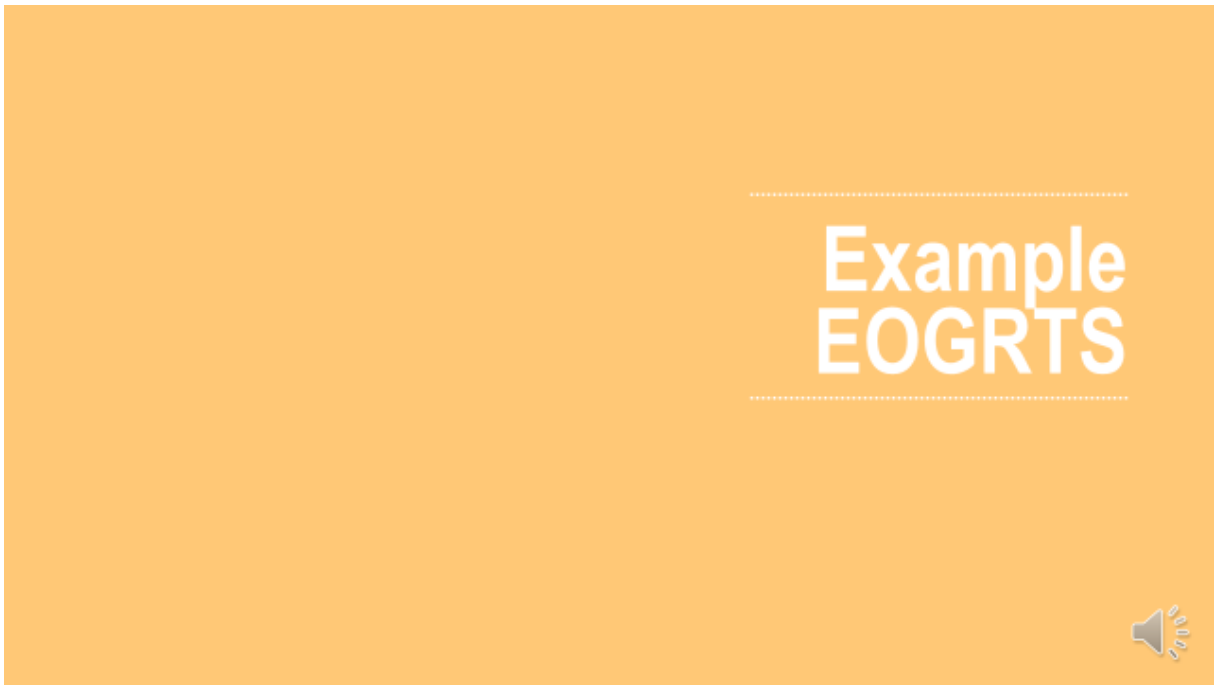
ECETOC, 2002, Technical Report No. 85. Recognition of, and differentiation between, adverse and non-adverse effects in toxicology studies

charles river



charles river


Examples


Example
EOGRTS

USE OF HISTORICAL CONTROL DATA

Example EOGRTS

Mean number of implantation sites:

Control	Low dose	Mid dose	High dose
13.1	12.7	12.4	10.8^{++}

+/++ Significant at 5% (+) or 1% (++) level

USE OF HISTORICAL CONTROL DATA

Example EOGRTS

Mean number of implantation sites:

Control	Low dose	Mid dose	High dose	HCD (min-max)
13.1	12.7	12.4	10.8^{++}	12.1 (11.3-12.8)

+/++ Significant at 5% (+) or 1% (++) level

- HCD:
  - Concurrent control mean slightly high
  - High dose mean slightly low
- Considered treatment-related
- But not adverse, as individual values were all within normal limits (next page)

Control	Low dose	Mid dose	High dose
13	15	12	13
11	11	15	10
16	14	13	12
15	9	13	14
16	12	10	10
16	16	12	12
13	10	10	9
11	14	9	10
13	4	11	11
13	14	16	12
12	12	10	12
10	15	15	9
14	14	10	6
15	13	11	11
14	14	13	11
12	13	13	11
9	15	14	11
	13	13	10
		15	11

Example Sex ratio

USE OF HISTORICAL CONTROL DATA

Example sex ratio

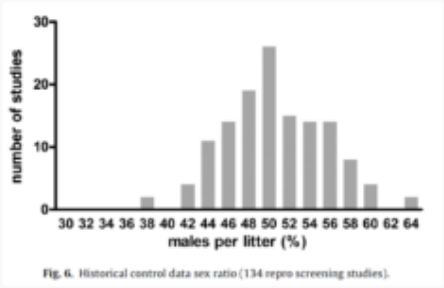


Fig. 6. Historical control data sex ratio (134 repro screening studies).

Beekhuijzen et al, 2014, The underestimated value of OECD 421 and 422 repro screening studies: Putting it in the right perspective

Living pups at first litter check % of males / females	Group 1	Group 2	Group 3	Group 4
	51 / 49	45 / 55	49 / 51	62 / 38

Not statistically significant, but Group 4 on border of HCD range

58 | EVERY STEP OF THE WAY

charles river

USE OF HISTORICAL CONTROL DATA

Example sex ratio

Group 1		Group 2		Group 3		Group 4	
M	F	M	F	M	F	M	F
6	6	5	3	7	5	11	3
7	6	4	8	4	6	8	4
5	6	3	5	7	5	8	6
11	3	5	5	4	7	6	4
7	5	3	7	6	9	7	6
6	6	4	7	11	2	6	4
6	4	7	3	5	8	2	2
3	8	5	8	6	7	8	5
6	7	9	5	4	4		
3	7			7	4		

39 | EVERY STEP OF THE WAY

charles river

USE OF HISTORICAL CONTROL DATA

Example sex ratio

Group 1		Group 2		Group 3		Group 4	
M	F	M	F	M	F	M	F
6	6	5	3	7	5	11	3
7	6	4	8	4	6	8	4
5	6	3	5	7	5	8	6
11	3	5	5	4	7	6	4
7	5	3	7	6	9	7	6
6	6	4	7	11	2	6	4
6	4	7	3	5	8	2	2
3	8	5	8	6	7	8	5
6	7	9	5	4	4		
3	7			7	4		

40 | EVERY STEP OF THE WAY

charles river

Example
Thyroid hormones

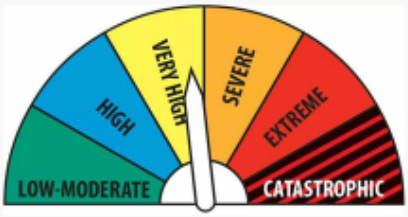
CHANGES IN THYROID HORMONES

- One of the conclusions from the 2017 SOT roundtable discussion was that *"in the absence of definitive biomarkers of altered neurodevelopment, general agreement exists that changes in TH are appropriate starting points for risk assessment to protect against potential downstream effects on neurodevelopment"*.

CHANGES IN THYROID HORMONES

- One of the conclusions from the 2017 SOT roundtable discussion was that *"in the absence of definitive biomarkers of altered neurodevelopment, general agreement exists that changes in TH are appropriate starting points for risk assessment to protect against potential downstream effects on neurodevelopment".*

- *"Disagreement persists regarding the magnitude of TH perturbation that could result in developmental hazard classification and labeling, which does not take into consideration dose response and exposure information".*



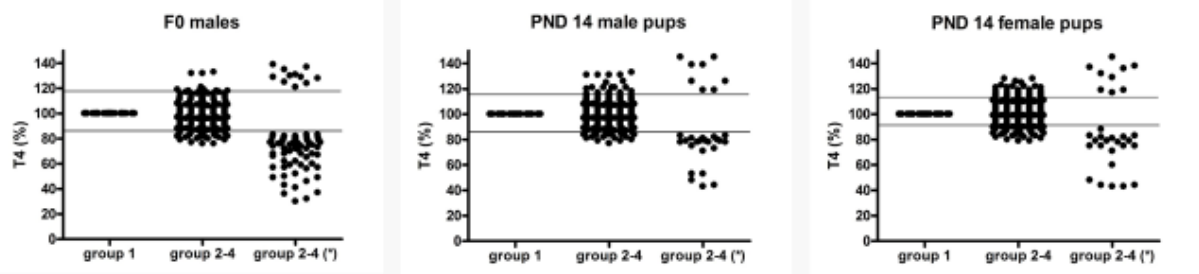
43 | EVERY STEP OF THE WAY

charles river

MAGNITUDE OF THYROID HORMONE PERTURBATION

Retrospective evaluation of 124 repro screening studies (OECD 421/422)

- In general, a statistical significant finding for T4 occurs beyond a 20% difference compared to the concurrent control group.



Relative T4 levels per study (i.e. concurrent control levels are set at 100%), clustered for the control group (Group 1; left panel), all treated groups without statistical significant effect (Group 2–4; middle panel) and all treated groups with statistical significant effect (p < 5%; Group 2–4(*); right panel).

44 | EVERY STEP OF THE WAY

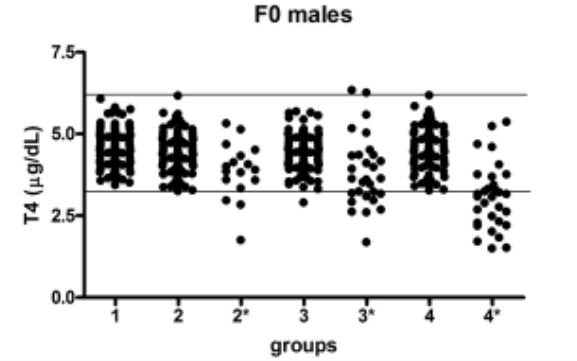
Beekhuijzen et al, 2019, A critical evaluation of thyroid hormone measurements in OECD test guideline studies: Is there any added value?

charles river

MAGNITUDE OF THYROID HORMONE PERTURBATION

Retrospective evaluation of 124 repro screening studies (OECD 421/422)

- Most T4 levels of Groups 2-4, although statistically significant, fall within the range of the control group means (Group 1).
- Due to the high variability in T4 levels.



Group mean values per study of F0-males for total T4 (µg/dL). The seven plots are T4 values in Group 1 (1), not statistically significant T4 values for Group 2, Group 3 and Group 4 (2, 3 and 4, respectively), and statistically significant T4 values for Group 2, Group 3 and Group 4 (p<5%; 2, 3* and 4*, respectively).*

45 | EVERY STEP OF THE WAY

Beekhuijzen et al, 2019, A critical evaluation of thyroid hormone measurements in OECD test guideline studies: Is there any added value?

charles river

MAGNITUDE OF THYROID HORMONE PERTURBATION

Retrospective evaluation of 124 repro screening studies (OECD 421/422)

- High frequency probably due to slight disturbances of normal homeostasis, leading to hormonal fluctuations. Moreover:
  - No pretest value
  - Only 1 measurement during treatment
  - Within toxicity study
- Therefore, possible treatment related effect should not be based on statistics only; **HCD should be taken into account.**

Frequency of statistical significant findings ($p < 5\%$) for T4 levels in F_0-males for 124 repro screening studies.

		T4 level statistically significant compared to concurrent control T4 level	
		No. of studies affected	Percentage
Occurence within a study	Group 2 only	5	4,0%
	Group 3 only	9	7,3%
	Group 4 only	11	8,9%
	Group 3 + 4	10	8,1%
	Group 2 + 3 + 4	8	6,5%
	Group 2 + 4	2	1,6%
	Group 2 + 3	2	1,6%
	Total	47	37,9%
All 124 studies combined	Group 2, total	17	13,7%
	Group 3, total	29	23,4%
	Group 4, total	31	25,0%

46 | EVERY STEP OF THE WAY

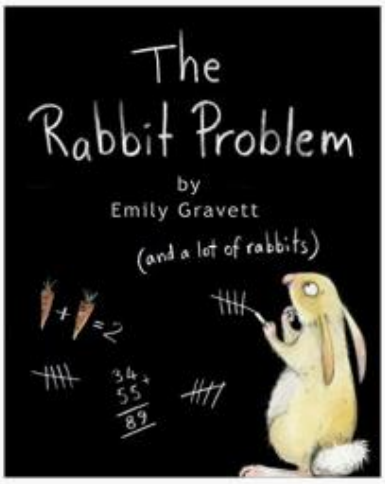
Beekhuijzen et al, 2019, A critical evaluation of thyroid hormone measurements in OECD test guideline studies: Is there any added value?

charles river


Example
Rabbits

DEVELOPMENTAL TOXICITY STUDY

In rats and rabbits

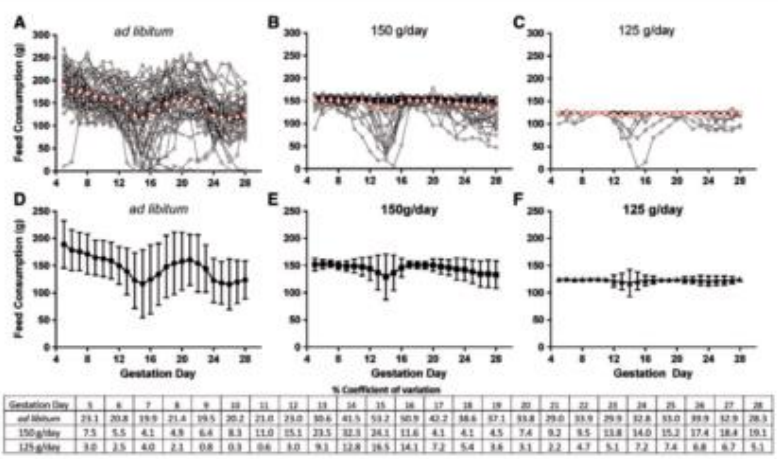


- Several challenges.
- Regular occasion of gastrointestinal toxicity in rabbits.
- Can lead to abortions.

48 | EVERY STEP OF THE WAY

charles river

HIGH VARIATION FOOD CONSUMPTION IN CONTROL RABBITS



Gestation Day	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28
ad libitum	23.1	20.8	19.9	21.4	19.5	20.2	21.0	23.0	30.6	41.5	53.2	50.9	42.2	38.6	37.1	33.8	29.0	31.9	29.9	32.8	33.0	39.9	32.9	28.3
150 g/day	7.5	5.5	4.1	4.9	6.4	8.3	11.0	15.1	23.5	32.3	24.1	11.6	4.1	4.1	4.5	7.4	9.2	9.5	13.8	14.0	15.2	17.4	18.4	19.1
125 g/day	3.0	2.5	4.0	2.1	0.8	0.3	0.6	3.0	9.1	12.8	16.5	14.1	7.2	5.4	3.6	3.1	2.2	4.7	5.1	7.2	7.4	6.8	6.7	5.1

FIG. 1. Feed consumption during GD 5–27 in untreated NZW rabbit does administered modified control diet (**A**) *ad libitum* (Study nos. 3–4), (**B**) 150 g/rabbit/day (Study nos. 1–2), or (**C**) 125 g/rabbit/day (Study nos. 5–6). Each data point represents a feed consumption measurement for an individual rabbit (n = 48–51 rabbits). The bold line represents the mean feed consumption. Panels D–F show the mean (SD) feed consumption for each group. The table includes the CV% for feed consumption of each feed level across each measured GD.

49 | EVERY STEP OF THE WAY

Hannas et al, 2016

charles river

TOXICITY DUE TO OILY VEHICLES

Feeding of vegetable fats and seeds lead to motility and functional depression of the GI tract in rabbits [*Johnson-Delaney, 2006*].

Example 2 ml/kg corn oil:

1.5 FOOD CONSUMPTION (G/ANIMAL/DAY)
FEMALES

F0-GENERATION

	POST COITUM									
DAYS ANIMAL	6-7	7-8	8-9	9-10	10-11	11-12	12-13	13-14	14-15	15-16
GROUP 1 (CONTROL)										
1	3	6	26	33	28	81	101	76	67	39
2	0	9	22	41	34	63	25	9	4	11
3	0	6	37	42	44	29	28	47	28	0
4	1	14	39	82	58	56	63	54	49	55
5	0	0	27	18	7	15	8	6	3	0
6	0	40	76	15	40	54	56	8	2	2

Important is to use a low dose volume so the amount of vehicle being dosed is limited.

50 | EVERY STEP OF THE WAY

charles river

USE OF HCD

51 | EVERY STEP OF THE WAY

charles river


Questions?
Manon.beekhuijzen@crl.com

3. Establishment and Use of Historical Control Data in Clinical Pathology

(Volker Strauss, on behalf of ESTP)

9th ESTP International Expert Workshop
"Assessment of Toxicological Relevance of Clinical Pathology Changes"

Establishment and Use of Historical Control Data in Clinical Pathology

Volker Strauss, BASF SE, Ludwigshafen, Germany
(volker.strauss@basf.com)

International Workshop on how to report, use and interpret historical control data in (eco)toxicity studies
3rd – 5th May 2022, virtual event

9th ESTP International Expert Workshop
"Assessment of Toxicological Relevance of Clinical Pathology Changes"

Introduction

- Parts of the presentation were presented at the 9th ESTP International Expert Workshop: Assessment of the Biological/Toxicological Relevance of Clinical Pathology Changes, 5th/6th Apr 2022
  https://www.eurotoxpath.org/meetings/index.php?id=workshop9
- New slides are marked (*)
- Disclosure: content of the new slides is not the official opinion of ESTP but the personal thoughts of the author

9th ESTP International Expert Workshop
"Assessment of Toxicological Relevance of Clinical Pathology Changes"

Who is member of the ESTP?

- ESTP: European Society of Toxicologic Pathology e.V.
- www.eurotoxpath.org
- Members: university graduate and working in the field of toxicologic pathology
- Goal: protecting man against harmful effects that may result from the intended use of active ingredients or additives, or which may be due to toxins at the workplace or in the environment

9th ESTP International Expert Workshop
"Assessment of Toxicological Relevance of Clinical Pathology Changes"

Content(*)

- Aims, issues and procedures of establishing historical control data (HCD)
- Minimum study numbers for establishing HCD intervals
- Identification of outlier values among HCD
- Increasing numbers of studies used for HCD
    - Enlarging time interval of studies used for HCD
    - Combining control groups of similar study types
- Thyroid hormone HCD
- Proposals for the use and establishment of HCD

9th ESTP International Expert Workshop
"Assessment of Toxicological Relevance of Clinical Pathology Changes"

Aims for using historical control intervals

- Evaluating if **study control data or pre-study values** are **outside historical control values**
- Assessing whether **statistical significant changes** among study values are **within normal variation** of historical control data (HCD)
- Assessing a **shift of clinical pathology values** in studies of the tox facility

9th ESTP International Expert Workshop
"Assessment of Toxicological Relevance of Clinical Pathology Changes"

Main issues for confidence in HCD

- Objections of regulatory bodies
  - HCD are not fitting to the submitted study because of **differences in study details**
  - Historical control **intervals are too great**, so that many statistically significant changes in a study may be discussed as within the normal variation

- To avoid this impression
  - Regulatory guidances (or scientific publications) for establishing historical control intervals
  - **Documented procedure** how to establish historical control intervals (e.g., SOP)
  - **Reasonable statistical methods** for calculating historical control intervals (e.g., ASVCP Guideline, 2011)
  - Consistent, **objective interpretation** of study values within or out of HCD intervals

9th ESTP International Expert Workshop
"Assessment of Toxicological Relevance of Clinical Pathology Changes"

Different procedures of establishing HCD

- Collecting **individual values**
  - Comparing mean and percentile range (confidence intervals) with study values
  - Above all for dog/non-human primates (NHP) studies with individual value interpretation
- Collecting **means/medians of prior control groups**
  - HCD interval established with min/max values after exclusion of outliers
  - Rationale: in rat/mice studies means/medians of dose groups are compared with study controls and therefore same procedure for HCD

9th ESTP International Expert Workshop
"Assessment of Toxicological Relevance of Clinical Pathology Changes"

Time interval of studies for establishing historical control data

- 5 years before actual study (industry standard)
- **2.5 years before until 2.5 years after actual study** (JMPR Guidance, 2015, focusing on tumor incidences in carcinogenicity studies; not feasible for clinical pathology data)
- Greater time interval if < 10 studies (or less than 20 individual values) are included (trend analysis)
- Regular or continous re-freshment of HCD, dependent on HCD database (and when changed method or species strain)

9th ESTP International Expert Workshop
"Assessment of Toxicological Relevance of Clinical Pathology Changes"

Considering number of studies and outlier values?

Without outlier statistics

Parameter:		WBC	WBC	WBC	WBC
Unit:		GIGA/L	GIGA/L	GIGA/L	GIGA/L
Study	sampling month				
1	09/10				2.22
2	09/10				1.86
3	09/10				2.98
4	09/10				2.96
5	10/10				1.53
6	03/11			1.32	1.32
7	09/11			3.31	3.31
8	06/13			2.99	2.99
9	06/14			2.63	2.63
10	08/14			3.60	3.60
11	08/15		4.60	4.60	4.60
12	08/15		3.10	3.10	3.10
13	03/16		1.65	1.65	1.65
14	12/16		2.20	2.20	2.20
15	04/17		4.68	4.68	4.68
16	10/18	2.21	2.21	2.21	2.21
17	10/18	2.70	2.70	2.70	2.70
18	11/18	2.33	2.33	2.33	2.33
19	04/19	2.68	2.68	2.68	2.68
20	01/20	3.78	3.78	3.78	3.78
N		5	10	15	20
Mean		2.74	2.99	2.92	2.77
Minimum		2.21	1.65	1.32	1.32
Maximum		3.78	4.68	4.68	4.68

With outlier statistics

Parameter:		WBC	WBC	WBC	WBC
Unit:		GIGA/L	GIGA/L	GIGA/L	GIGA/L
Study	sampling month				
1	09/10				2.22
2	09/10				1.86
3	09/10				2.98
4	09/10				2.96
5	10/10				1.53
6	03/11			1.32	1.32
7	09/11			3.31	3.31
8	06/13			2.99	2.99
9	06/14			2.63	2.63
10	08/14			3.60	3.60
11	08/15		4.60	4.60	*4.60
12	08/15		3.10	3.10	3.10
13	03/16		1.65	1.65	1.65
14	12/16		2.20	2.20	2.20
15	04/17		4.68	4.68	*4.68
16	10/18	2.21	2.21	2.21	2.21
17	10/18	2.70	2.70	2.70	2.70
18	11/18	2.33	2.33	2.33	2.33
19	04/19	2.68	2.68	2.68	2.68
20	01/20	*3.78	3.78	3.78	3.78
N (wo. outliers)		4	10	15	18
Mean		2.48	2.99	2.92	2.56
Tukey's interquartile fences:		2.03 - 3.54	0.19 - 5.67	0.48 - 5.23	0.79 - 4.57
Minimum (wo. outliers)		2.21	1.65	1.32	1.32
Maximum (wo. outliers)		2.70	4.68	4.68	3.78

- Mean white blood cell (WBC) counts are shown
- Female C57BL/6 J Rj mice
- 10 weeks old
- 28d dietary study
- 10-16 hours fasted
- Isoflurane anesthesia
- Retro-bulbar blood sampling
- ADVIA120 instrument
- ***x.xx = outlier**

- Too few studies used for HCDs may result in a skewed HCD interval
- Without outlier statistics risk of too great HCD intervals

9th ESTP International Expert Workshop
"Assessment of Toxicological Relevance of Clinical Pathology Changes"

Minimum sample size for establishing reference intervals

Sample size	Data distribution (innate or transformed)	Statistical method
≥ 120	Not applicable	Nonparametric with 90% CI of ref. limits
40 ≤ x <120	Gaussian	Robust with 90% CI of ref. limits
		Parametric with 90% CI of ref. limits
	Non-Gaussian	Robust with 90% CI (preferred) of ref. limits
		Nonparametric[a]
20 ≤ x <40	Gaussian	Parametric with 90% CI of ref. limits[b]
	Non-Gaussian	Robust with 90% CI of ref. limits[b]
10 ≤ x < 20	Not applicable	Do not calculate reference intervals[b]
< 10	Not applicable	Do not report reference values

according to ASVCP Guideline, 2011

Confidence interval (CI)
[a]Cannot determine 90% CI nonparametrically with <120 reference sample, alternative methods required, e.g., bootstrap.
[b]Include the following information: histogram, mean or median, minimum and maximum

Similar recommendations in Reference Value Advisor, Ecole Nationale Vétérinaire, Toulouse, France (Geffre et al, 2011)

9th ESTP International Expert Workshop
"Assessment of Toxicological Relevance of Clinical Pathology Changes"

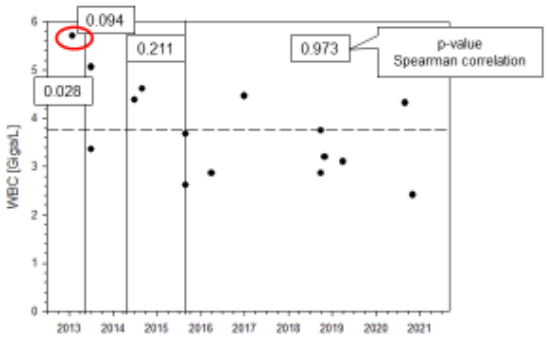
What can be done in case of too few studies for establishing HCD? (*)

- **Enlarging time interval** of studies used for HCD of more than 5 years (after trend analysis)
- Combining control groups of **similar study types** (e.g., dietary and gavage studies)

9th ESTP International Expert Workshop
"Assessment of Toxicological Relevance of Clinical Pathology Changes"

Trend analysis for including HCD > 5 years (*)

Date	WBC [giga/L]	N	p-value
Jan 13	5.70	15	0.028
Jun 13	5.06	14	0.094
Jun 13	3.36	13	0.241
Jun 14	4.38	12	0.211
Aug 14	4.61	11	0.440
Aug 15	3.67	10	0.973
Aug 15	2.62	9	0.881
Mrz 16	2.86	8	0.713
Dez 16	4.46	7	0.268
Okt 18	2.87	6	0.742
Okt 18	3.74	5	0.505
Nov 18	3.20	4	0.600
Apr 19	3.10	3	0.667
Sep 20	4.32		
Nov 20	2.41		



- Risk of too broad reference intervals by including older HCD
- Step-wise backward Spearman rank correlation for excluding HCD with p-value <0.05 showing a trend of the data

Date	Jan. 13	Jun. 13	Jun. 13	Jun. 14	Aug. 14	Aug. 15	Aug. 15	Mrz. 16	Dez. 16	Okt. 18	Okt. 18	Nov. 18	Apr. 19
N	15	14	13	12	11	10	9	8	7	6	5	4	3
Mean WBC	3.76	3.62	3.51	3.52	3.44	3.33	3.29	3.37	3.44	3.27	3.35	3.25	3.28
Minimum WBC	2.41	2.41	2.41	2.41	2.41	2.41	2.41	2.41	2.41	2.41	2.41	2.41	2.41
Maximum WBC	5.70	5.06	4.61	4.61	4.61	4.46	4.46	4.46	4.46	4.32	4.32	4.32	4.32
Max - Min	3.29	2.65	2.20	2.20	2.20	2.05	2.05	2.05	2.05	1.91	1.91	1.91	1.91

- Mean white blood cell (WBC) counts are shown
- Male C57BL/6 J Rj mice
- 10 weeks old
- 28d dietary study
- 10-16 hours fasted
- Isoflurane anesthesia
- Retro-bulbar blood sampling
- ADVIA120 instrument

9th ESTP International Expert Workshop
"Assessment of Toxicological Relevance of Clinical Pathology Changes"

Combining studies to increase sample size for establishing HCD

Consider:

- Animal strain, animal supplier
- Sex
- **Age**
- **Administration route** (and vehicle, diet)
- Diet/fasting
- Single-housed or group-housed
- Blood sampling site
- Anesthesia method
- **Parameter measurement method** (instrument change)
- Others

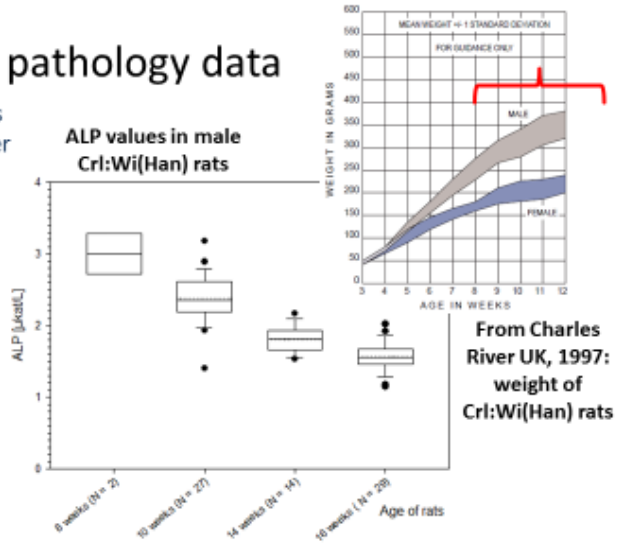
9th ESTP International Expert Workshop
"Assessment of Toxicological Relevance of Clinical Pathology Changes"

Age dependency of clinical pathology data

- Following values of 40 clinical pathology parameters were different in 10-week old rats compared to older ones (Student-Newman-Keuls test):
  - HGB, absolute NEUT, UREA, CREA, GLUC, TBIL, PROT, GLOB were lower
  - PLT, absolute LYMPH and BASO, INP, ALP were higher

- Age dependency of clin path values above all in the growing phase of young rats have to be considered

  - Oral 14-day rat studies; study control mean values were compared
  - Box plots: solid median lines, dotted mean lines, box length: 25 and 75 percentiles, whiskers: 10 and 90 percentiles, single dots: outliers
  - N = number of studies with 4 to 5 rats per group



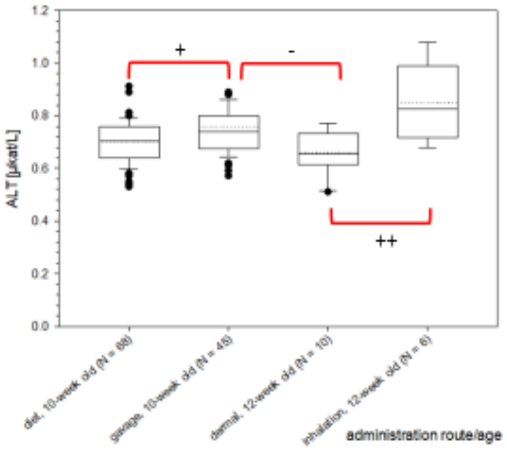
From Charles River UK, 1997: weight of Crl:Wi(Han) rats

9th ESTP International Expert Workshop
"Assessment of Toxicological Relevance of Clinical Pathology Changes"

Administration route dependency of clinical pathology data

- Comparison of HCD of 40 clinical pathology parameters established with study control means
- 28d oral gavage studies resulted in statistically significant higher ALT, AST and globulin values in males and females compared to dietary studies. However, differences are marginal
- All other hematology and clinical chemistry parameters were not different between gavage and dietary studies
- 28d inhalation studies resulted in higher ALT, AST and triglyceride values compared to dermal studies
- No difference between dermal and oral administration studies
  - Box plots: solid median lines, dotted mean lines, box length: 25 and 75 percentiles, whiskers: 10 and 90 percentiles, single dots: outliers
  - Student-Newman-Keuls test: - $p > 0.05$; + $p \leq 0.05$; ++ $p \leq 0.01$
  - N = number of studies with 5 to 10 rats per group

ALT values in 28-day studies with male Crl:Wi(Han) rats



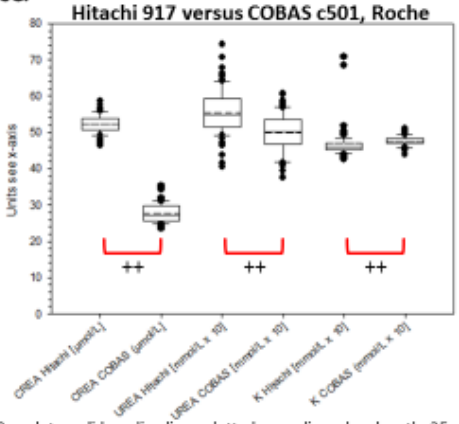
9th ESTP International Expert Workshop
"Assessment of Toxicological Relevance of Clinical Pathology Changes"

Laboratory method dependency of clinical pathology data

- 90d rat studies (N = 10 male rats per group, study mean values)
- 21 clinical chemistry values were compared
- Aug 2013, **Hitachi 917**, Roche replaced by **COBAS c501**, Roche
- Hitachi values: Mar 2003 – Jul 2013 (N = 81); COBAS values: Aug 2013 – Apr 2021 (N = 70)
- Creatinine: Hitachi: Jaffé, COBAS: enzymatic method
- Potassium: both instruments: indirect ion-selective electrodes
- Urea: both instruments: kinetik Urease/GLDH test
- Potential bias because of different time interval of studies for comparison

- For creatinine new HCD collection necessary after method change
- For urea and potassium
  - significant differences although overlap of 25. – 75. percentiles
  - During validation (20 samples per sex) no difference was observed

Male Crl:Wi(Han) rats from 90d studies: Hitachi 917 versus COBAS c501, Roche



- Box plots: solid median lines, dotted mean lines, box length: 25 and 75 percentiles, whiskers: 10 and 90 percentiles, single dots: outliers
- Mann-Whitney-U-test: ++ $p \leq 0.01$

9th ESTP International Expert Workshop
"Assessment of Toxicological Relevance of Clinical Pathology Changes"

Interlaboratory variance of clinical pathology data

- Male Crl:Wi(Han) rats; individual hematology data of controls are shown
- Hematology measured with **ADVIA 120**, Siemens
- Age was different, but both groups were adult rats
- Bias may be possible because of different time intervals (2008 versus 2013-2018)

		Charles River, 2008, 17 weeks and older (N = 167)				BASF, 2013 - 2018, 18 weeks old (N = 669)			
Parameter	Unit	Mean	SD	Mean -1.95 SD	Mean + 1.95 SD	Mean	SD	Mean -1.95 SD	Mean + 1.95 SD
Hematocrit	L/L	0.442	0.034	0.376	0.508	0.423	0.022	0.380	0.466
Hemoglobin	mmol/L	9.6	0.6	8.4	10.8	9.0	0.5	8.0	10.0
RBC	Tera/L	8.69	0.66	7.40	9.98	8.55	0.51	7.56	9.54
Reticulocytes	Giga/L	166.5	28.8	110.3	222.7	139.8	27.0	87.2	192.5
WBC	Giga/L	4.28	2.14	0.11	8.45	5.64	1.58	2.56	8.72
Neutrophils	Giga/L	0.87	0.42	0.05	1.69	1.14	0.36	0.44	1.84
Lymphocytes	Giga/L	3.19	1.89	0.00	6.88	4.24	1.43	1.45	7.03

- Above all mean reticulocyte counts and low borders of WBC, neutrophil and lymphocyte counts were different

9th ESTP International Expert Workshop
"Assessment of Toxicological Relevance of Clinical Pathology Changes"

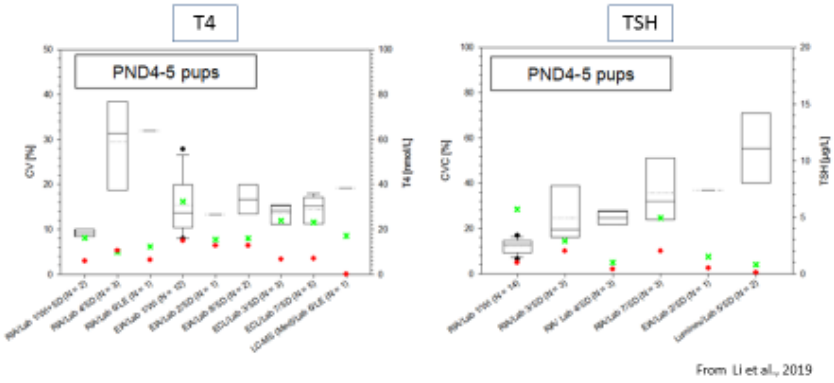
T4 and TSH measurement in rat toxicity studies (*)

- Total T4:
  - Human medicine reference method: isotope dilution-liquid chromatography/tandem mass spectrometry (ID-LC/tandem MS)

    (International Federation of Clinical Chemistry and Laboratory Medicine IFCC, IFCC Scientific Division Working Group for Standardization of Thyroid Function Tests WG-STFT)
  - **Routine method:** immunoassays (RIA or ELISA)
- Thyroid stimulating hormone (TSH):
  - No reference method; **routine method:** immunoassays
  - Human medicine: harmonization of TSH measurements standardized against the 2nd IRP WHO Reference Standard 80/558 (Roche, Elecsys)
  - No international reference standard for rat TSH
- Allowed inter-individual variation (CV) in study control groups: for **T3/T4 25%**; for **TSH 35%** (OECD TG 408, 2018)
- Critical issues with hormone measurement in toxicity studies: BfR expert hearing (Kucheryavenko et al., 2019)
- Different HCD intervals for (T3), T4 and TSH in rats according mandatory measurements in OECD TG
  - Different ages: PND4, PND13, PND22, PND90, adult rats
  - For female rats: nulliparous rats, pregnant rats end of pregnancy, rats during lactation (LD14)

9th ESTP International Expert Workshop
"Assessment of Toxicological Relevance of Clinical Pathology Changes"

Comparison of mean study control T4 and TSH values of various laboratories (*)



From Li et al., 2019

- T4/TSH measurement in rat fetuses/pups is challenging
  - Minimum sample volume
  - Low T4 values compared to adults
- Most labs reach CV <25% for T4 and <35% for TSH (OECD TG 408) also for HCD

CV[%], left y-axis
T4/TSH mean, right y-axis
T4/TSH LOQ, right y-axis
N = number of control groups; Wi = Wistar rats; SD = Sprague-Dawley rats; LOQ = limit of quantification

9th ESTP International Expert Workshop
"Assessment of Toxicological Relevance of Clinical Pathology Changes"

Proposals for establishment of clinical pathology HCD

- Using studies **5 years before the actual study** for calculation of the HCD
  - If only few studies, expand the time interval (trend analysis)
- Exclude statistically identified **outliers** (outlier statistics, use of percentiles)
- Combine controls of studies with **not exactly the same design** (e.g. different administration route etc.) not before confirming that HCD are not affected by different design
- **Don't use HCD established in external labs or textbook data** for arguing that significant changes in toxicity studies are within the normal variation, also when the parameters were measured with the same methods/instruments

9th ESTP International Expert Workshop
"Assessment of Toxicological Relevance of Clinical Pathology Changes"

Proposals for the use of clinical pathology HCD

- The **concurrent control group of the study or pre-study values** remain the most relevant comparator to determine test article related effects
- HCD are an adjunct to sound scientific judgement
- HCD enable evaluation of **outlying study control or pre-test data**
  - It should trigger further investigation of the potential factor causing this
- HCD provide information on the **normal level of biological variation** in order to assess if a statistically significant change in a study is within normal variation or not
- The utility of HCD is **limited to the test site** where the HCD were generated

9th ESTP International Expert Workshop
"Assessment of Toxicological Relevance of Clinical Pathology Changes"

References

- ASVCP Quality Assurance and Laboratory Standards Committee (QALS): Guidelines for the Determination of Reference Intervals in Veterinary Species and other related topics (2011), http://www.asvcp.org/pubs/pdf/RI%20Guidelines%20For%20ASVCP%20website.pdf
- Reference values: a review
  Geffré, A. et al. Reference values: a review, Vet Clin Pathol 38/3 (2009) 288–298
- Geffré A. et al. Reference Value Advisor: a new freeware set of macroinstructions to calculate reference intervals with Microsoft Excel, Vet Clin Pathol 40/1 (2011) 107–112
- Kluxen, F.M. et al. Using historical control data in bioassays for regulatory toxicology. Regulatory Toxicology and Pharmacology Volume 125, October 2021
- Kucheryavenko, O. et al. Report from the BfR expert hearing on practicability of hormonal measurements: recommendations for experimental design of toxicological studies with integrated hormonal end points, Arch Toxicol (2019), 93(4):1157-1167
- Li, A. et al. Practical Considerations for Developmental Thyroid Toxicity Assessments: What's working, What's not, and How can we do better? Regulatory Pharmacology and Toxicology, (2019), 106,111-136

volker.strauss@basf.com

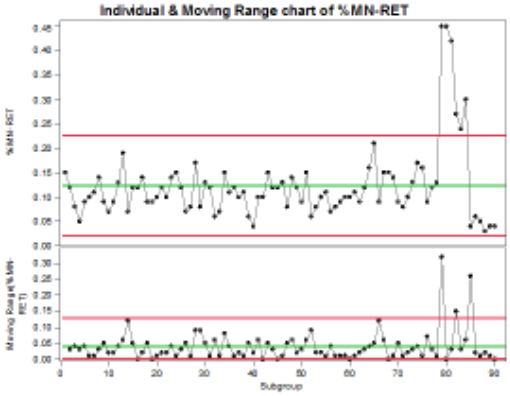
4. Assessing the Quality of Historical Control Distributions and Calculating Useful Intervals: Genetic Toxicology Examples

(Stephen Dertinger, Litronlabs)

Assessing the Quality of Historical Control Distributions and Calculating Useful Intervals: Genetic Toxicology Examples

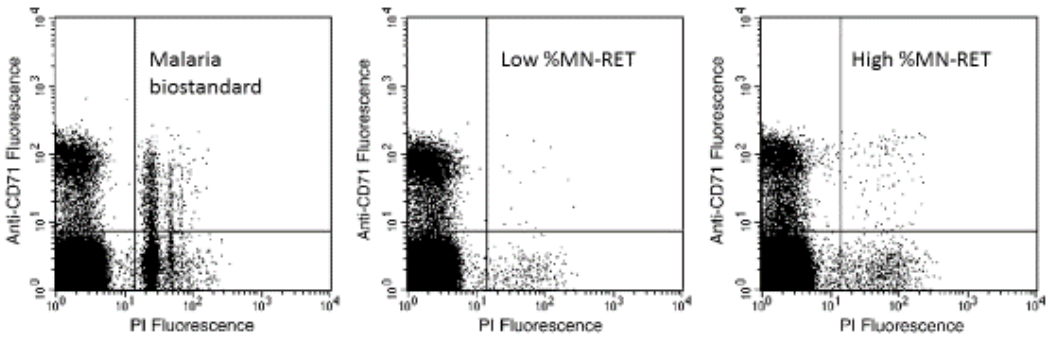
EFSA Workshop, May 4, 2022

Stephen D. Dertinger, Ph.D.
Director of Research
Litron Laboratories
sdertinger@litronlabs.com



Disclosure

- S.D. works for Litron Laboratories, a company that sells reagent kits and offers testing services based on flow cytometric analysis of genetic toxicology endpoints, including the *in vivo* micronucleus assay that will be discussed here



Acknowledgements

• This presentation draws heavily from two ongoing efforts:

Development of an OECD TG for the *Pig-a* Gene Mutation Assay

Special thanks to:
Lead author Bob Heflich, FDA-NCTR
David Lovell, St. George's University of London
Carol Gleason, BMS

Bob Heflich — Carol Gleason — David Lovell

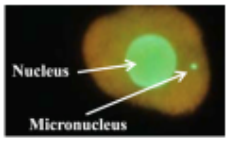
IWGT Workgroup "Statistical Approaches & Data Interpretation"

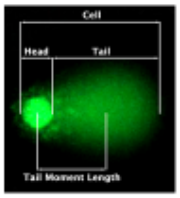
Stephen Dertinger, Litron, chair
Kristine Witt, NIH-NIEHS, co-chair
Carol Beevers, Broughton Group, rapporteur
Andreas Zeller, Roche
Bob Heflich, FDA-NCTR
David Lovell, St. George's University of London
George Douglas & Andrew Williams, Health Canada
Dingzhou (Dean) Li, Pfizer
Daniel Roberts, CRL
Robert Smith, Labcorp
Yoshifumi Uno, MB Medience
Changhui Zhou, InnoStar

While this presentation has benefitted from discussions with this workgroup it should not be considered official IWGT output or consensus findings

Outline

- Use of Historical Control Distributions (HCD) in regulatory genetic toxicology studies
  - Note we'll focus on historical NEGATIVE control distributions for this presentation
- *In vivo* micronucleus example
  - Initial considerations
  - Tools for evaluating sources of variation
  - Calculating useful interval(s)
- *In vivo* Comet example
- Conclusions (spoiler alert: we should use HCD in a flexible, nuanced manner)





Historical Control Distributions

- Genetic toxicology OECD Test Guidelines have harmonized their language regarding HCD and their uses
  - One component of demonstrating laboratory proficiency
  - One component of demonstrating study validity
  - ★ One of three assessments made to judge whether a particular study's response data are "clearly negative" or "clearly positive"
    - A. Pair-wise test that considers concurrent vehicle/solvent control data
    - B. Trend test
    - C. Do the study data fall above or below an upper bound limit value derived from HCD?

"Criterion C"

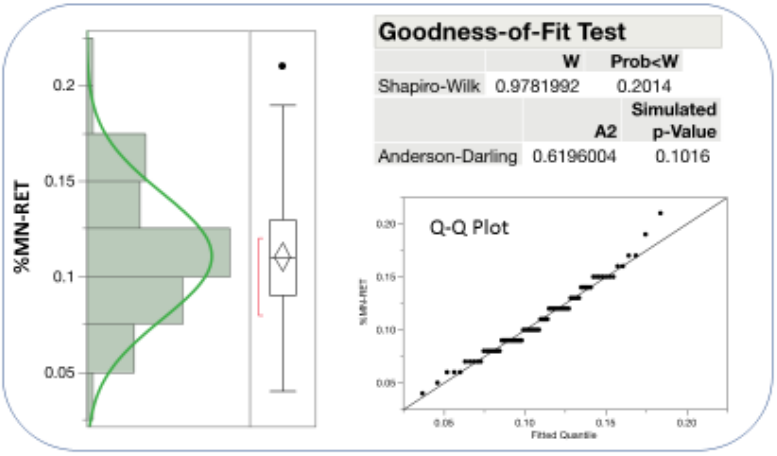
In vivo Micronucleated Reticulocyte (MN-RET) Example

- %MN-RET* as a case study:
  - Evaluate distribution
  - Assess whether assay appears to be "under control"
  - Calculate a useful upper bound value that describes elevated frequencies
- Data = male and female Crl:CD rats; 5 weeks old; pooled males and females given the similarity of values between sexes
  - Total rats = 78; 13 studies over 14 month period of time
  - To simulate an assay that has drifted to an "out of control" status, some of the analyses use the 78 actual rat MN-RET frequencies plus 12 simulated values (six that are 3x higher and six that are 1/3rd lower than actual values)

*data from Dertinger et al., Environ. Mol. Mutagen. 60 (2019) 704-739; available upon request

Evaluate Distribution

- Do the data approximate normal distribution?
  - Some assessment tools (e.g., Nelson Rules) and interval calculations (e.g., Control Limits, Tolerance Intervals) **assume normality**; transform as necessary if you intend to make use of these methods

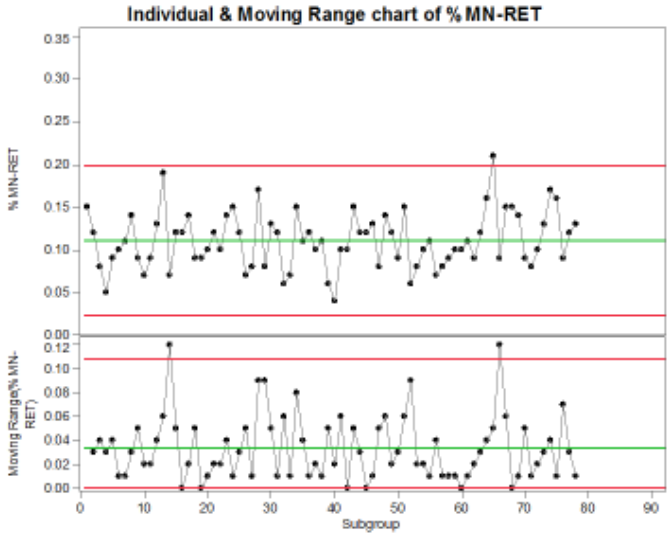


Assess Quality of Historical Control Distribution

- There are a **number of useful approaches** for evaluating the quality of historical control data
- We'll look at each of the following, in turn:
  - Qualitative & semi-quantitative assessments
  - Methods used in the fields of manufacturing, process control
    - Control charts, with or without Nelson Rules
    - Stability Index
  - Variance Component Estimates [e.g., REstricted Maximum Likelihood (REML) analyses, Nested Anova]

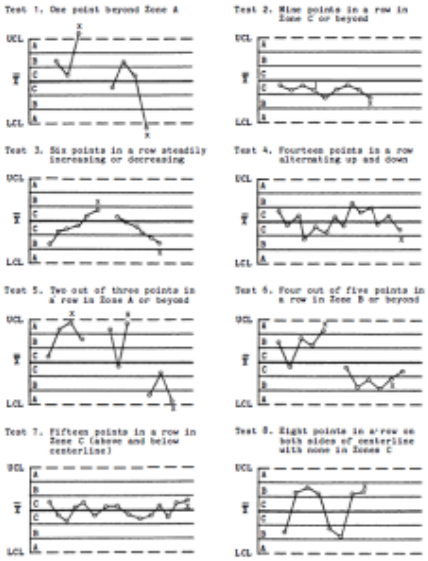
Qualitative & Semi-Qualitative Assessments

- Are the data consistent with published results from proficient labs?
- Is the level of variation across samples within a study and across studies comparable to published results from proficient labs?
- Is there obvious drift with respect to time?
- Control charts can help with these qualitative & semi-quantitative assessments



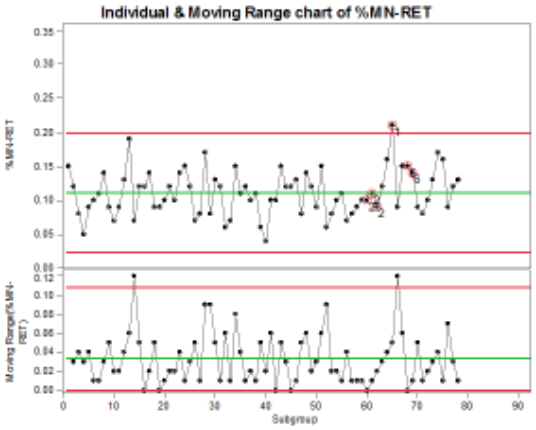
Control Charts with Nelson Rules, from Wiki…

- Nelson rules are a method in process control for determining whether some measured variable is out of control (unpredictable versus consistent)
- First published by Lloyd Nelson in the Journal of Quality Technology, 1984
- The rules are applied to a control chart on which the magnitude of some variable is plotted against time

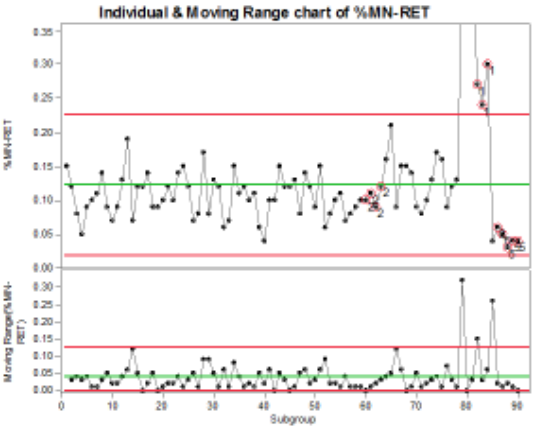


Control Charts with Nelson Rules, cont.

Actual Data (n=78);
Process appears to be "under control,"
relatively few Nelson Rules violations

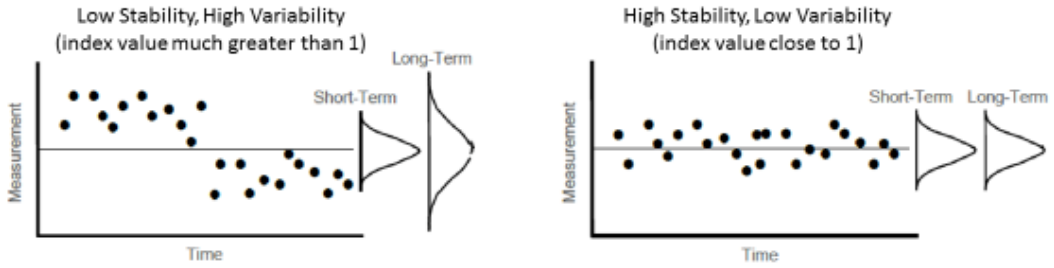


Actual Data + 12 Simulated Samples;
Process has drifted to "out of control" status,
many Nelson Rules violations



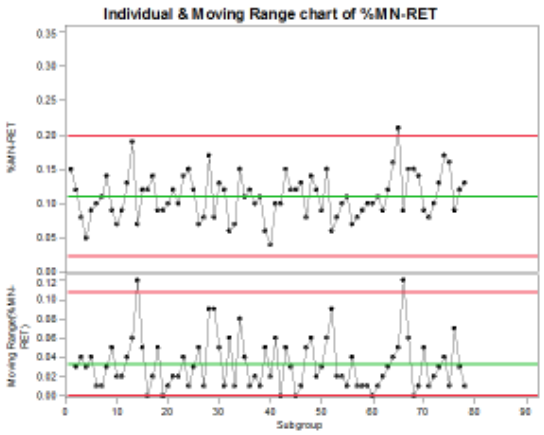
Stability Index

- Manufacturing and Process Control disciplines have developed a variety of tools for evaluating the stability (conversely, the variability) of a process
- One simple metric that might be leveraged for evaluating historical negative control data is the "Stability Index"
- Stability Index = Long-Term Sigma/Short-Term Sigma; close to 1.0 is evidence of stability, i.e., low variability

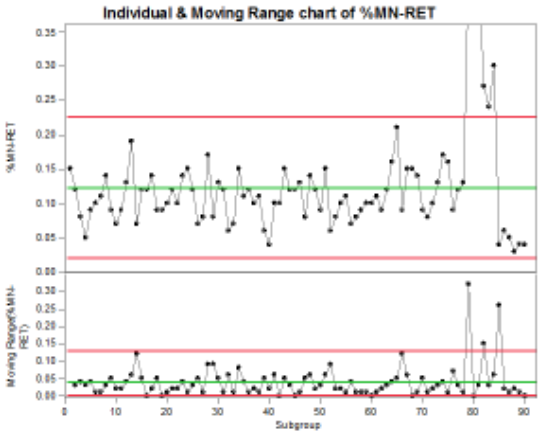


Stability Index, cont.

Actual Data (n=78);
Stability Index = **1.12**



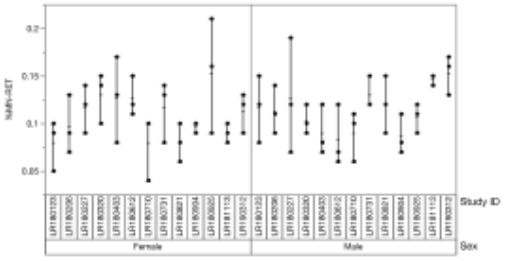
Actual Data + 12 "Out of Control" Samples:
Stability Index = **2.19**



Sources of Variation

- Variance Component Estimates via REsidual Maximum Likelihood (REML), Nested Anova, & Bayesian models may be useful for quantifying sources of variation

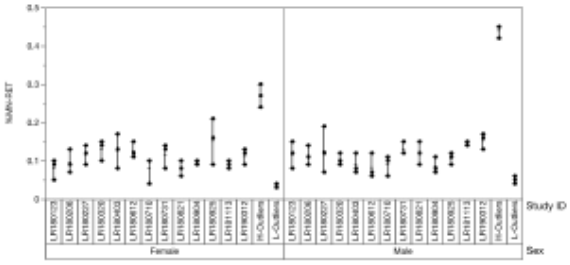
Actual Data (n=78);
Animal to animal variation dominates,
79.7% of total variation observed



Variance Components

Component	Var Component	% of Total	20 40 60 80	Sqrt(Var Comp)
Sex	0.00002568	2.5		0.00507
Study ID[Sex]	0.00018662	17.9		0.01366
Within	0.00083308	79.7		0.02886
Total	0.00104538	100.0		0.03233

Actual Data + 12 "Out of Control" Samples;
Now inter-study variation dominates,
81.9% of total variation



Variance Components

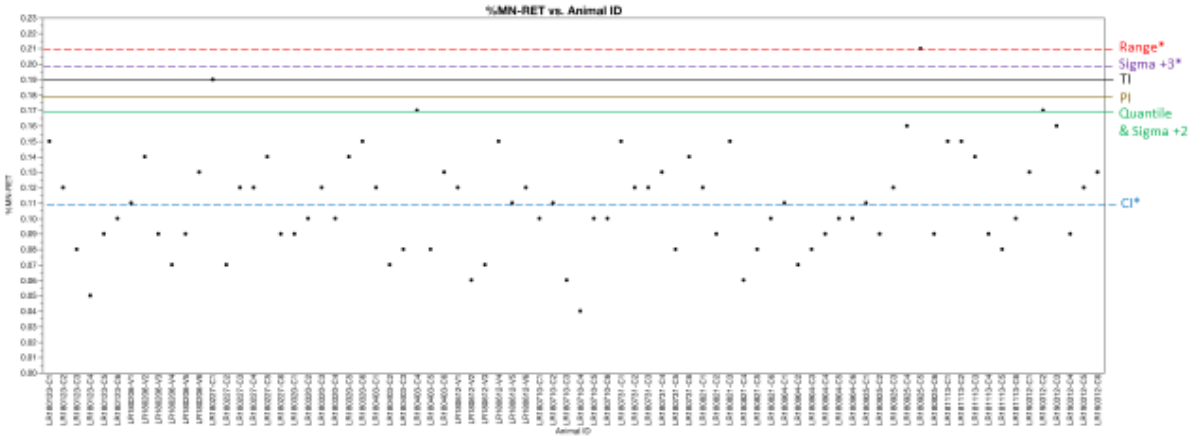
Component	Var Component	% of Total	20 40 60 80	Sqrt(Var Comp)
Sex	0.00022811	4.1		0.01510
Study ID[Sex]	0.00454540	81.9		0.06742
Within	0.00077742	14.0		0.02788
Total	0.00555094	100.0		0.07450

Calculation of Intervals

- As stated earlier, intervals that describe the distribution of the historical control data are useful for a variety of purposes
- BUT… it is premature to calculate and utilize intervals for the purposes described in genetic toxicology OECD TGs **until/unless an assay has been found to be "under control"**
  - Qualitative assessments
  - Control charts (consider supplementing with Nelson Rules, Stability Index)
  - Variance Component Estimates (e.g., REML, Nested Anova, Bayesian)
  - Etc.
- The following slide describes several **less appropriate** and several **more appropriate** means of calculating intervals

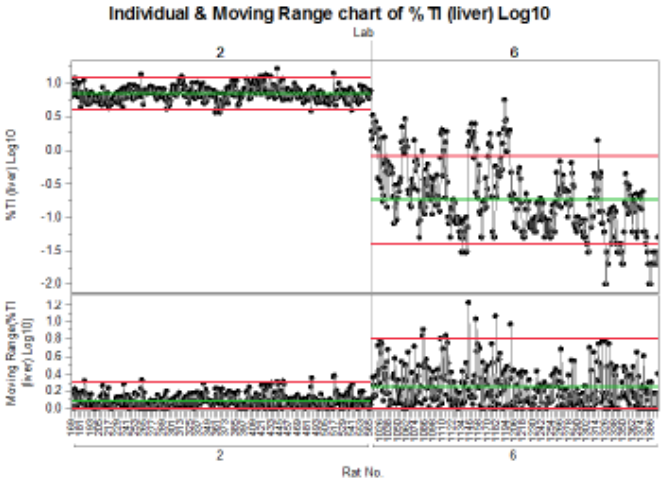
Calculation of intervals; let's focus on the derivation of an upper bound value for "criterion C" assessment; n = 78 (actual) %MN-RET values

- Generally inappropriate intervals for our intended purpose(s)*
  - Range = 0.21%
  - 95% Confidence interval = 0.11%
  - Control Limit, Sigma +3 = 0.20%
- More useful intervals for our intended purpose(s)
  - 95% Quantile = 0.17% (no assumptions about normality)
  - Warning Limit, Sigma +2 = 17%
  - 95% Prediction interval = 18%
  - 95% Tolerance interval, 95% coverage = 19%



Liver Comet Example: Controls Charts with Stability Index

- Two anonymized labs that provided %TI data for IWGT exercises



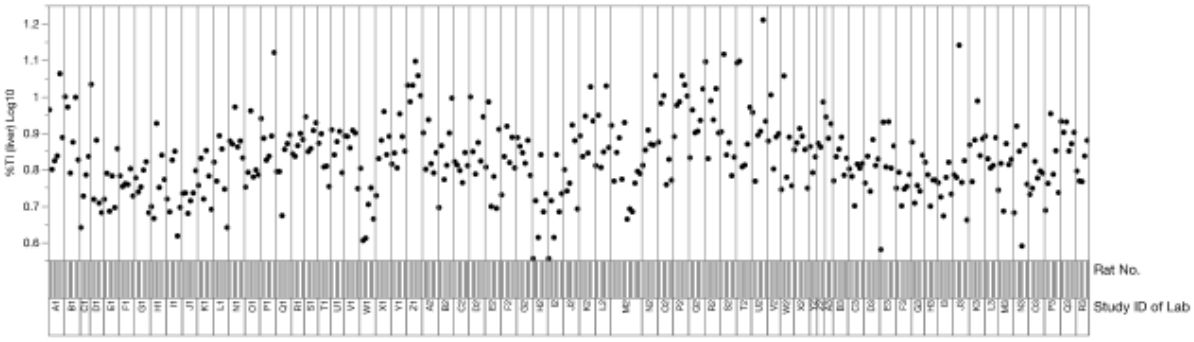
% TI (liver) Log10 Limit Summaries

Points plotted	Lab	LCL	Avg	UCL	Limits Sigma
Individual	2	0.589581	0.833547	1.077514	Moving Range
Individual	6	-1.3978	-0.74225	-0.08671	Moving Range
Moving Range	2	0	0.091762	0.299744	Moving Range
Moving Range	6	0	0.246567	0.805419	Moving Range

Process Sigma Report

Summary	Lab	Value
N	2	398
N Subgroups	2	398
Sample Mean	2	0.833547
Within Sigma	2	0.081322
Overall Sigma	2	0.102495
Stability Index	2	1.260356
N	6	382
N Subgroups	6	382
Sample Mean	6	-0.74225
Within Sigma	6	0.218514
Overall Sigma	6	0.542773
Stability Index	6	2.483926

Liver Comet Example: Variance Component Estimates, Lab 2



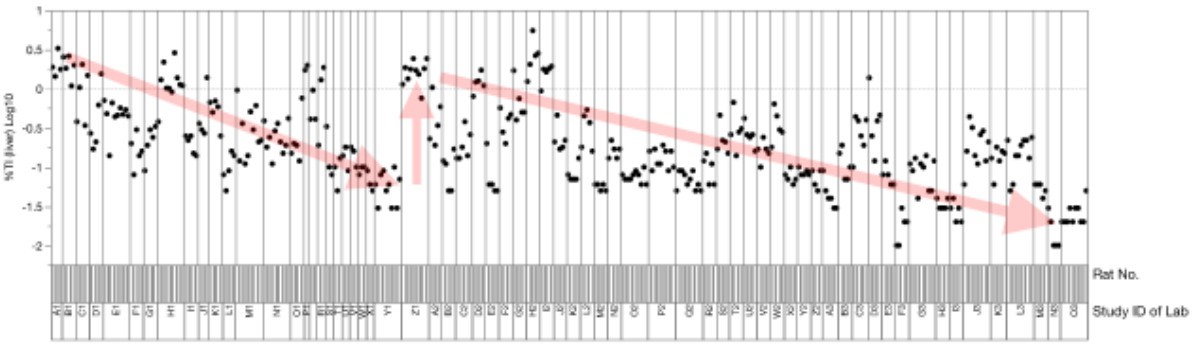
REML Variance Component Estimates

Random Effect	Var Component	Std Error	95% Lower	95% Upper	Pct of Total
Study ID of Lab	0.0038286	0.0008597	0.0025769	0.0062826	36.403
Rat No.[Study ID of Lab]	0.0066886	0.0005209	0.0057745	0.0078395	63.597
Total	0.0105172	0.0009586	0.008865	0.0126812	100.000

-2 LogLikelihood = -756.7062245

Inter-animal variation dominates

Liver Comet Example: Variance Component Estimates, Lab 6



REML Variance Component Estimates

Random Effect	Var Component	Std Error	95% Lower	95% Upper	Pct of Total
Study ID of Lab	0.2592957	0.0466786	0.1875611	0.3819998	85.839
Rat No.[Study ID of Lab]	0.0427754	0.0034091	0.0368077	0.0503302	14.161
Total	0.3020711	0.0467502	0.2279184	0.4195745	100.000

-2 LogLikelihood = 117.58666118

Study ID is predominate source of variation

Circling Back to Criterion C

- HCD is one of three assessments made to judge whether a particular study's response data are "clearly negative" or "clearly positive"
  - A. Pair-wise test that considered concurrent vehicle/solvent control data
  - B. Trend test
  - C. Do the study data fall above or below an upper bound limit value derived from HCD?
- The following are personal opinions, more research is needed:
  - HCD are generally thought of as a proxy for biological variability
  - By extension, HCD are used to evaluate the "biological relevance" of increase(s) identified by criteria A and B
  - Criterion C CAN BE USEFUL when it can be shown variability is primarily explained by biological variability
  - Criterion C is considerably LESS USEFUL when variability is dominated by inter-study nuisance factors—e.g. poorly controlled tissue harvest times, electrophoresis conditions, etc. etc.

Conclusions

- Regulatory genetic toxicology and OECD Test Guidelines make use HCD in several manners
- Important to assess the quality of HCD, and a number of qualitative, semi-quantitative, and quantitative approaches can be employed
- Rather than rigidly applying criterion C, more nuance should be employed
  - For instance, does the HCD describe inter-animal variation? If so, go ahead and apply criterion C; otherwise do not place much weight on it
- There will need to be more detailed reporting of HCD and the type(s) of quality assessments undertaken for all regulatory study stakeholders to gain confidence in their use
- More work is necessary, stay tuned for
  - IWGT Workgroup output
  - Results from a HESI-Genetic Toxicology Technical Committee survey

Thank you for your attention! Questions?

Presentations Day 3

1. ANSES experience - pesticide evaluation in the area of mammalian toxicology

(Adeline Cavelier, Bertrand Desprez, ANSES)

anses

INTERNATIONAL WORKSHOP ON HOW TO REPORT, USE AND INTERPRET HISTORICAL CONTROL DATA IN (ECO)TOXICITY STUDIES

ANSES EXPERIENCE - PESTICIDE EVALUATION IN THE AREA OF MAMMALIAN TOXICOLOGY

ADELINE CAVELIER & BERTRAND DESPREZ
Regulated Products Assessment Department – Toxicological Evaluation Unit of Plant Protection Products

INVESTIGATE, EVALUATE, PROTECT

3-5 May 2022

Regulatory context – 1/4

anses

Commission Regulation (EU) No 283/2013 setting out the data requirements for active substances, in accordance with Regulation (EC) No 1107/2009

Section 5. Toxicology and Metabolism studies

Where available, historical control data shall be provided routinely. The data submitted shall be for endpoints that could represent critical adverse effects, and shall be strain-specific and from the laboratory which carried out the index study. They shall cover a five-year period, centred as closely as possible on the date of the index study.

Detailed information only available for:

- Long-term toxicity and carcinogenicity (Section 5.5)
- Reproductive toxicity (Section 5.6)

2

Regulatory context – 2/4

anses

5.5 Long-term toxicity and carcinogenicity

Where submitted, historical control data shall be from the same species and strain, maintained under similar conditions in the same laboratory and shall be from contemporaneous studies. Additional historical control data from other laboratories may be reported separately as supplementary information.

The information on historical control data provided shall include:

(a) identification of species and strain, name of the supplier, and specific colony identification, if the supplier has more than one geographical location;

(b) name of the laboratory and the dates when the study was performed;

(c) description of the general conditions under which animals were maintained, including the type or brand of diet and, where possible, the amount consumed;

(d) approximate age, in days, and weight of the control animals at the beginning of the study and at the time of killing or death;

(e) description of the control group mortality pattern observed during or at the end of the study, and other pertinent observations (such as diseases, infections);

(f) name of the laboratory and the examining scientists responsible for gathering and interpreting the pathological data from the study;

(g) a statement of the nature of the tumours that may have been combined to produce any of the incidence data.

The historical control data shall be presented on a study by study basis giving absolute values plus percentage and relative or transformed values where these are helpful in the evaluation. If combined or summary data are submitted, these shall contain information on the range of values, the mean, median and, if applicable, standard deviation.

3

Regulatory context – 3/4

anses

5.6 Reproductive toxicity

While the standard reference point for treatment responses shall be concurrent control data, historical control data may be helpful in the interpretation of particular reproductive studies. Where submitted, historical control data shall be from the same species and strain, maintained under similar conditions in the same laboratory and shall be from contemporaneous studies.

The information on historical control data provided shall include:

(a) identification of species and strain, name of the supplier, and specific colony identification, if the supplier has more than one geographical location;

(b) name of the laboratory and the dates when the study was performed;

(c) description of the general conditions under which animals were maintained, including the type or brand of diet and, where possible, the amount consumed;

(d) approximate age, in days, and weight of the control animals at the beginning of the study and at the time of killing or death;

(e) description of the control group mortality pattern observed during or at the end of the study, and other pertinent observations (such as diseases, infections);

(f) name of the laboratory and the examining scientists responsible for gathering and interpreting the pathological data from the study.

The historical control data shall be presented on a study by study basis giving absolute values plus percentage and relative or transformed values where these are helpful in the evaluation. If combined or summary data are submitted, these shall contain information on the range of values, the mean, median and, if applicable, standard deviation.

4

Regulatory context – 4/4

anses

EFSA Administrative Guidance on submission of dossiers and assessment reports for the peer-review of pesticide active substances - EFSA Supporting publication 2019:EN-1612

HCD are necessary to follow changes in the biology of the used test species and to differentiate the way to evaluate test results. HCD represent a summary of the observations made on the untreated or control groups from individual studies and a complete assessment of their relevance should be provided by the applicant in the dossier based on the criteria as set out in Commission Regulation (EU) No 283/2013:

- the incidences of effects for control animals in studies with the same design conducted by the same laboratory; summarised by species, sex, route of administration and vehicle. If study via diet, the diet should be mentioned with reference to the diet characteristics.

- the data for control animals compiled from the concurrent five-year period.

Therefore the following information should be provided:

- the mean, the median, the SD and range of incidences among studies of the effect,

- the number and the dates of studies summarised,

- the use of percentiles could be further considered for HCD of growth or survival (presented as curves),

- Single values (mean, median, SD and range) from those studies that fulfil criteria as set out in Commission Regulation (EU) No 283/2013.

Note that this document only refers to the expected minimum amount of details when reporting HCD. However, if the HCD are intended to be used for the evaluation of the appropriateness of the study's control group, the applicant should refer to the OECD GD 116 for the data set that should be reported and included in the statistical analysis when using HCD.

5

How are used HCD ? – 1/3

anses

Concurrent control data take always precedence on HCD and is the most relevant comparison group to conclude on the treatment-relationship of an effect

However, **HCD may be helpful in the assessment and interpretation of toxicological studies**:

→ In general, HCD are mostly used to assess the following effects:

- Histopathological findings including tumours
- Malformations/variations in fetuses
- Developmental parameters (developmental landmarks ++)
- Gestational parameters (e.g. implantation loss, corpora lutea)
- Functional reproductive parameters (oestrus cycles, sperm parameters)
- + Genotoxicity studies

- Rarely used for e.g. body weight gain, clinical chemistry parameters…

6

How are used HCD ? – 2/3

anses

During the assessment of a toxicity study, HCD are mainly used:

- **For assessing the reliability of a study**
  - → comparison of the incidences in the concurrent control group with incidences in HCD
  - → mainly for genotoxicity studies: HCD are part of acceptability criteria in most of the OECD TG

- **For the interpretation of rare findings**, e.g. tumours, malformations
  - → use of HCD as a way to have an idea on whether lesions are rather rare or occur regularly

- **For the interpretation of borderline findings**, i.e. marginally increased incidence and/or severity of non-neoplastic lesions

- **For the interpretation of a genotoxicity study** - HCD are part of evaluation criteria in most of the OECD TG

HCD

- Represent data obtained on a larger group size (number of animals/cells) than the concurrent control group
- Give a better overview of the biological variation of the finding of interest

7

How are used HCD ? – 3/3

anses

HCD should **ALWAYS** be used in a **Weight of Evidence approach** to decide on the treatment-relationship of an effect:

As a first step: Comparison of the incidence/severity in the treated groups versus incidence/severity in the **concurrent control group**

Then: if needed, comparison to incidence/severity in the **HCD**

- Magnitude of the effect
- Dose-response relationship (proportionnality with the dose not expected)
- Statistical significance – pairwise comparison and trend analysis
- Biological plausibility, e.g. continuum from pre-neoplastic lesions to malignant tumours

HCD should not be used to dismiss an adverse effect if:

- Comparison to the concurrent control group confirms the effect
- And/or dose-relationship is obvious
- And/or statistical significance is noted

8

Which information are needed on HCD?

anses

Currently, the **MINIMAL** information that should be provided to consider HCD as relevant are:

- Same species and strain, same sex
- Same laboratory *and breeder*
- Same route of administration (diet ≠ gavage)
- Same type of study and study duration, *same age of the animals*
- Number of studies

- Studies included in the HCD should be contemporary to the assessed study:

+/- 5 years centred as closely as possible around the date of the index study

(experimental/in-life dates of the study) (e.g. HCD from 2008-2012 if study conducted in 2010)

- Reporting of results: at least mean, min-max range, standard deviation

Litter and fetal incidences for developmental toxicity studies

In absence of these data and/or if criteria not fulfilled

→ **HCD are not considered relevant or considered of low relevance**

9

What is often missing… but considered needed? – 1/2

anses

In the submitted dossiers, some information are missing

However, to our point of view, they should be provided to improve relevance of HCD:

- Breeder
- Age and weight of the animals
- Information on environmental factors, caging protocol, stress conditions
- Diet characteristics in the case of dietary administration
- Vehicle (for genotoxicity studies, studies with different vehicles are often combined)
- Diagnostic criteria (histopathology)
- Staining method (developmental toxicity studies)
- Standardised terminology
- OECD and GLP status

What is often missing… but considered needed? – 2/2

anses

Presentation/reporting of HCD:

- Detailed statistical evaluation + raw data
- Information on HCD distribution: 'mean' and 'min-max range' are useful information but are not considered sufficient since distribution is an important point to consider (identification of outliers)
- Ideally, graphical representation of the distribution should be provided
- If only '% incidence' available: not considered sufficient, incidences in ratios (population with effect / size of the population) are also needed
- Calculation of confidence limits should be harmonized

► Example, issues and solutions in the next slides

11

Improving statistics from HCD – 1/5

anses

What is shown in the next slides applies to any HCD reporting proportions or %

The chosen example uses micronucleus data for the purpose of the presentation, but considerations made would apply to any type of proportion, for instance % of carcinomes

Several issues are raised… but solutions are presented too!

- Issue #1: poorly summarized HCD
- Issue #2: using test guideline-recommended 95% confidence limits
- Issue #3: variability considerations on HCD by using 95% confidence intervals on percentages or proportions
- Issue #4: getting all raw data too

12

Improving statistics from HCD – 2/5

anses

When dealing with percentages (%) or proportions, most of the time HCD are presented as:

- minimum-maximum range, or
- minimum-maximum range + mean, or
- minimum-maximum range + mean + SD
- Possibly, number of studies are also specified

This provides a *certain idea* of HCD data BUT it is however insufficient to have a *good idea* of HCD and HCD distribution and 95% confidence limits (95% confidence limits)

Confidence limits are even wrong most of the time : « Mean ± 3.SD and round to zero if negative (!) »

Example, Micronucleus test

Historical Range for Vehicle Control Cultures

	4 hour exposure without S9	4 hour exposure with S9	24 hour exposure without S9
	% binucleate cells with micronuclei	% binucleate cells with micronuclei	% binucleate cells with micronuclei
Minimum	0.05	0.05	0.15
Maximum	1.20	1.30	0.90
Mean	0.56	0.51	0.47
Standard Deviation	0.29	0.29	0.19
95% Control Limits	0 – 1.43	0 – 1.38	0 – 1.04
Number of Experiments	50	50	50

Here (4h, –S9)
Mean = 0.56%
SD = 0.29%
Upper bound = 0.56 - 3×0.29 = 1.43
Lower bound = 0.56 + 3×0.29 = -0.31
rounded to 0 (!)

Issue #1
Rounding to zero results in artificial shortening of the 95%CL
The presented interval does not even have a 95% coverage according to the chosen methodology
Truncated CL implies information loss

Issue #2
The OECD test guideline 487 recommends the use of Poisson distribution – which is a way to deal easily with proportions
95%CL(Poisson) = mean ± 2× √mean (because in Poisson distribution SD = √mean)

Here, if 2000 cells were used, the mean at 0.56% would be corresponding to 11 cells out of 2000 cells

13

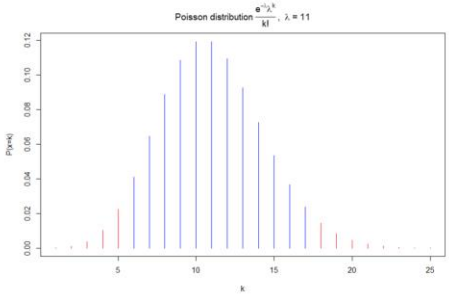
Improving statistics from HCD – 3/5

anses

Example, Micronucleus test

Historical Range for Vehicle Control Cultures

	4 hour exposure without S9	4 hour exposure with S9	24 hour exposure without S9
	% binucleate cells with micronuclei	% binucleate cells with micronuclei	% binucleate cells with micronuclei
Minimum	0.05	0.05	0.15
Maximum	1.20	1.30	0.90
Mean	0.56	0.51	0.47
Standard Deviation	0.29	0.29	0.19
95% Control Limits	0 – 1.43	0 – 1.38	0 – 1.04
Number of Experiments	50	50	50



Issue #2 (continued)
Here, if 2000 cells were used, the mean at 0.56% would be corresponding to 11 cells out of 2000 cells

95% confidence limits =
11 ± 2×√11
i.e., 95CL = [4.4 – 17.6] cells out of 2000
i.e., 95CL = [0.21 – 0.88] %
Here the coverage is indeed 95% contrary to what is shown in the Table

Providing % is not sufficient, what should be provided is the real number of cells with the effect out of the number of cells tested.
Presenting in % can still be presented for group comparison purposes

Issue #3 Variability considerations in HCD by using 95% CI or 95% CL
Poisson is an approximation
What should be used is proportion confidence interval i.e., binomial confidence interval
Here, x=11 « successes » (cells with micronuclei) out of n=2000 « attempts » (total cells tested)
The 95% confidence interval is [0.29—0.99]%
Again the coverage is 95%, contrary to the Table, good consistency with Poisson interval

```
binom.agresti.coull(x=11, n=2000)
        method  x    n   mean     lower      upper
agresti-coull 11 2000 0.0055 0.0029435 0.00995246
```

14

Improving statistics from HCD – 3/5

anses

Example, Micronucleus test

Historical Range for Vehicle Control Cultures

	4 hour exposure without S9	4 hour exposure with S9	24 hour exposure without S9
	% binucleate cells with micronuclei	% binucleate cells with micronuclei	% binucleate cells with micronuclei
Minimum	0.05	0.05	0.15
Maximum	1.20	1.30	0.90
Mean	0.56	0.51	0.47
Standard Deviation	0.29	0.29	0.19
95% Control Limits	0 – 1.43	0 – 1.38	0 – 1.04
Number of Experiments	50	50	50

Key message: Why presenting percentages only is a problem ?

Example of 1%
1% = ratio of 10 / 1000 → 95% CI [0.52% — 1.86%]
1% = ratio of 100 / 10 000 → 95% CI [0.82% — 1.22%]

Issue #2 (continued)
Here, if 2000 cells were used, the mean at 0.56% would be corresponding to 11 cells out of 2000 cells

95% confidence limits =
$11 \pm 2\times\sqrt{11}$
i.e., 95CL = [4.4 – 17.6] cells out of 2000
i.e., 95CL = [0.21 – 0.88] %
Here the coverage is indeed 95% contrary to what is shown in the Table

Providing % is not sufficient, what should be provided is the real number of cells with the effect out of the number of cells tested.
Presenting in % can still be presented for group comparison purposes

Issue #3 Variability considerations in HCD by using 95% CI or 95% CL
Poisson is an approximation
What should be used is proportion confidence interval i.e., binomial confidence interval
Here, x=11 « successes » (cells with micronuclei) out of n=2000 « attempts » (total cells tested)
The 95% confidence interval is [0.29—0.99]%
Again the coverage is 95%, contrary to the Table, good consistency with Poisson interval

```
binom.agresti.coull(x=11, n=2000)
        method  x    n   mean     lower      upper
agresti-coull 11 2000 0.0055 0.0029435 0.00995246
```

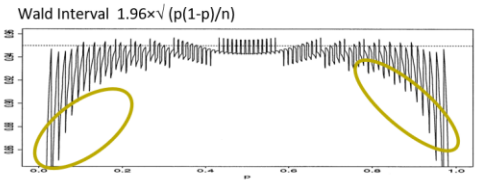
14

Improving statistics from HCD – 4/5

anses

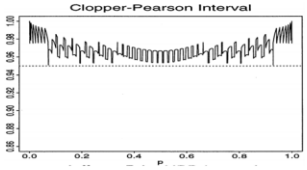
Issue #3 (continued)
What should be used is confidence interval for proportions i.e., binomial confidence interval
There are significant shortcomings in working directly on % instead of x out of n (x/n ratio)
The well-known formula $p \pm 1.96\times\sqrt{(p(1-p)/n)}$ has been reported to not perform correctly (coverage <<95%) when proportions are close to 0 (which is the case in negative HCD) or close to 1
Brown, Cai and DasGupta, Statistical Science 2001, Vol.16, No. 2 101—183

Wald Interval $1.96\times\sqrt{(p(1-p)/n)}$



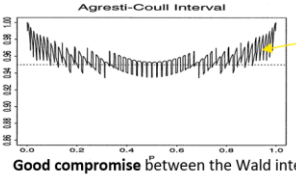
Coverage << 95% when for proportions close to 0 or 1
Important consideration for negative HCD 95%CL!

Clopper-Pearson Interval



Too conservative >>95% consistently

Agresti-Coull Interval



Good compromise between the Wald interval that underperforms and the Clopper-Peason that is too conservative

This interval was shown 2 slides ago
Good consistency with 95%CL that can be derived from Poisson distribution
Consequence : HCD 95%CL should at least use and present correctly the one calculated from Poisson law. Binomial confidence intervals are a must

15

Improving statistics from HCD – 4/5

anses

Issue #3 (continued)

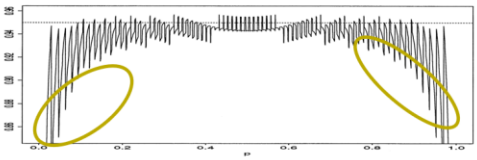
What should be used is confidence interval for proportions i.e., binomial confidence interval

There are significant shortcomings in working directly on % instead of x out of n (x/n ratio)

The well-known formula $p \pm 1.96 \times \sqrt{(p(1-p)/n)}$ has been reported to not perform correctly (coverage <<95%) when proportions are close to 0 (which is the case in negative HCD) or close to 1

Brown, Cai and DasGupta, Statistical Science 2001, Vol.16, No. 2 101—183

Wald Interval $1.96 \times \sqrt{(p(1-p)/n)}$

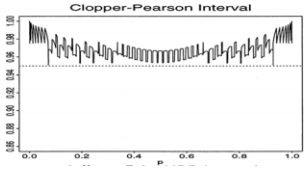


Coverage << 95% when for proportions close to 0 or 1

Important consideration for negative HCD 95%CL!

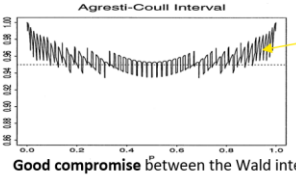
Key message : deciding on how to establish 95% CI is not trival for HCD and in the end how HCD will be relied upon

Clopper-Pearson Interval



Too conservative >>95% consistently

Agresti-Coull Interval



Good compromise between the Wald interval that underperforms and the Clopper-Peason that is too conservative

This interval was shown 2 slides ago

Good consistency with 95%CL that can be derived from Poisson distribution

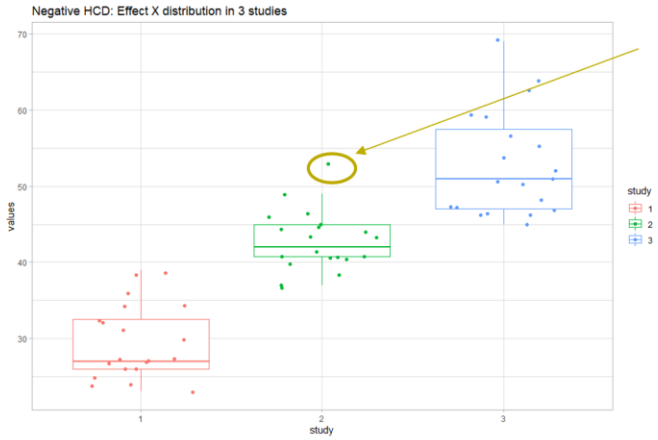
Consequence : HCD 95%CL should at least use and present correctly the one calculated from Poisson law. Binomial confidence intervals are a must

15

Improving statistics from HCD – 5/5

anses

Issue #4 raw data for HCD
Min.-max. range, mean and 95%CL (even correctly calculated) are not yet sufficient
Also raw data are needed in addition to summarized data
One should be able to study distributions from available studies and identify outliers



This is an outlier in Study 2
Depending on the period chosen for HCD to be used, this information may be important

This allows to have an idea on how much presented data are reliable

16

Conclusion

anses

Although concurrent control group is the most appropriate comparator for assessing treatment-relationship of a finding, **HCD may in certain cases be helpful**

For genotoxicity studies, **HCD are mandatory** – part of the acceptability and interpretation criteria

Most of the cases:

- Information needed for assessing relevance of available HCD are missing
- Reporting of HCD is incomplete

Need for harmonisation:

- use of HCD
- minimal information to be provided to assess the reliability of HCD
- reporting of the HCD/improving statistics: in terms of statistical analysis and of reporting

► Mean, median and min-max (% and corresponding ratios), SD, number of studies, choice of confidence interval types, graphical distributions, raw data

17

2. Experience from the European pesticide, biocide and C&L evaluation – Competent Authority perspective

(Susanne Rudzok, BfR)

GERMAN FEDERAL INSTITUTE FOR RISK ASSESSMENT

20 YEARS BfR German Federal Institute for Risk Assessment

Experience from the European pesticides, biocides and C&L evaluation Competent Authority perspective

Dr. Susanne Rudzok

Pesticides, Biocides and C&L evaluation

Department of Pesticide Safety
Unit: Toxicology of Active Substance and their Metabolites

Human Health – Hazard Assessement under different regulations – each with different data requirements

- PPP — REGULATION (EC) No 1107/2009 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 21 October 2009 concerning the placing of plant protection products on the market and repealing Council Directives 79/117/EEC and 91/414/EEC
- BPR — REGULATION (EU) No 528/2012 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 22 May 2012 concerning the making available on the market and use of biocidal products (Text with EEA relevance)
- C&L — 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, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006 (Text with EEA relevance)

5th May 2022, Workshop HCD page 2 BfR

Survey

Question	Answer	Eintrag in Survey BPR	Eintrag in Survey PPP	Eintrag in Survey CLH
General considerations on historical control data (HCD)				
For which purpose do you consider information from HCD				
	To create reference ranges to flag outliers during the evaluation			
	To create reference ranges for normal background variability	x	x	x
	For comparison to controls of the study under consideration to assess study reliability	x	x	x
	For comparison to controls of the study under evaluation to identify genetic drift in species / strains	x	x	x
	For comparison of animals of a single strain sourced from different suppliers			
	To replace concurrent control data by HCD			
	To integrate HCD in the statistical analysis of the study results			
	Other, please specify	typically, referenceranges from third party sources are used rather than created)		
	Never consider HCD			
From which sources do you consider compilation of HCD?				
	Internal company database			
	Proprietary information from contract organization / laboratory conducting the study	x	x	x
	Proprietary information from contract organization at location other than study conduct			
	Registry of Industrial Toxicology Animal-data (RITA)			
	National Toxicology Programme (NTP)			
	Open access information from CROs (e.g. Charles Rivers compilations)			
	Published literature (peer reviewed journals)	x	x	x
	Other, please specify	from RITA /NTP / CROs(Charles River) are used only assupplementary source ofinform		
When considering compiling / providing / asking for HCD, which regulatory documents or guidance do you consult?				
	Body weight and body weight gains			
	Food consumption			
	Water consumption			
	In-life observations (e.g. clinical signs)			
	Ocular examination	x	x	x
	Haematology, clinical chemistry, urinalysis	x	x	x
	Mortality			
	Organ weights	x	x	x
	Gross pathology findings			
	Histopathology findings	x	x	x
	Foetal pathology findings (e.g. malformations, variations)	x	x	x
	Developmental parameters (e.g. survival, body weight, developmental landmarks)	x	x	x
	Gestational parameters (e.g. pre-/ post-implantation loss, corpora lutea, implantation sites)	x	x	x
	Functional reproductive parameters (oestrus cycle / semen analysis)	x	x	x
	Functional and behavioural parameters (e.g. Functional Observational Battery (FOB), motor activity)	x	x	x
	Other, please specify	use of HCD for general parameters such asfood / water consumption / mortality are considered only inexceptional cases; HCD for Haematology/ Clinical Chemistry/Urinalysis used with great caution		

5th May 2022, Workshop HCD page 3 BfR

PPP data requirements on HCD

Commission Regulation (EU) No 283/2013 of 1 March 2013 setting out the data requirements for active substances, in accordance with Regulation (EC) No 1107/2009

Where available, historical control data shall be provided routinely. The data submitted shall be for endpoints that could represent critical adverse effects, and shall be strain-specific and from the laboratory which carried out the index study. They shall cover a five-year period, centred as closely as possible on the date of the index study.

Where submitted, historical control data shall be from the same species and strain, maintained under similar conditions in the same laboratory and shall be from contemporaneous studies. Additional historical control data from other laboratories may be reported separately as supplementary information.

The historical control data shall be presented on a study by study basis giving absolute values plus percentage and relative or transformed values where these are helpful in the evaluation. If combined or summary data are submitted, these shall contain information on the range of values, the mean, median and, if applicable, standard deviation.

5th May 2022, Workshop HCD page 4 BfR

PPP data requirements on HCD (EU No 283/2013)

In Theory

The information on historical control data provided shall include:

(a) identification of species and strain, name of the supplier, and specific colony identification, if the supplier has more than one geographical location;

(b) name of the laboratory and the dates when the study was performed;

(c) description of the general conditions under which animals were maintained, including the type or brand of diet and, where possible, the amount consumed;

(d) approximate age, in days, and weight of the control animals at the beginning of the study and at the time of killing or death;

(e) description of the control group mortality pattern observed during or at the end of the study, and other pertinent observations (such as diseases, infections);

(f) name of the laboratory and the examining scientists responsible for gathering and interpreting the pathological data from the study;

(g) a statement of the nature of the tumours that may have been combined to produce any of the incidence data.

5th May 2022, Workshop HCD page 5 BfR

BPR data requirements on HCD

Guidance on the Biocidal Products Regulation

Volume III: Human health
Part A: Information requirements
Version 2, March 2022

General considerations for animal data reporting

Where submitted, historical control data should be from the same species and strain, maintained under similar conditions in the same laboratory and should be from contemporaneous studies (within a period of five years, centred as closely as possible on the date of the study). Additional historical control data not fulfilling these conditions, or from other laboratories may be reported separately as supplementary information.

The historical control data should be presented on a study-by-study basis giving absolute values plus percentage and relative or transformed values where these are helpful in the evaluation. If combined or summary data are submitted, these should contain information on the number of studies included and whether the current study is included, the range of values, the mean, median and, if applicable, standard deviation.

If the appropriateness of the control group of the study is in question, please refer to the considerations in OECD GD 116 (section 4.22) on the relevant details in analysing the historical control data.

5th May 2022, Workshop HCD page 6 BfR

BPR data requirements on HCD

The information on historical control data provided should include[6]:

(a) identification of species and strain, name of the supplier, and specific identification if the supplier has more than one geographical location;

(b) name of the laboratory and the dates when the study was performed;

(c) description of the general conditions under which animals were maintained, including the type or brand of diet and, where possible, the amount consumed;

(d) approximate age, in days, and weight of the control animals at the beginning of the study and at the time of sacrifice or death;

(e) description of the control group mortality pattern observed during or at the end of the study, and other pertinent observations (such as diseases, infections);

(f) name of the laboratory and the examining scientists responsible for gathering and interpreting the pathological data from the study;

(g) for carcinogenicity studies: a statement of the nature of the tumours that may have been combined to produce any of the incidence data.

5th May 2022, Workshop HCD page 7 BfR

BPR data requirements on HCD

The information on historical control data provided should include[6]:

(a) identification of species and strain, name of the supplier, and specific identification if the supplier has more than one geographical location;

(b) name of the laboratory and the dates when the study was performed;

(c) description of the general conditions under which animals were maintained, including the type or brand of diet and, where possible, the amount consumed;

(d) approximate age, in days, and weight of the control animals at the beginning of the study and at the time of sacrifice or death;

(e) description of the control group mortality pattern observed during or at the end of the study, and other pertinent observations (such as diseases, infections);

(f) name of the laboratory and the examining scientists responsible for gathering and interpreting the pathological data from the study;

(g) for carcinogenicity studies: a statement of the nature of the tumours that may have been combined to produce any of the incidence data.

[6] This information will enable the assessment of the relevance of the historical data and the effects observed in the study provided.

If some of the elements listed above are missing, this must be considered in assessing the relevance of the historical control data.

5th May 2022, Workshop HCD page 8 BfR

Reality – Example 1 - Clomazone

Developmental toxicity study rats

Incidences of major external, visceral and skeletal malformations

Parameters	Doses [mg/kg bw/d]				Historical control ranges (%)
	Control	250	500	750	
Major external malformations					
Number of foetus examined	265	254	260	204	-
Forelimbs flexed at wrist	0.0	0.0	0.0	0.5[c]	0.0-2.1
Anasarca	0.0	0.0	0.4[a]	0.5[a]	na.
Foetuses with major malformations (%)	0	0	1 (0.4)	2 (1.0)	0.0-4.0
Dams with major malformed fetuses (%)	0	0	1 (4)	2 (10)	0.0-11.0

[c] also indicated as arthrogryposis in study report, dam different from the dams with major malformed fetuses under skeletal malformations

Incidence for "forelimbs flexed at wrist" (arthrogryposis) seems to be within HCD

However, only submission of „% high range" data

HCD Clomazone

11 studies

Of the *external* observations a single study defined the high range of 2.1% of forelimbs flexed at wrist.

- further inspection of this study would be needed

Incidence is out of HCD

submission of only „% high range" should not be accepted

Table 6.6.2/06-1: Historical Data of Prenatal Developmental Toxicity Study in Wistar Rats – External Observation (%)

Study No.	2523/1998	2642/1999	2723/1999	2438/1997	2676/1999	3217/2001	3268/2001	3369/2002	3557/02	3634/03	3479/04	% low range	% high range
No. of litters examined	147	24	18	21	21	27	24	21	24	20	23		
No. of fetuses examined	728	228	141	229	248	323	269	241	279	221	267		
Parameters													
NORMAL VARIANT													
Dead fetus	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.40	0.00	0.00	0.00	0.00	0.40
Haemorrhagic spot	0.4	2.6	2.1	1.7	9.3	2.2	1.90	3.30	1.08	2.26	0.75	0.40	9.30
MINOR ANOMALIES													
Haemorrhagic patch	0.3	0.9	1.4	1.3	7.7	1.2	0.40	0.80	0.00	1.81	1.12	0.00	7.70
Subcutaneous edema on ventral aspect of neck	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.36	0.00	0.00	0.00	0.36
Fetus small	0.1	0.00	0.00	0.00	0.40	0.00	0.00	0.00	0.00	0.45	0.37	0.00	0.45
Tongue - protruding	0.1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10
Hydrocephaly	0.00	0.00	0.7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.70
Fore limb flexed at wrist	0.00	0.00	9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.33
MAJOR MALFORMATIONS													
Malformed fetus	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.45	0.00	0.00	0.45
Cleft palate	0.1	0.00	2.1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.10
F limb-flexed at wrist	0.00	0.00	2.1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.10
Fetuses with major external malformations (%)	1 (0.1)	0 (0.0)	2 (4.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.00)	0 (0.00)	0 (0.00)	1 (0.45)	0 (0.00)	0.00	2.00
Dams with major malformed fetuses (%)	1 (0.7)	0 (0.0)	2 (11.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.00)	0 (0.00)	0 (0.00)	1 (5)	0 (0.00)	0.00	11.00

Reality – Example 2a - Diuron

Rat, 2-yr carcinogenicity study , Uterus				
Dose (mg/kg bw/d)	**0**	**1.7**	**17**	**203**
Number of animals	48	50	50	50
Polyp (b)	7 15 %	7 14 %	6 12 %	3 6 %
Fibromyoma (b)	0	1 2 %	0	0
Leiomyosarcoma (m)	0	1 2 %	0	0
Endometrium sarcoma (m)	0	0	0	2 4 %
Adenocarcinoma (m)	5 10 %	5 10 %	5 10 %	**10 #** 20 %
Squamous epithelial carcinoma (m)	0	0	1 2 %	1 2 %

b=benign; m= malignant; # for $p \leq 0.05$ (Cochrane Armitage linear trend test, one-sided)

Reality – Example 2a - Diuron

Rat, 2-yr carcinogenicity study , Uterus				
Dose (mg/kg bw/d)	**0**	**1.7**	**17**	**203**
Number of animals	48	50	50	50
Polyp (b)	7 15 %	7 14 %	6 12 %	3 6 %
Fibromyoma (b)	0	1 2 %	0	0
Leiomyosarcoma (m)	0	1 2 %	0	0
Endometrium sarcoma (m)	0	0	0	2 4 %
Adenocarcinoma (m)	5 10 %	5 10 %	5 10 %	**10 #** 20 %
Squamous epithelial carcinoma (m)	0	0	1 2 %	1 2 %

b=benign; m= malignant; # for p ≤ 0.05 (Cochrane Armitage linear trend test, one-sided)

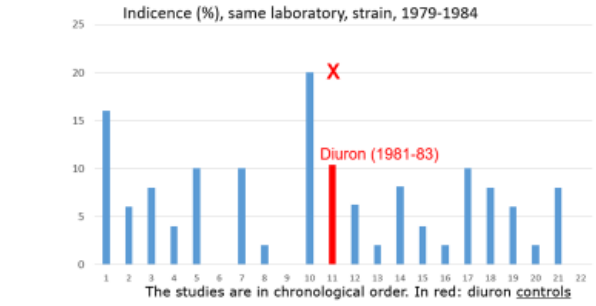
Historical Control Data Compilation provided by the Applicant:

Uterus adenocarcinoma Incidence	
8/50	
3/50	
4/50	
2/50	
5/50	
0/49	
5/50	
1/49	
0/49	
10/50	
5/48	Diuron study
3/48	
1/50	
4/49	
2/50	
1/50	
5/50	
4/50	
3/50	
1/50	
4/50	

- ✓ 21 studies (including Diuron-study)
- ✓ 5 yr, centered closely to index study
- ✓ laboratory name, strain

Example 2a – Diuron - **Evaluation for C&L**

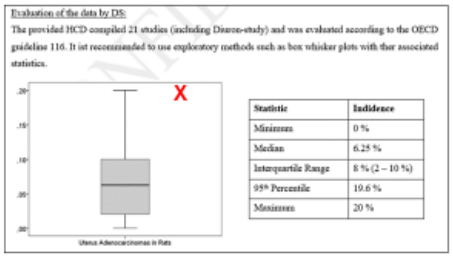
Comparison with historical control data



- Incidence of 20% at the top dose is at the upper end of HCD, far above mean of 8%
- HCD 1975-1980: 0-16%, mean 8%; HCD 1975-1994: 0-20%, mean 6%

Rat, 2-yr carcinogenicity study , Uterus				
Dose (mg/kg bw/d)	**0**	**1.7**	**17**	**203**
Number of animals	48	50	50	50
Adenocarcinoma (m)	5 10 %	5 10 %	5 10 %	**10 #** 20 %

b=benign; m= malignant; **# for p ≤ 0.05 (Cochrane Armitage linear trend test, one-sided**



Example 2b – Diuron - **Evaluation for C&L**

Statistics
- Not statistically significant in group-wise comparison
- Borderline significance for trend (in one-sided by not two-sided test)

Dose-Response
- Only top dose affected - no clear dose response documented
- Descrease in benign tumors – trend towards increased malignancy (DS interpretation) or no effect on overall incidence of neoplasia (IND interpretation)

Tumor latency
- Not specified

HCD
- Borderline

➡ **No strong evidence for classification**

Mechanistic considerations
- endocrine MoA might be suspected but speculative

Reality – Example 2b- Diuron

Mice, 2-yr carcinogenicity study , Mammary gland				
Dose (mg/kg bw/d)	**0**	**7.5**	**77.5**	**867**
Number of animals	39	32	44	39
Adenocarcinoma (m)	2	1	1	6*
	5 %	3 %	2 %	15 %
Anaplastic carcinoma (m)	0	1	0	0
		3 %		

m= malignant; * statistically significant trend $p \leq 0.05$ (Peto trend test); statistically evaluation by study director

Historical Control Data Compilation provided by the Applicant:

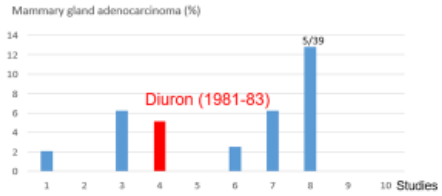
Mammary gland carcinoma, Incidence	
1/48	study duration 21 months
0/47	study duration 21 months
3/48	study duration 21 months
2/39	study duration 24 months (Diuron)
0/49	study duration 21 months
1/40	study duration 20 months
3/48	study duration 21 months
5/39	study duration 21 months
0/45	study duration 21 months
0/50	study duration 21 months

- ✓ 10 studies (including Diuron-study)
- ✓ 5 yr centered closely to index study
- ✓ laboratory name, strain

Example 2b – Diuron - **Evaluation for C&L**

Comparison with historical control data

Indicence (%), same laboratory, strain, 1981-1984



The studies are in chronological order. In red: diuron controls

- HCD are limited: low number of studies, 20-21-month studies instead of 24 month in the diuron study
- Incidence of 15% (6/39) at the diuron top dose is slightly above HCD and is far above HCD mean (3.2%)

Mice, 2-yr carcinogenicity study , Mammary gland				
Dose (mg/kg bw/d)	**0**	**7.5**	**77.5**	**867**
Number of animals	39	32	44	39
Adenocarcinoma (m)	2	1	1	6*
	5 %	3 %	2 %	15 %

m= malignant; * statistically significant trend $p \leq 0.05$ (Peto trend test); statistically evaluation by study director

Bomhard (1993) showed that there is no high biological variability of spontaneously occurring tumors in the mammary gland of NMRI mice:

In **16 studies** a total of 717 control group female mice were examined displaying:

- 1 case of carcinosarcoma,
- 4 cases of adenoacanthoma (highest incidence in a single study 1/45) and
- 2 cases of adenomas (highest incidence in a single study 1/47) and
- 22 cases of carcinomas with a **highest incidence in a single study of 6.3 % (3/48).**

Example 2b – Diuron - **Evaluation by RAC**

Statistics
- Statistically significant (trend test only)

Dose-Response
- Top dose only

Pre-neoplastic findings
- The mammary gland of some of the animals were activated and produced secretions, which could lead to cystic enlargements of mammary structures

Tumour Latency and survival
- No mammary gland tumours at interim kill
- The tumours occurred at roughly the same time as control group.

➡ **Relevant for classification**

supported by incidences above limited relevant HCD and clearly above less representative HCD

Summary

- Historical control data is used in the assessment of active substances (pesticides & biocides)
- 1st challenge: representative HCD
  - 5 year period, centered as closely as possible to the date of the index study
  - Sufficient number of studies

Summary

- Historical control data is used in the assessment of active substances (pesticides & biocides)
- 1st challenge: representative HCD
  - 5 year period, centered as closely as possible to the date of the index study
  - Sufficient number of studies
- 2nd challenge: data gaps
  - Laboratory name, strain, weight, age
  - Information e.g. on infections that could be the reason for a control study with higher background
  - Only summary data, not generally including percentiles, mean or median
  - Frequently not appropriate to compare incidences of malformations, resorptions etc.

Summary

- Historical control data is used in the assessment of active substances (pesticides & biocides)
- 1st challenge: representative HCD
  - 5 year period, centered as closely as possible to the date of the index study
  - Sufficient number of studies
- 2nd challenge: data gaps
  - Laboratory name, strain, weight, age
  - Information e.g. on infections that could be the reason for a control study with higher background
  - Only summary data, not generally including percentiles, mean or median
  - Frequently not appropriate to compare incidences of malformations, resorptions etc.
- 3rd challenge: Interpretation of the HCD
  - Applicants often use HCD in the argumentation to negate experimental results using HCD range
  - Competent authorities can use range to evaluate the concurrent control

Summary

- Historical control data is used in the assessment of active substances (pesticides & biocides)
- 1st challenge: representative HCD
  - 5 year period, centered as closely as possible to the date of the index study
  - Sufficient number of studies
- 2nd challenge: data gaps
  - Laboratory name, strain, weight, age
  - Information e.g. on infections that could be the reason for a control study with higher background
  - Only summary data, not generally including percentiles, mean or median
  - Frequently not appropriate to compare incidences of malformations, resorptions etc.
- 3rd challenge: Interpretation of the HCD
  - Applicants often use HCD in the argumentation to negate experimental results using HCD range
  - Competent authorities can use range to evaluate the concurrent control
- 4th challenge: Assessment of the Relevance of the HCD for the observed effect

  → HCD is only one aspect in the overall WoE evaluation of the observed effects

5th May 2022, Workshop HCD page 21 BfR

GERMAN FEDERAL INSTITUTE FOR RISK ASSESSMENT

BfR German Federal Institute for Risk Assessment

Thank you for your attention

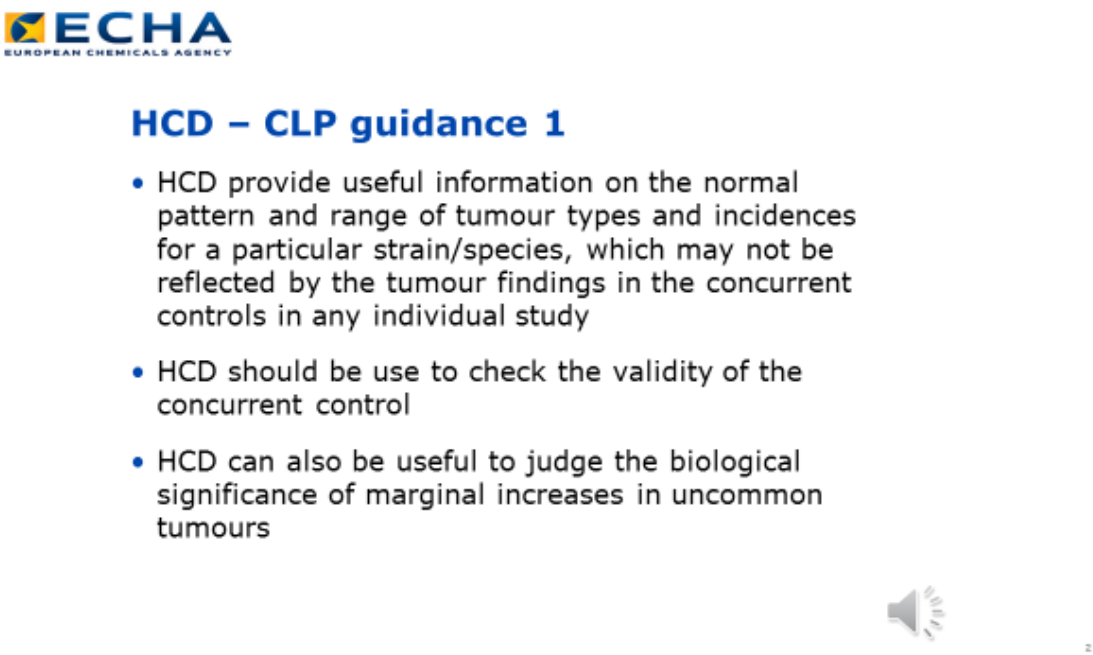
Dr. Susanne Rudzok

20 YEARS Identify Risks – Protect Health

German Federal Institute for Risk Assessment
Max-Dohrn-Straße 8-10 • 10589 Berlin, GERMANY
Phone +49 30 - 184 12 - 0 • Fax +49 30 - 184 12 – 99 0 99
bfr@bfr.bund.de • www.bfr.bund.de/en

3. Historical control data in CLH dossiers

(Chiara Perazzolo, ECHA)



ECHA
EUROPEAN CHEMICALS AGENCY

HCD – CLP guidance 2

- Use of HCD should be on a case by case basis with due consideration of the[ir] appropriateness and relevance
- HCD must be from the same animal strain/species, and ideally, be from the same laboratory
- HCD should be contemporary to the study being evaluated (e.g. within a period of up to around 5 years of the study)

ECHA
EUROPEAN CHEMICALS AGENCY

H

Often, the CLH dossier includes only the ranges

~ Used to check the validity of the concurrent control

✓ Used to check increase of uncommon tumours

✗ Due consideration of appropriateness and relevance

HCD must be …

✓ same animal strain/species

✓~ ideally from the same laboratory

✓~ HCD should be contemporary to the study being evaluated (e.g. within a period of up to around 5 years of the study)

ECHA
EUROPEAN CHEMICALS AGENCY

Comparison with HCD instead of concurrent control

90 d study mice in STOT RE

Dose (ppm)	0	1000	3000	10000
Atrophy of the seminiferous tubules of the testes	0	0	1	1
Hepatic microgranuloma (F)	0	0	0	2

- Seminiferous tubules atrophy: *The very low incidence of this finding was within the laboratory historical control range (0/10 – 1/10). Therefore, the observation of atrophy of the seminiferous tubules in this study is considered to be not treatment-related.*
- Hepatic microgranulomas (F): *were within the laboratory historical control range (0/10-2/10) and are therefore considered to be not treatment-related.*

ECHA.EUROPA.EU

ECHA EUROPEAN CHEMICALS AGENCY

Comparison with HCD instead of concurrent control

90 d study mice in STOT RE

Dose (ppm)	0	1000	3000	10000
Atrophy of the seminiferous tubules of the testes	0	0	1	1
Hepatic microgranuloma (F)	0	0	0	2

- Seminiferous tubules atrophy: *The very low incidence of this finding was within the laboratory historical control range (0/10 – 1/10). Therefore, the observation of atrophy of the seminiferous tubules in*

Seminiferous tubules atrophy[1] was observed at ≈409, 1387 mg/kg bw/d

above the guidance values for STOR RE 2 classification (max 100 mg/kg bw/d)

ECHA.EUROPA.EU

ECHA EUROPEAN CHEMICALS AGENCY

Comparison with HCD instead of concurrent control

90 d study mice in STOT RE

Dose (ppm)	0	1000	3000	10000
Atrophy of the seminiferous tubules of the testes	0	0	1	1
Hepatic microgranuloma (F)	0	0	0	2

Hepatic microgranuloma (F) was observed at ≈ 1555 mg/kg bw/d

well above the guidance values for STOR RE 2 classification (max 100 mg/kg bw/d)

- Hepatic microgranulomas (F): *were within the laboratory historical control range (0/10-2/10) and are therefore considered to be not treatment-related.*

ECHA.EUROPA.EU

Example: a bit of everything

ECHA
EUROPEAN CHEMICALS AGENCY

Time range

- HCD were provided from the laboratory where the carcinogenicity study in rats was carried out. This included the **incidences** of **hepatocellular adenoma and carcinoma** in control male F344 rats in studies carried out from **1978 – 2011**. The incidence ranges of adenoma and carcinoma during this period were **0 – 12% and 0 – 4%, respectively**

10

ECHA
EUROPEAN CHEMICALS AGENCY

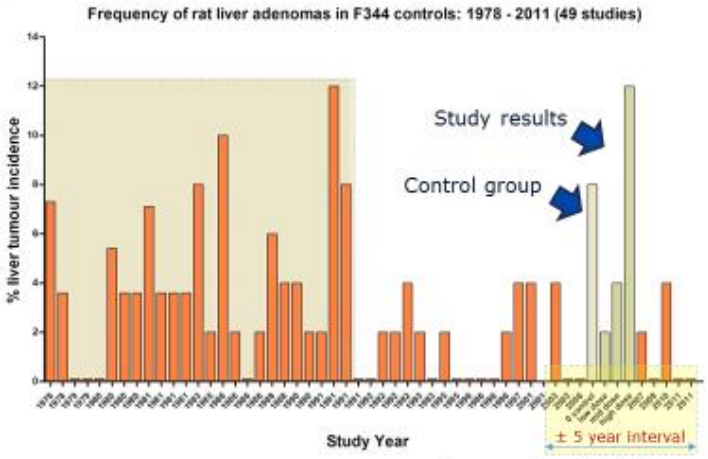
HCD appropriateness and relevance

- Closer analysis of the HCD showed that the **majority of the higher incidences of adenoma and carcinoma** occurred between the years **1980 and 1986**, which indicates that tumour incidences in control animals may have changed with time. Taking this into account, and utilising only the **studies within a 5 year time period of the concurrent study, the incidence of adenoma ranged from 0 – 4% and carcinoma incidence was 0%**

11

ECHA EUROPEAN CHEMICALS AGENCY

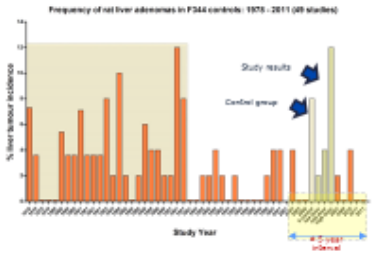
HCD incidence adenomas per year



Concurrent control: 8%

ECHA EUROPEAN CHEMICALS AGENCY

Adenomas over time

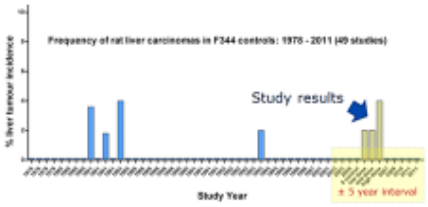


	5 years HCD	Full HCD range (1978-2011)
# studies	10	49
Mean	1.25%	2.7%
Range	0 - 4%	0 - 12%

HCD of the full time span (1978-2011) are not representative of the conditions at the time of the study
→ provide inaccurate information

ECHA EUROPEAN CHEMICALS AGENCY

Carcinomas over time



	5 years HCD	Full HCD range (1978-2011)
# studies	10	49
Mean	0%	0.25%
Range	0%	0 - 4%

HCD of the full time span (1978-2011) are not representative of the conditions at the time of the study → provide inaccurate information

14

ECHA EUROPEAN CHEMICALS AGENCY

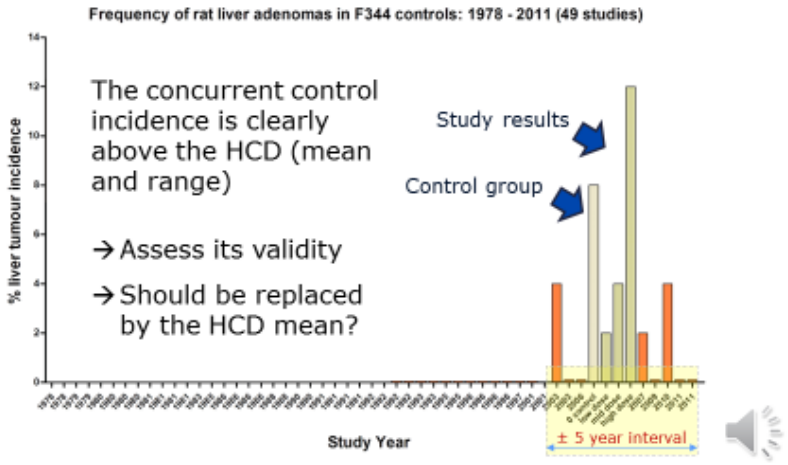
HCD from a different laboratory

- The Applicant provided further examples of HCD for spontaneous hepatocellular adenoma and carcinoma in male F344 rats taken from national databases. These included a paper by the **US NTP** that indicated **maximum incidences of adenoma and carcinoma in this strain of male rats of 10% and 6%, respectively** (Haseman, et al., 1998) and a report by **Charles River** showing incidences of **hepatocellular adenoma and carcinoma of 4.3% and 3.3%**, respectively (Lang, 1990).

15

ECHA
EUROPEAN CHEMICALS AGENCY

Checking concurrent control



16

ECHA
EUROPEAN CHEMICALS AGENCY

CLH dossier: DS conclusion

- In a carcinogenicity study in rats, a small increase in the incidence of liver tumours was observed in males treated with […]. The increases observed were **above the concurrent control values; and above the contemporary laboratory control incidence of 0%**.

+ Used the appropriate time range (5 years)

- No discussion and conclusion on the validity of the concurrent control

17

Example:

From the same laboratory

18

ECHA
EUROPEAN CHEMICALS AGENCY

Up to few years ago …

- In the absence of HCD from the in house laboratory other sources were included
- … and sometimes even if they were available
- Not seen recently

19

ECHA
EUROPEAN CHEMICALS AGENCY

Reproductive toxicity study in rabbits

Cleft palate (1 foetus from one litter) was reported in the high dose. No HCD from the laboratory.

Published HCD: cleft palate foetus/litter 4/4 (0.052%/0.35%)

Although these HCD should be used with care as they relate to a different laboratory, it indicates that cleft palate is a rare malformation in the rabbit but a single incidence cannot be attributed with certainty to the treatment.

Good use of HCD from another laboratory

20

ECHA
EUROPEAN CHEMICALS AGENCY

HCD in the CLH dossier

Table: Incidence of Thyroid Tumours in Males from 2-Year Rat Study

Dose Level (ppm) (mg/kg bw)	0 (0)	10 (0.50)	100 (5.0)	1000 (51)	3000 (150)	(SD) BR rat HCD
Adenoma (%)	3/50 (6)	0/50	0/50	1/50 (2)	5/50 (10)	1.7 – 12%
Carcinoma (%)	0/50	0/50	1/50 (2)	1/50 (2)	2/50 (4)	0.9 – 3.9%
Adenoma and/or carcinoma	3/50	0/50	1/50	2/50	7/50	0.0 – 14%

Incidences of thyroid follicular tumours is within historical control data as published by Charles River Labs (23 studies, 1995 – 2001).

Follicular cell tumours of the thyroid gland were also slightly increased in high dose males but **did not exceed historical control data incidences**

21

ECHA
EUROPEAN CHEMICALS AGENCY

HCD in the RAC opinion

Table: *Incidence of thyroid tumours in male rats (carcinogenicity study, 1998)*

Tumour type	Tumour incidence (%) Dose (mg/kg bw/d)						
	0	0.5	5	51	150	HCD*	HCD**
Thyroid follicular tumours							
Adenoma	6	0	0	2	10	1.7-12	0-8
Carcinoma	0	0	2	2	4	0.9-3.9	0-2
Adenoma and/or carcinoma	6	0	2	4	14	0-14	

*Charles river (SD)BR rat HCD (23 studies, 1995-2001), ** Historical control range from the laboratory

HCD form the same laboratory were not initially included in the CLH dossier but provided upon request

22

ECHA
EUROPEAN CHEMICALS AGENCY

Historical control data in CLH dossiers

Improved use over time, however

- **Submission of ranges only**
- **Comparison with the range**
- Assessment of validity not explicitly stated
- Use HCD in addition to concurrent control available / not needed

23

4. Acceptability and Use of Historical Control Data in Toxicological Studies

(Thomas Hofmann, on behalf of Crop Life Europe)

CropLife EUROPE

Acceptability and Use of Historical Control Data in Toxicological Studies

Workshop 03rd to 5th May 2022

Thomas Hofmann, Jean-Christophe Garcin, Felix Kluxen

Introduction

CropLife EUROPE

- **History**
  - Historical Control Data (HCD): Pool of control responses in bioassays
  - Originally introduced by the US NTP (carcinogenicity studies)
  - Since 1970's: Growing collection of HCD's by commercial laboratories
- **Aim of toxicological bioassays**
  - Hazard identification: inherent toxicological properties
  - Hazard characterization: dose response relationship
  - Risk Assessment: Exposure scenarios of various target populations

Introduction

CropLife EUROPE

- **Sensitivity and variation of Bioassays**
  - High variability → low sensitivity: Hazard cannot be distinguished
  - Low variability: any difference from control may be identified as hazard
  - Main uses of HCD in regulatory toxicology
    - Quality assurance for the test system
    - Identify aberrant control groups
    - Distinguishing true responses from chance/aberrant findings
    - Help to judge the relevance of findings (background variability)
    - Addressing the statistical multiple comparison problem

Examples for the use of HCD

CropLife EUROPE

- **General considerations**
  - Regulatory guidelines prescribe the general design of toxicology studies and the parameters to be measured
  - Usually, one control group and three dose groups
  - Statistical evaluation
  - Statistically significant differences may suggest a compound-related effect in first instance or *vice versa*
    - However, this may not always be true

Examples for the use of HCD

CropLife EUROPE

- **Example: Incidence of microphthalmia in an embryo-fetal study in rats**

Group	Control	Low	Mid	High
Fetuses	0/134	0/154	0/131	3/131
Litters	0	0	0	3

  - Statistical analysis: No significant differences between the groups
  - HCD: Microphthalmia is a rare malformation in rats

Examples for the use of HCD

CropLife EUROPE

- **Example: Incidence of fused lung lobes in an embryo-fetal study in rabbits**

Group	Control	Low	Mid	High
Fetuses	2/124 (1.6%)	4/142 (2.8%)	12/126 (9.5%)	15/131 (11.5%)
Litters	2/19 (11%)	3/20 (15%)	8/20 (40%)	10/20 (50%)

  - Statistical analysis
    - Significantly higher incidences at mid and high dose
  - Apparent dose-dependency suggests treatment-relationship

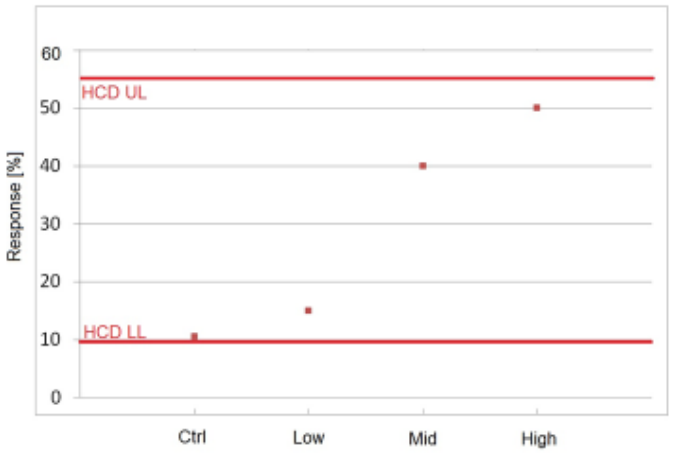
Examples for the use of HCD

CropLife EUROPE

- **However, a look to historical control incidences of the strain reveals**
  - Fused lung lobes occurred
    - From 2/142 (1.4%) fetuses to 25/142 (17.6%) fetuses
    - From 2/20 (10%) litters to 11/20 (55%) litters
  - → Control incidence of present study are at the lower border
  - → Incidence in the dose groups are within the historical range
- **Puts the relevance of these findings into a different perspective**

Examples for the use of HCD

CropLife EUROPE



Historical Data – How should they be used

CropLife EUROPE

- **These examples demonstrate the importance of HCD - however**
  - They should be used in a scientifically justified manner
  - Not intended to dismiss 'true' findings
  - Not intended to use the worst control animal as comparator for an uncomfortable finding in the high dose group
  - Establish the real importance of the bioassay
  - Note: an aberrant control group could skew the impression in both directions (e.g. a high incidence in the control group could obscure real findings in the high dose group)

The Statistical Multi Comparison Problem

CropLife EUROPE

- **Statistics usually perform Null Hypothesis Significance Testing**
  - Assumption: No differences between the groups
  - Only type-1 error (α, usually set at 5%) is controlled
  - If 3 assays are analysed, the chance of a statistically significant difference is ($1 - 0.95^3 = 14.3\%$) if there is no treatment-related effect
  - In a combined chronic toxicity and carcinogenicity test
    - Approx. 1000 tests → 50 statistically significant differences due to chance
- **Using HCD addresses the multiple comparison problem**
  - If an effect exceeds the HCD limits, more relevance is assigned

Aspects of Historical Data – Time Period

CropLife EUROPE

- **Length of the time period which should be covered**
  - Depends on the data type
  - Continuous data (e.g. hematology) or count data (e.g. uterine contents)
    - Data always measured
    - Time period of 5 years may be advisable, but an adequate number of studies should also be covered
    - 5-year period of *historical* data is *retrospective* by definition, i.e. not 2.5 years retrospective and 2.5 years in the future
    - Should this period of 5 years be increased if the number of available studies is limited? (i.e. dependent rather on the number of studies)

Aspects of Historical Data – Time Period

CropLife EUROPE

- **Length of the time period which should be covered**
  - Binary data (e.g. tumor incidences, malformations)
  - Situation is completely different
    - Data on a distinct finding are *not* necessarily always observed
    - By restricting the period to 5 years, valuable historical control data information may not be used
    - Especially information regarding rare malformation types are at risk of not being used, which would have contributed significantly to the assessment of the study

Aspects of Historical Data – Time Period

CropLife EUROPE

- **What was the rationale to narrow down the range of historical control data, and why to 5 years?**
  - Was it just a case of 'lets choose a number?'
  - Was it based on data analysis? If so, then we would expect dependence on:
    - The data type
    - The frequency of finding (common / rare)
    - The number of available studies
  - Five years are limiting too much in many cases and do not allow scientifically correct use of historical control data

Combination of Routes of Administration

CropLife EUROPE

- **Feeding and gavage studies**
  - Combining gavage and feeding studies is often critizised by authorities
  - But why should it make a difference?
  - Historical data reflect the genetic background and the biological variation
  - It is very unlikely that fetal morphology data are influenced by the route of oral administration
  - Therefore, there is no reason not to combine different administration routes together
  - This does not apply for other administration routes

Biologic Variability / Genetic Drift

CropLife EUROPE

- **Trend towards lower life expectancy in CD rats in the late 1980's**
  - Maybe caused by heterozygosity (selection pressures)
- **CRL introduced a new breeding system**
- **Ultimately new strains were created (repopulation)**
- **Genetic Drift was minimized**
- **The effect of genetic drift seems not to play a major role**

Conclusion / Outlook

CropLife EUROPE

- **Historical Control Data are a powerful tool that helps to establish the real importance of the bioassay**
- **Main uses of HCD in regulatory toxicology are:**
  - Quality assurance for the test system
  - Identify aberrant control groups
  - Distinguishing true responses from chance findings
  - Help to judge the relevance of findings (background variability)
  - Addressing the statistical multiple comparison problem

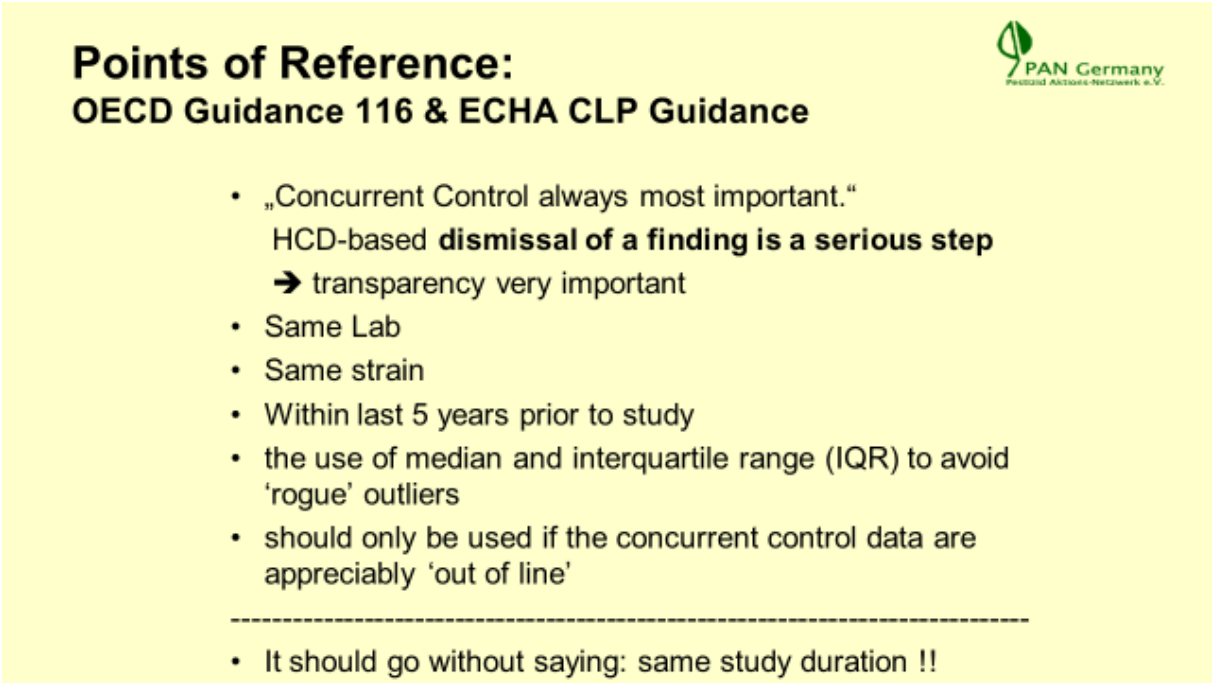
Conclusion / Outlook

CropLife EUROPE

- **They should be used in a scientifically sound way and not as a 'trick' to dismiss uncomfortable findings**
- The time period of historical control data should be
  - Retrospective from the time of report signature
  - Dependent on the data type under consideration (Continuous, Count, Binary)
  - Dependent on the frequency of the finding (common versus rare)
  - Dependent on the number of studies available
- **Gavage and feeding studies can be combined**
- **The effect of genetic drift seems not to play a major role**

5. Experience from the European pesticide evaluation – NGO perspective

(Peter Clausing, PAN Germany)



Transparency: a precondition for a proper evaluation

PAN Germany

1. Often date of in-life phase of study is missing in RARs, sometimes even date of study report
2. Median and Interquartile Range never provided (or any other method of outlier exclusion
3. Frequently: insufficient description of HCD-source
   – no clear statement whether from same lab
   – time span not mentioned
   – no statement what strain(s) were used as HCD
   – always Arithmetic Mean and „simple" Range used, instead of Median and IQR (or showing: individual data)

Importance of details for Comparability

PAN Germany

Immune System

- Stressful housing (wire-bottom vs. solid floor cages)
- Stressful housing (individual vs. group housing)

Genetic background

- **„Fischer 344" rats:** F344/DuCrl vs. F344/N vs. F344/Nhsd
- **„CD-1" mice:** Crl:CD-1 vs. Crj:CD-1 vs. Hsd:ICR (CD-1)

Avoid arbitrary discreditation of HCD (Pirimicarb)

PAN Germany

RAR 2021, Vol. 3, pp. 104-107

- Papillary cyst adenoma
  0 – 0 – 1,7 – 5.4 – **5,4%**
  HCD: 0.0 – **2.1%** (6 studies)
- Malignant liver nodules
  6.9 – 10.2 – 22.0 – 13.8 – **29.8%**
  HCD: 0.0 – **8.3%** (6 studies)
- **Dismissed "with the limitations in the historical control data"**, presumably because HCD from 6 studies an insufficient number

Phosmet

PAN Germany

RAR (2017, Vol 3, p.121) – study conclusion:

"These historical controls showed that the significance of the increased liver cell adenomas … was considered questionable" because it was lower than the HCD **(from a single other study!)**

Importance of Median and IQR

PAN Germany

Glyphosate (RAR/CLH 2021, Volume 1)

Malignant lymphomas 12% incidence (1997 mouse study)
HCD of 12 studies, mean: 6.3%, range: 3.8–**19.2%**;
11 of the 12 studies had an incidence of 6% or lower!!

- p. 289: **HCD were exceeded**, if outlier study discarded
- p. 310: **General conclusion** (Section on comparison with CLP criteria): … HCD is available, showing that the incidence at the top dose level was **within HCD range**.

Two more examples of flawed use Compound (year of assessment report)

PAN Germany

- Dimoxystrobin (2017): HCD from 22 years
- Glyphosate (2015): HCD from 22 years, 7 different laboratories and sometimes different substrains, housing on wire-bottom cages compared with "shoe box" housing on bedding

Conclusion

PAN Germany

The old way

"… is no longer considered valid as only HCD should be considered which ..." Glyphosate (RAR/CLH 2021, Vol. 1, p. 296)

- for re-assessment (old studies) – extreme care necessary when using HCD
- provide full transparency
- use Median and Interquartile Range (IQR) if the number of studies is sufficient and no other means of excluding outliers is available

6. Avian reproductive toxicity studies – ecotoxicologist perspective

(Manousos Foudoulakis (Corteva), Thomas Bean (FMC))

European Crop Protection

CLE Terrestrial Vertebrates ahT

Avian reproductive toxicity studies – ecotoxicologist perspective

Manousos Foudoulakis (Corteva Agriscience), Thomas Bean (FMC)

EFSA Workshop: 05 May 2022

Ecotoxicological relevance

European Crop Protection

- Limited guidance for ecotox (EFSA B&M 2009)-Well established in mamm toxicology
- …it is recommended that guidance is produced with mammalian toxicology experts

PPR / Pesticide Peer Review Unit

Minutes of the 3rd Meeting of the Working Group on the revision EFSA (2009) Guidance Document 'Risk Assessment for Birds and Mammals'

Held on 26 - 28 March 2019, Ede

efsa

5.10. Update on draft sections

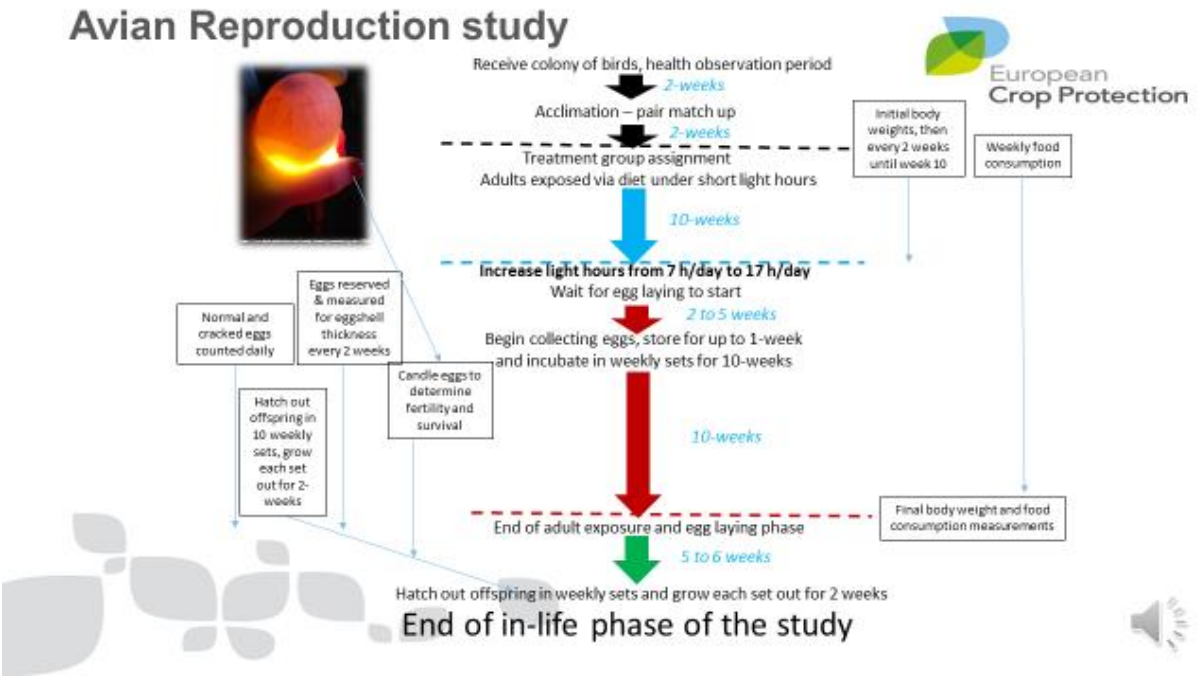
With regard to the consideration of historical control data for endpoint setting, an update on the analysis of the Valverde-Garcia et al 2018 paper was given. It was noted that this is a crosscutting issue relevant also for toxicology. It was suggested that it would be better to consider the use of historical control data, together with toxicologists, outside the guidance document for birds and mammals. The WG expressed concern that this may not be available in the short-term and therefore

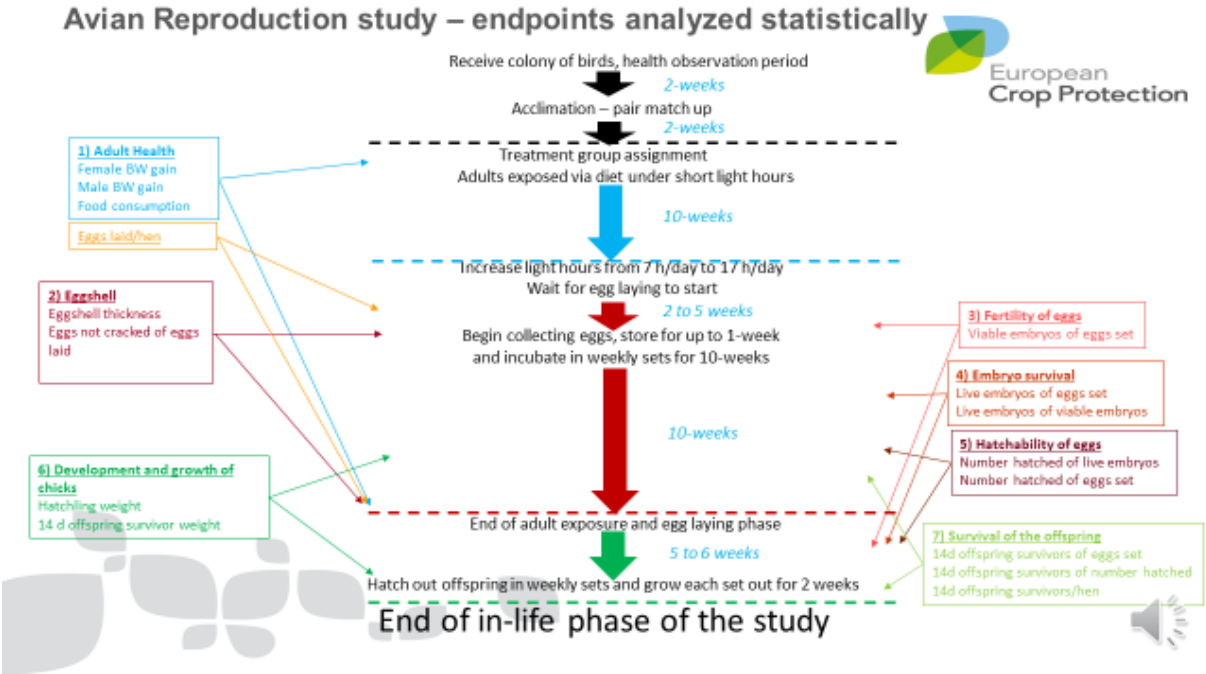
Historical control data

In EFSA (2009a), guidance was provided regarding how to use historical control data (HCD). Concerns and questions have been raised regarding the appropriateness of this guidance during both the peer review process and the public consultation. These concerns can be briefly summarised as follows:

i. Need for more guidance on evaluation and use of HCD.
ii. Need for the consideration of the biological relevance of the HCD.
iii. Consistency with mammalian toxicology, for example the timeframe needed to be considered.
iv. Relevance of HCD compared to the concurrent control, for example should one take precedence over the other

It is appreciated that HCD has the potential to make the most use of existing data and hence possibly reduce the likelihood of repeating vertebrate studies, however it is currently not possible to recommend a way forward regarding how to interpret or use HCD. It is recommended that guidance is produced with mammalian toxicology experts.


Avian Reproduction study
European Crop Protection
Receive colony of birds, health observation period
2-weeks
Acclimation – pair match up
2-weeks
Treatment group assignment
Adults exposed via diet under short light hours
Initial body weights, then every 2 weeks until week 10
Weekly food consumption
10-weeks
Increase light hours from 7 h/day to 17 h/day
Wait for egg laying to start
2 to 5 weeks
Begin collecting eggs, store for up to 1-week and incubate in weekly sets for 10-weeks
Eggs reserved & measured for eggshell thickness every 2 weeks
Normal and cracked eggs counted daily
Candle eggs to determine fertility and survival
Hatch out offspring in 10 weekly sets, grow each set out for 2-weeks
10-weeks
End of adult exposure and egg laying phase
Final body weight and food consumption measurements
5 to 6 weeks
Hatch out offspring in weekly sets and grow each set out for 2 weeks
End of in-life phase of the study


Avian Reproduction study – endpoints analyzed statistically
European Crop Protection
Receive colony of birds, health observation period
2-weeks
Acclimation – pair match up
2-weeks
Treatment group assignment
Adults exposed via diet under short light hours
1) Adult Health
Female BW gain
Male BW gain
Food consumption
Eggs laid/hen
10-weeks
Increase light hours from 7 h/day to 17 h/day
Wait for egg laying to start
2 to 5 weeks
2) Eggshell
Eggshell thickness
Eggs not cracked of eggs laid
Begin collecting eggs, store for up to 1-week and incubate in weekly sets for 10-weeks
3) Fertility of eggs
Viable embryos of eggs set
4) Embryo survival
Live embryos of eggs set
Live embryos of viable embryos
10-weeks
5) Hatchability of eggs
Number hatched of live embryos
Number hatched of eggs set
6) Development and growth of chicks
Hatchling weight
14 d offspring survivor weight
End of adult exposure and egg laying phase
7) Survival of the offspring
14d offspring survivors of eggs set
14d offspring survivors of number hatched
14d offspring survivors/hen
5 to 6 weeks
Hatch out offspring in weekly sets and grow each set out for 2 weeks
End of in-life phase of the study

Draft EFSA B&M GD 2021 update

- **Valverde et al., 2018;** Brooks et al., 2019; Temple et al., 2020
- *Whilst Valverde-Garcia et al. provides a useful starting point, it is does not provide a way forward to interpret, and hence use, HCD*
- *…it is currently not possible to recommend a way forward regarding how to interpret or use*
- *EFSA comment on Valverde et al. (2018): approach seems logical, ultimately this approach means potential effects could be excluded as they are not within the range of the HCD.* EFSA misinterpret Valverde
- *In the meantime, it is recommended to compare data from treated animals with data from concurrent study control. HCD **may** be used to determine if concurrent control animals are performing within the margins of normal variability for the species and strain*
- *The above data should cover **a five-year period**, centered as closely as possible on the date of the index study. If more guidance becomes available, that should be considered to complement the information reported here.*

Valverde et al. 2018

European Crop Protection

- Studies conducted between 1985 and May 2016 (32 years)
  - 301 bobwhite quail studies (from 18 different suppliers)
  - 292 mallard duck studies (from 1 supplier)
- EAG-Easton has been performing studies in bobwhite quail and mallard duck since 1978
- Greater standardisation since 1984
  - Finalisation of OECD TG 206
- Data available: Mean ± sd for each study (no raw data or values per pen)
- Valverde et al (2018) investigated the utility of historical control data for interpreting avian reproduction studies, including power analyses to document the size effect that could be expected to be found statistically significant

Green et al., 2022


Environmental Sciences Europe

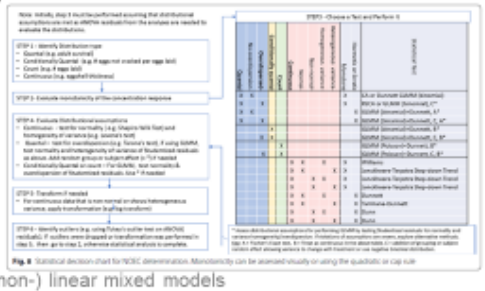
DISCUSSION Open Access

Statistical analysis of avian reproduction studies

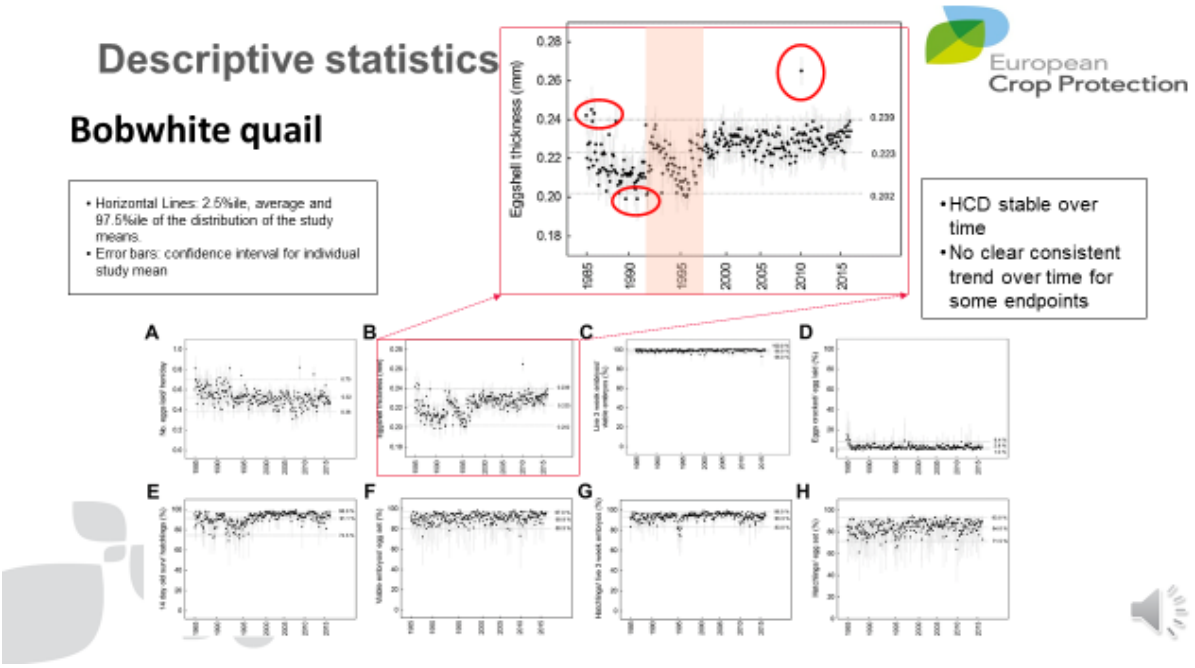
European Crop Protection

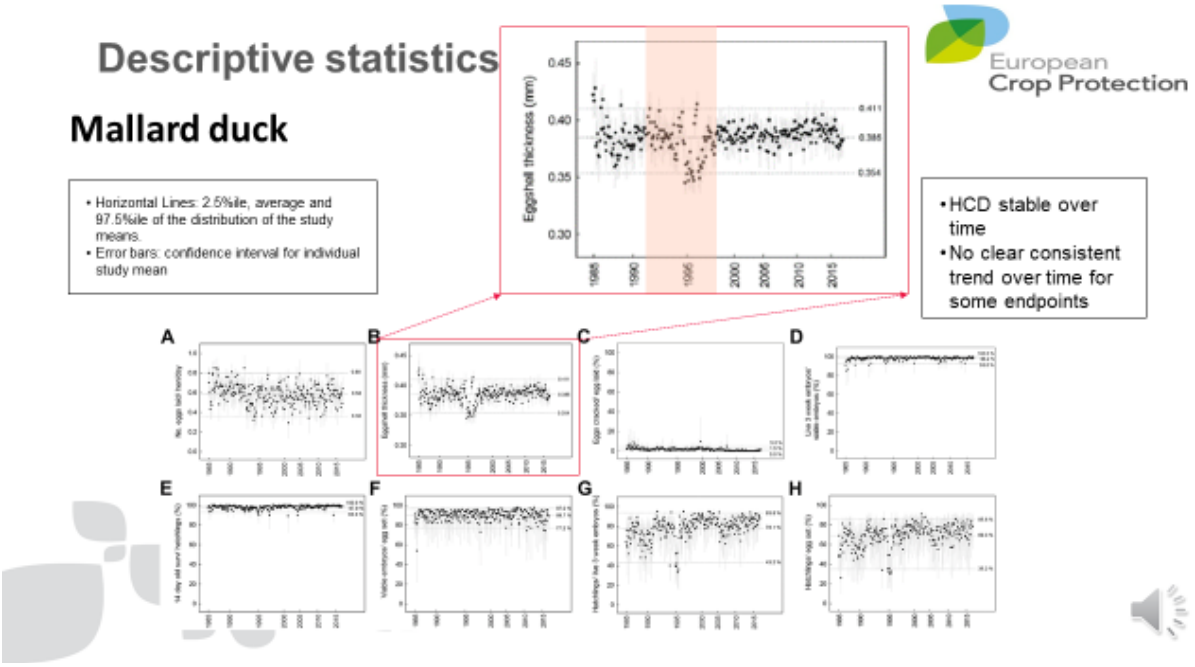
- Evaluation tools to make the most of the data we can collect, continues and extends Valverde et al. 2018
- **HCD as part of a holistic approach**
- Recommended statistical protocol

1. Assess distribution
2. Determine presence, meaning, and impact of outlier
3. Assess concentration–response monotonicity
4. **Use historical control data: EAG-Easton + Smithers HCD**
   - Smithers studies conducted between:
     - 2001–2020 for quail
     - 2004–2019 for mallard
5. Transform responses to meet test requirements or use generalized (non-) linear mixed models
6. Use regression or BMD methodology where supported by data
7. Use Model Averaging where possible for BMDx calculations
8. Assess need for special regression models
9. Consider an alternative to NOEC and BMD



- Biological relevance. Of significant importance topic in ecotoxicology


Descriptive statistics
Bobwhite quail
European Crop Protection
• Horizontal Lines: 2.5%ile, average and 97.5%ile of the distribution of the study means.
• Error bars: confidence interval for individual study mean
Eggshell thickness (mm)
0.28
0.26
0.24
0.22
0.20
0.18
0.239
0.223
0.202
1985
1990
1995
2000
2005
2010
2015
•HCD stable over time
•No clear consistent trend over time for some endpoints
A
B
C
D
E
F
G
H

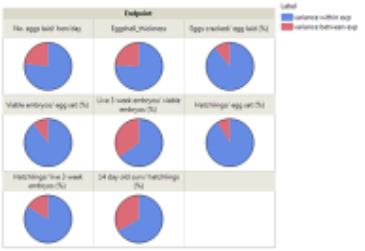

Descriptive statistics
Mallard duck
European Crop Protection
• Horizontal Lines: 2.5%ile, average and 97.5%ile of the distribution of the study means.
• Error bars: confidence interval for individual study mean
Eggshell thickness (mm)
0.45
0.40
0.35
0.30
0.411
0.386
0.354
1985
1990
1995
2000
2005
2010
2015
•HCD stable over time
•No clear consistent trend over time for some endpoints
A
B
C
D
E
F
G
H

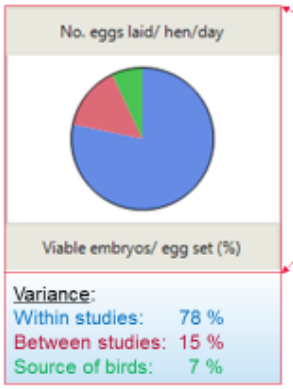
Main source of variation is within study variability

European Crop Protection

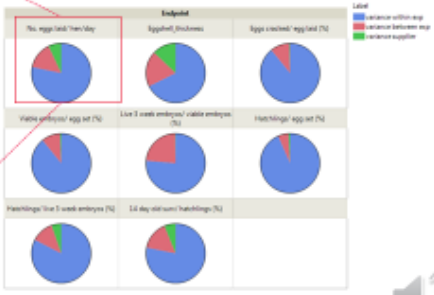
- Intrinsic & extrinsic sources of variation?
- Variance components analysis show that the observed variability is likely due to intrinsic biological variation and typical experimental variation rather than other potentially controllable factors

Mallard duck dataset (1985 - 2016)





Bobwhite quail dataset (1985 - 2016)

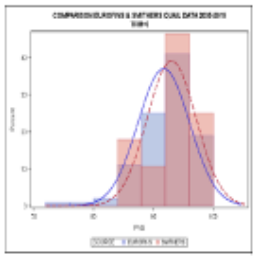


Can we pool HCD for birds?

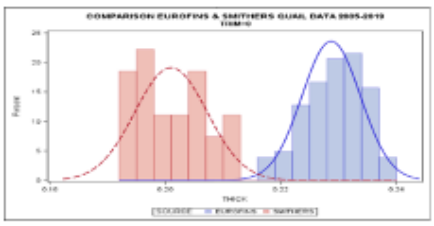
European Crop Protection

- We compared data for Smithers and Eurofins from 2005-2019
- Good agreement among the labs for most endpoints

% viable quail eggs per eggs set (quail)



Eggshell thickness (quail)



- **Data from the laboratory conducting the concurrent assay is more useful than from other laboratories for birds. HCD could be pooled across labs on an endpoint-by-endpoint basis**

Power analysis HCD

European Crop Protection

- Reinterpretation of the study over time is common, and generally more conservative endpoints are selected
- Wide range of normal responses
- Distributions vary greatly
- Careful interpretation is needed
- HCD can remove bias

Table 1 MDD% for mallard and quail

MDD%			Response	Abbrv			
Mallard					Quail		
p50	p75	p90			p50	p75	p90
18	23	35	Eggs laid per hen	EL	24	28	38
1	3	3	Eggs not cracked/number eggs laid	ENC_EL	2	3	4
8	16	24	Live embryos/number eggs set	LE_ES	8	16	25
2	4	9	Live embryos/number viable embryos	LE_VE	1	2	3
18	22	25	Number hatched/number eggs set	NH_ES	13	18	25
10	15	20	Number hatched/number live embryos	NH_LE	5	9	13
18	22	25	14-day survivors/number eggs set	HS_ES	14	19	26
2	2	3	14-day survivors/number hatched	HS_NH	4	7	10
26	29	35	Number of 14-day survivors per hen	HS	30	34	39
6	7	7	Hatchling body wt (g)	HATWT	5	6	7
6	7	8	14 Day survivor BW (g)	SURVWT	7	8	8
4	5	6	Eggshell thickness (mm)	THICK	5	6	6
8	9	12	Adult food consumption (g/bird/d)	FOOD	8	8	9
94	203	293	Adult male body weight gain (g)	WTGAINM	81	101	155
45	51	58	Adult female body weight gain (g)	WTGAINF	35	42	70
25	29	35	Number of hatchlings per hen (#/hen)	NH	29	34	38

pn = nth quantile of distribution of MDD% for indicated responses, n = 50, 75, 90

Calculations assume 18 cages of 2 birds each in every treatment group and within-study CVs from historical control data from two frequently used testing labs

Case study 1 - informal statistical reasoning can be misleading

European Crop Protection

- Regulator concluded number of eggs hatched of eggs set concluded that LOEC was 10 ppm based on >10% effect from concurrent control at 10 and 35 ppm (not based on statistically significance)

Skewness observed for 2 highest treatment groups (**not normal distribution**)

Group	Conc	Count	Mean		Median		Std
1	0	15	85.29		88.46		8.70
2	4	14	82.04	>10%	82.43	<8%	9.01
3	10	13	72.02		83.02		25.86
4	35	15	74.30		86.54		27.87

Eggs hatched per eggs set (HATCH_ES)

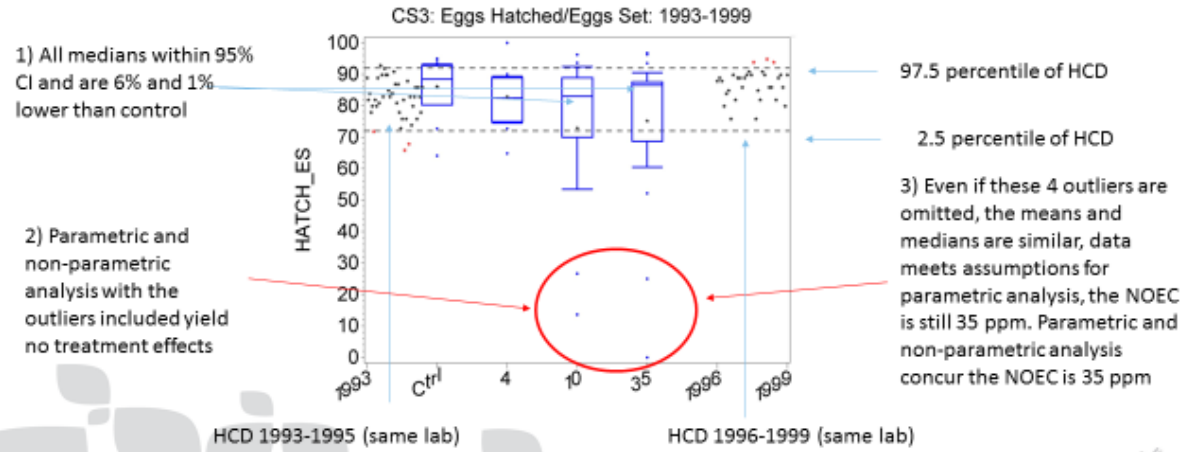
Standard deviations not homogenous

DISCUSSION Open Access

Statistical analysis of avian reproduction studies

Case Study 1 – NOEC of 35 ppm justified from multiple lines of evidence using HCD

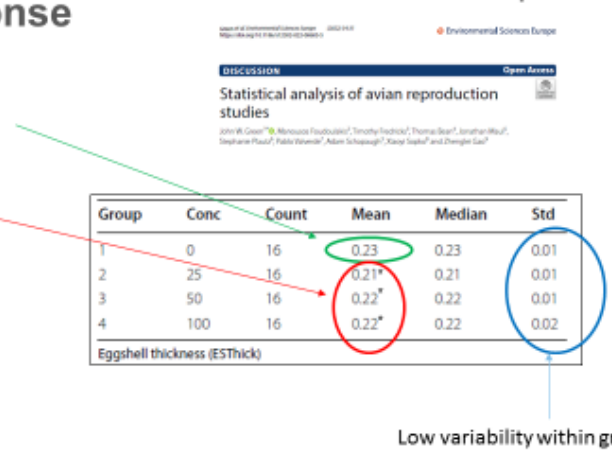
European Crop Protection



Case Study 2 – Eggshell thickness a low variability response

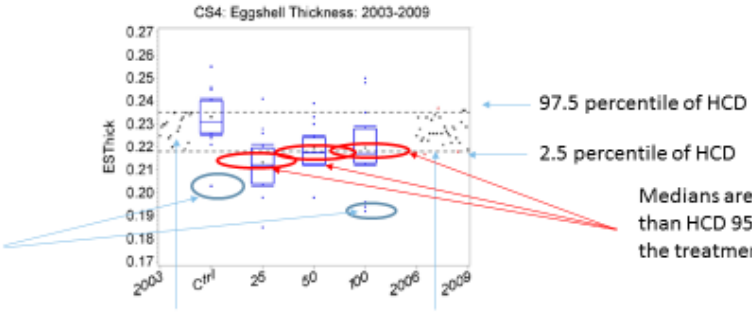
European Crop Protection

- Eggshell thickness drops from a relatively high control to a flat, non-monotonic response across the treatment groups
- Non-parametric tests run and NOEC is <25 ppm with all groups sig. different from the control
- But effects are slight and <10%



Case Study 2 – There is no question about statistical significance, the question here is about biological relevance of slight effects (maximum effect of 6%)

European Crop Protection



Even if these 3 "Tukey outlier test determined" outliers are removed and parametric stats are run, the NOEC is still <25 ppm

HCD 2003-2005(same lab) HCD 2006-2009 (same lab)

- Is 6% difference biological relevant?
- A decrease of less than 18% or 22% in eggshell thickness, is not biologically important in terms of population effect (EFSA Birds & Mammals GD 2009, EFSA biological relevance 2017, Green et al. 2022)
- Is there inter analyst variation in such a sensitive measurement?

Factors to Consider When Interpreting Slight Effects on Eggshell Thickness

European Crop Protection

SETAC NORTH AMERICA 42ND ANNUAL MEETING
14–18 NOVEMBER 2021 • VIRTUAL CONFERENCE

Figure 1: A digital Mitutoyo micrometer is used to measure 5 points on an eggshell, 3 points on one half and 2 points on the other half.

- Eggshell thickness has a low coefficient of variation resulting in effects as low as a 3% thinning from the control being reported as statistically significant. Subsequently, eggshell thickness has been one of the endpoints that most commonly drove avian reproduction study conclusions
- Preliminary results indicate inter analyst variation is greater than many of the slight effects that have driven study conclusions
- For northern bobwhite the range of measurements for an individual analyst is 2% with multiple analysts then it is generally 3% but could be as much as 9% for any individual egg
- For mallard an individual analyst's 10 repeated measurements on the same 45 eggs ranged by 7% on average and by almost 13% for the most variable eggshell
- **Due to inherent measurement variability, effects on eggshell thickness of <10% should not be interpreted as biologically significant**

What HCD time span should be used?

European Crop Protection

- Usually recommended a 5-year span centered on the starting date of the study
- **What HCD time span should be used?**
- **How many HCD studies are needed?**
  - These 2 questions are related (EFSA does not appear to relate them)
  - How many studies will be available for HCD ranges going forwards?
  - Labs doing less avian work will have still smaller numbers
  - For a new study, there will be no HCD data for dates later than the current study
    - The use of the previous 5 years or previous 20 studies are obvious alternatives in this case

What HCD time span should be used?

European Crop Protection

- Time span will **depend**:

1. **Number of studies**

- It would be best to have 20 or more studies from the HCD where possible, approximately equally split on both sides of the concurrent study date

2. **Reality check**

- Suggested 5th and 95th percentiles of the HCD are dependent on the number of studies in the HCD and a reality check would include assessing the data using several time spans, such as ± 2, ± 3 and ± 5 years in the HCD to make sure these percentiles are not overly influenced by the size or the time span of the HCD
- Note also that 5% of 20 studies is 1, so the 5% and 95% bounds on a smaller HCD are of questionable relevance

Conclusion

European Crop Protection

- Due to the nature of the test, the lack of positive control and limited number of test item treatment levels, variation that can complicate the interpretation of individual studies is not uncommon
- A key principle of ecotoxicology testing is to understand and define the baseline for this standardised test design
- Relative stability of both species' responses over a considerable time
- Variance components analysis show that the observed variability is likely due to intrinsic biological variation
- Case studies indicate determining the true treatment-related effect can be challenging while HCD can be useful in interpreting ecotoxicology studies

Conclusion

European Crop Protection

- It would be best to have 20 or more studies from the HCD where possible. A reality check assessing the data using several time spans is also proposed
- Biological relevance or/and inherent measurement variation should be always considered
- Draft EFSA B&M GD update 2021 seeks a way forward regarding how to interpret or use HCD. Green et al., 2022 continues & extends Valverde et at 2018 work. A toolbox of different approaches (statistics, biological relevance, BMD, HCD) can improve interpretation of standard avian reproduction studies by providing further context in the data assessment

- **For birds, the use of HCD should be more routinely considered to help decision making, aid data interpretation, following principles established for mammalian toxicology in a holistic assessment**
- Industry will continue to contribute the data needed for all stakeholders to have the confidence to integrate HCD into the holistic toolbox

Terrestrial Vertebrate Ad-Hoc Team

References – Terr Vert CLE projects

European Crop Protection

- Brooks Amy C., Foudoulakis Manousos, Schuster Hanna S., Wheeler James R. (2019). Historical control data for the interpretation of ecotoxicity data, are we missing a trick? Ecotoxiciolgy 28:10:1198-1209. DOI: 10.1007/s10646-019-02128-9
- EFSA Scientific Committee, Hardy A, Benford D, Halldorsson T, Jeger MJ, Knutsen HK, More S, Naegeli H, Noteborn H, Ockleford C, Ricci A, Rychen G, Schlatter JR, Silano V, Solecki R, Turck D, Younes M, Bresson J-L, Griffin J, Hougaard Benekou S, van Loveren H, Luttik R, Messean A, Penninks A, Ru G, Stegeman JA, van der Werf W, Westendorf J, Woutersen RA, Barizzone F, Bottex B, Lanzoni A, Georgiadis N and Alexander J (2017). Guidance on the assessment of the biological relevance of data in scientific assessments. EFSA Journal 2017;15(8):4970, 73 pp. https://doi.org/10.2903/j.efsa.2017.4970
- Green John, W., Foudoulakis Manousos, Fredricks Timothy, Bean Thomas, Maul Jonathan, Plautz Stephanie, Valverde Pablo, Schapaugh Adam, Sopko Xiaoyi and Gao Zhenglei (2022). Statistical analysis of avian reproduction studies. Environmental Sciences Europe: 34:31. https://doi.org/10.1186/s12302-022-00603-5
- Stanfield Kelley and Bean Thomas (2021). Factors to Consider When Interpreting Slight Effects on Eggshell Thickness. Poster Presentation at SETAC North America 42nd Annual meeting – Virtual Conference, Abstract no. 03.07.17. https://setac.confex.com/setac/sna2021/meetingapp.cgi/Paper/5909
- Temple, Diana, Springer Timothy, Gallagher Sean, du Hoffmann Gunther and Wheeler, James R. (2020). The value of avian gross pathology in identifying endocrine disrupting properties. Regulator Toxicology and Pharmacology: 112: 104577. https://doi.org/10.1016/j.yrtph.2020.104577
- Valverde-Garcia Pablo, Springer Tim, Kramer Vince, Foudoulakis Manousos, Wheeler James R. (2018). An avian reproduction study historical control database; A tool for data interpretation. Regulatory Toxicology and Pharmacology. 92: 295.302. https://doi.org/10.1016/j.yrtph.2017.12.004