ATC comments on the proposed Harmonised Classification and Labelling for:

Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene (EC 270-128-1)

Submission to the ECHA Public Consultation

Document 156
April 2024

ATC - Technical Committee of Petroleum Additive Manufacturers in Europe AISBL
Registered in Belgium: 0694709743
Registered Address: Avenue de Tervueren 188A, box 4, B-1150 Brussels, Belgium

Introduction

The Technical Committee of Petroleum Additive Manufacturers in Europe (ATC) was established in 1974 for member companies to discuss topics of a technical and statutory nature which are of concern to our industry. ATC works to develop sound scientific responses to regulatory changes, for the benefit of end consumers and environmental protection.

ATC's position is that the proposed Harmonised Classification and Labelling (CLH) of Reprotoxicity Category 1B for EC 270-128-1 (which will lead to an unwarranted and overly conservative Restriction on consumer uses per REACH Annex XVII), is not justified based on the available toxicology data for the reasons summarised below and in the annexes to this document.

Along with EC 701-385-4 (subject to a parallel CLH proposal), this substance has proven to be a highly effective antioxidant for lubricants which are essential for transport, power generation and a range of other industries. Use of this substance in lubricants allows equipment and vehicle manufacturers to comply with increasingly stringent fuel efficiency and emission targets, to enhance hardware durability, and to reduce the use of chemicals and mineral oils, thereby benefitting the European economy, society and environment and contributing to the sustainability goals of the EU Green Deal.

Summary of ATC's position on the proposed classifications

ATC disagrees with the assessment of the dossier submitter (DS) and does not consider the classification of Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene as Repr. 1B (H360 FD) and Aquatic Chronic 2 (H411) warranted based on the reasoning provided in this document. Multiple shortcomings were identified in data presentation and evaluation by the DS. These affect the overall assessment of toxicological relevance and severity of the findings observed. Since a correct and complete compilation of all available evidence and a scientifically sound data assessment is essential for a robust decision on classification, ATC asks the DS, rapporteur and the Risk Assessment Committee (RAC) to consider the comments provided in Annex 1 – Comments to the CLH dossier as well as further data evaluations presented in Annexes 1-11 before concluding on the proposed hazard classification for the substance.

1. General note on the read-across proposed by the DS

The DS has proposed to conduct a read-across between this substance and Reaction products of diphenylamine with nonene, branched (EC 701-385-4) and states that this read-across is currently also proposed in the registration dossier of Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene. While this is correct, it must be mentioned that the registrants have already decided to remove this read-across based on different results obtained in the OECD TG 421 bridging studies that were conducted with both substances after the read-across proposal. Studies to fulfill Annex VII and VIII data requirements (which are required to perform a REACH-compliant dossier update) are already ongoing and results are expected at the end of 2024. Once these data are obtained, a REACH-compliant dossier update with removal of the existing read-across will be performed.

The data obtained from the OECD TG 421 studies show distinct differences between the two substances, particularly in the areas most likely to be responsible for the mode of action proposed for Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene.

Therefore, the read-across approach proposed for the reproductive and developmental endpoints is considered not sufficiently reliable.

For further information on why the read-across is no longer supported, please refer to the comments submitted for Reaction products of diphenylamine with nonene, branched evaluated in parallel.

2. Toxicity to reproduction: sexual function and fertility

Lower numbers of implantation sites with subsequently smaller litter sizes, decreased ovary weight and altered cyclicity are identified by the DS as basis for the proposed classification regarding fertility. These are discussed in turn below and in more detail in Annexes 2-4.

Implantation sites and litter size

While the DS evaluates the available toxicological data as consistent and clear evidence for a reduced number of implantation sites caused by the substance, they make scientific errors by disregarding absence of statistical significance on some occasions, modifying historical control ranges to artificially generate toxicological relevance, and ignoring accompanying systemic toxicity. Implantation sites and subsequent litter size changes reached statistical significance in studies conducted with the substance only at the highest dose level and accompanied by clear systemic toxicity. The mean number of implantation sites, if statistically significantly reduced, was only slightly below the historical control range considered relevant for data evaluation by the testing laboratory in all studies with the substance. A more detailed discussion of the available data on implantation sites and subsequent impact on litter size is provided in Annex 2 – Detailed information on effects on implantation sites and systemic toxicity**.**

Ovary weights

A statistically significantly decreased absolute ovary weight is identified by the DS in the available OECD testing guideline (TG) 421 and 443 studies with the substance (BASF SE, 2021, 2020a). However, the DS disregards the decreased body weights clearly impacting on organ weights – in line with this, relative ovary weight is not statistically significantly affected by the substance in any of the studies. Moreover, no histopathological correlation was observed in any of the studies with the substance. More details on this are provided in Annex 3 – Detailed information on ovary parameters**.** Applying internationally accepted standards for data interpretation, these changes in ovary weights should not be considered as toxicologically relevant and most importantly, not considered relevant to support a classification as Repr. 1B.

Estrus cycle changes

Altered cyclicity is another argument raised by the DS to substantiate the need to classify the substance as reproductive toxicant, Repr. 1B. Significantly increased estrus cycle length was found in the high dose groups of the OECD TG 421 study and one of three cohorts of the OECD TG 443 study performed with the substance. No changes in estrus cycle length were found in the other two cohorts of the OECD TG443 study up to the highest dose tested. The increased estrus cycle lengths were associated with a slight increase in the mean percentage of days in diestrus stage in single animals. More details are provided in Annex 4 – Detailed information on data for proposed estrus cycle changes**.** Based on the inconsistent observation of this finding and the low magnitude of the effect, the biological relevance as well as the evidence for a direct substance-related effect are questionable.

Mode of action

Taken together, these borderline effects do not provide robust evidence for a clear, intrinsic effect of the substance on female fertility or sexual function. Moreover, as described in detail by the DS, systemic toxicity observed in all studies includes decreased body weights and food consumption. The liver is identified as the primary target organ of the substance throughout all available toxicological studies with repeated dosing with the substance, as was also stated by the DS. In line with the observed liver changes, thyroid hormones, organ weights and histopathology are altered as secondary consequences to enhanced hepatic metabolic capacities. Thyroid hormone imbalances have been shown in the literature to interfere with female fertility including estrus cyclicity, ovary weights and uterine receptivity affecting early implantation. A summary of related literature and discussion of the available data is provided in Annex 2, section 4 **(**Mode of action for effects on fertility**)**. In summary, on review of the liver effects and clear secondary thyroid toxicity evident in the studies for this substance, as shown by the changes in thyroid hormone levels (TSH and/or T4) and consequently thyroid organ changes, it is likely that the female fertility findings including slightly affected cyclicity, implantations and ovary weights seen in the top dose of these studies are secondary to these clear changes in the liver and thyroid, particularly given the lack of other specific impacts on any other reproductive parameters.

Comparison with CLP criteria

According to the CLP Annex I section 3.7.2, classification of a substance in Category 1B requires data showing “clear evidence of an adverse effect on sexual function and fertility or on development in the absence of other toxic effects, or if occurring together with other toxic effects the adverse effect on reproduction is considered not to be a secondary non-specific consequence of other toxic effects.” In contrast, a substance shall be classified in Category 2 for reproductive toxicity when there is “some evidence” of an adverse effect on sexual function and fertility, or on development, “and where the evidence is not sufficiently convincing to place the substance in Category 1”. Also, for classification in Category 2, the relevant “effects shall have been observed in the absence of other toxic effects, or if occurring together with other toxic effects the adverse effect on reproduction is considered not to be a secondary non-specific consequence of the other toxic effects.”

Section 3.7.2.2.1 further specifies that “classification as a reproductive toxicant is intended to be used for substances which have an intrinsic, specific property to produce an adverse effect on reproduction”. Section 3.7.2.4.3 states that “Classification shall not automatically be discounted for substances that produce developmental toxicity only in association with maternal toxicity, even if a specific maternally-mediated mechanism has been demonstrated. In such a case, classification in Category 2 may be considered more appropriate than Category 1.”

Effects observed were a slight but statistically significant reduction in the number of implantation sites in both a reproduction/developmental toxicity screening study and an extended one-generation reproduction toxicity study with subsequently reduced litter sizes. The findings were accompanied by pronounced maternal toxicity with effects on liver and thyroid. A mode of action has been proposed, demonstrating that primary maternal toxicity may lead to secondary effects on fertility as described in Annex 2, section 4: Mode of action for effects on fertility.

Conclusion

The available studies provide some evidence for an effect of the substance on female fertility together with maternal toxicity. The mode of action proposed shows these effects are indicative of a secondary, non-specific mechanism as a consequence of pronounced maternal toxicity. The mode of action is especially relevant as it is most likely species-specific and of low relevance for humans. In line with the criteria laid down in Regulation (EC) 1272/2008 (CLP), classification as Category 2 (H361f) is warranted.

3. Toxicity to reproduction: developmental toxicity

Changes in neurohistopathology and neuromorphometric parameters together with findings in auditory startle response (ASR), effects on viability, and pup growth are identified by the DS as basis for the proposed classification regarding development. In their evaluation of the data, the DS made several scientific errors in data assessment, statistical evaluation, and scientific data interpretation as described below and in more detail in Annexes 5-8.

Structural abnormality

The DS concludes that there is clear evidence for structural abnormalities in the central nervous system of animals exposed during the developmental period based on neurohistopathological and neuromorphometric findings in the high dose group of the OECD TG 443 study performed with the substance. While males of the developmental neurotoxicity (DNT) cohort investigated at postnatal day (PND) 77 did show increased incidences of axonal degeneration in the thoracic cord, the concurrent controls also showed an unusually high background of this parameter. Additionally, this finding was not observed in females and no clear dose-dependent increase of severity was observed. Additional investigations in pups of the DNT cohort at PND 22 could not confirm this observation, either. Therefore, it was concluded by the laboratory that this finding is due to neurotoxicity following repeated exposure rather than developmental neurotoxicity. In the OECD TG 422 study with the substance (WIL Research Europe, 2014), no histopathological effect on spinal cord was reported, however based on the technical processing conducted in a neuropathological investigation, such a finding is much more likely to be detected in the DNT of an OECD TG 443 as compared to a standard repeated dose toxicity study. An expert statement supporting this will be submitted by the REACH SIEF for the substance. Decreased brain length observed was minimal and in the same range as reduced absolute brain weights, whereas relative brain weights were normal. The neuromorphometric changes in corpus callosum observed were not considered statistically significant in the study report, however the DS conducted its own statistical analysis of the data showing statistical significance in a two-way ANOVA (sex, treatment). Of note, no other brain morphometric parameter investigated was impacted. However, investigation of the historical control data available on neuromorphometric parameters showed the findings were within the historical control data of the laboratory as well as within the coefficient of variation for this parameter and can therefore not be regarded as biologically relevant. Further information can be found in Annex 5 – Detailed information on neurohistopathology and neuromorphometrics.

The DS claims that there is some evidence that the substance induces functional neurological deficiency based on changes in the ASR investigated as part of the OECD TG 443 DNT module. In the analysis of this data package, the DS has made several grave mistakes which distort the data reported in the CLH dossier and do not reflect the data available in the study report. The relative decrease in ASR of mid and high dose group animals as compared to concurrent controls stems only from one animal in the control

group, which showed an extremely high max. amplitude. Exclusion of this outlier or use of the median (instead of mean) led to the conclusion that no decrease in mean max. amplitudes was observed in control vs. high dose animals. Further, the DS analyzed habituation response of the animals by subtracting mean values generated in blocks 1 and 5 of the ASR evaluation. Due to the high intragroup variability, this analysis is scientifically not appropriate. Rather, the habituation of single animals must be evaluated, and the behavioral trend assessed. No relevant changes in habituation were observed between control and high dose groups. Further information can be found in Annex 6 – Detailed information on proposed neurological functional deficits. Based on these data as well as on the lack of findings in the functional observational battery, it can be concluded that the exposure of animals with the substance during developmental stages does not lead to a functional impairment.

Death of the developing organism

The DS concludes that there is slight evidence that the substance induces death of the developing organism based on reduction of pre- and postnatal viability in the high dose animals of the OECD TG 422 study. Indeed, the post-implantation loss calculated by the DS from the implantation site and litter size data of this study shows higher values for all treated groups as compared to the control group. No frame of reference (historical control data) is available for this study, but the values are within the range of the naturally observed variations of the same strain in the other studies available. In addition, an increased postnatal mortality was observed in the high dose group of the OECD TG 422 study with the substance, which was largely attributed to one female, however as two other litters were affected as well with one pup each, a relationship with treatment could not be excluded. No substance-induced pre- or postnatal mortality was observed in any other study available with the substance, also at comparable dose levels. Therefore, these effects presumed relevant by the DS are of questionable biological relevance, could not be reproduced in subsequent studies and should therefore not be considered for classification purposes. Further information on data regarding this section can be found in Annex 7 – Discussion of proposed information death of the developing organism.

Altered growth

Evidence for effects of the substance on growth of the developing organism is claimed based on reduction of postnatal weights in the OECD TG 421 study. Importantly, the reductions in fetal postnatal weights were limited to the mid and high dose group and only occurred together with pronounced maternal toxicity (reduced body weight and food consumption during gestation). It is well-known that maternal toxicity can impact several developmental parameters, including fetal weights as a secondary unspecific mechanism. For further information on the data available for Reaction products of diphenylamine with nonene, branched please refer to Annex 8 – Detailed discussion of proposed effects on pup growth. Based on the well-established secondary nature of this effect, this line of argumentation should be disregarded in accordance with Regulation (EC) 1272/2008 (CLP), Annex I, Section 3.7.2.4.3.

Comparison with CLP criteria

According to the CLP Annex I section 3.7.2, classification of a substance in Category 1B requires data showing “clear evidence of an adverse effect on sexual function and fertility or on development in the absence of other toxic effects, or if occurring together with other toxic effects the adverse effect on reproduction is considered not to be a secondary non-specific consequence of other toxic effects.” In contrast, a substance shall be classified in Category 2

for reproductive toxicity when there is “some evidence” of an adverse effect on sexual function and fertility, or on development, “and where the evidence is not sufficiently convincing to place the substance in Category 1”. Also, for classification in Category 2, the relevant “effects shall have been observed in the absence of other toxic effects, or if occurring together with other toxic effects the adverse effect on reproduction is considered not to be a secondary non-specific consequence of the other toxic effects.”

Section 3.7.1.4 of Annex I, CLP state “The major manifestations of developmental toxicity include (1) death of the developing organism, (2) structural abnormality, (3) altered growth, and (4) functional deficiency.”

- Death of the developing organism: The findings claimed by the DS were attributable to a single animal and were not repeated in subsequent studies Therefore, these findings should not be taken into consideration for decision on classification.
- Structural abnormality: These findings were observed in the context of unusually high background, were only minimal or slight in occurrence and in additional investigations it could be shown that these findings were due to repeated rather than developmental exposure.
- Altered growth: The data for the substance under evaluation are limited to reduced fetal weights in reproduction studies at doses which caused pronounced maternal toxicity. Based on the well-established link between maternal and fetal body weights, these effects are not considered a direct, substance-specific effect but rather unspecific consequence of maternal toxicity and should be disregarded for classification purposes.
- Functional deficiency: Analyzing this data package, the DS has made several grave mistakes which distort the data reported in the CLH dossier and do not reflect the data available in the study report. Based on thorough scientific analysis of the data, no functional deficiency could be identified.

Conclusion

The available studies provide some evidence for a neurotoxic effect following repeated exposure, however no correlation to the developmental period could be established. In the absence of functional impairment, classification as Category 1B (H360D) is inappropriate. Effects on pup viability and growth of the organism only occurred in conjunction with excessive maternal toxicity and should therefore be disregarded based on the criteria laid down in Regulation (EC) 1272/2008 (CLP).

4. Bioaccumulation

For bioaccumulation, the DS derived a BCF of 2219 L/kg using an OECD 305 study conducted in 2000 on the MNDPA constituent of the substance. The experimental study was performed to meet the requirements of the Japanese legislation (Chemical Substances Control Law of Japan (MITI)) and it was not designed to meet the requirements of the updated OECD 305 TG or the criteria required to evaluate the ‘B’ criterion of the PBT assessment under REACH. There are, therefore, a number of shortcomings in the study, as recognized by the DS in the CLH report (e.g. use of surfactant and dissolvent, measurements were made for a group of 2 fish instead of individually). However despite this the DS has re-assessed this data against these new criteria and has produced an unrealistic worst-case evaluation.

Bioaccumulation potential has been assessed by the Registrants using experimental data, with support of QSAR modelling. An OECD TG 305 bioaccumulation study in fish was performed with the test substance Mono-Nonyl diphenylamine (MNDPA, EC248-295-7) at the

MITI institute in 2000. Different modelling approaches identified MNDPA as the most critical substance affecting bioaccumulation when used as read across for the registered UVCB. Considering the available experimental data and the QSAR data of the different models, the BCF based on experimental data is used to conservatively assess the bioaccumulation potential and used for the chemical safety assessment. This value as calculated by the study directors as 1730 L/kg. The position of the Registrants is that the substance is not bioaccumulative, as per Annex XIII of regulation 1907/2006/EC and the Guidance on information requirements and chemical safety assessment Chapter R.11 (PBT assessment, May 2008).

There are many issues when reassessing an older study under new criteria and inherent uncertainty in these estimations. In addition, the registrants consider that the shortcomings in the performed fish study contribute to an unrealistic inaccurate evaluation which cannot be used to robustly assess bioaccumulation. It is also true that the data are close to the threshold of 2000, and a valid assessment of “B” is crucial to the PBT assessment. Thus, the Registrants consider that the multiple uncertainties would best be dealt with by re-performing the bioaccumulation study in fish - as per the current testing proposal. Annex 9 – Detailed information on bioaccumulation in fish provides further discussion on this endpoint.

Conclusion

The performed bioaccumulation study revealed several shortcomings when compared against today's relevant criteria. In sum, these shortcomings impair a robust and valid evaluation of the bioaccumulation potential in fish. Therefore, the registrants strongly recommend to re-perform the bioaccumulation study in fish.

5. Chronic aquatic toxicity

A long-term daphnia study (OECD TG 211) is available for this substance based on nominal loading rates and without analytical confirmation. Since the lowest applied test concentrations reduced reproduction in the range of 4 to 8 % (statistically significant), no NOEC or NOELR could be determined. However, an EC10 could be determined which in line with regulatory guidance is preferred over NOEC values for classification and labelling. Since no analytical monitoring was performed, the DS has chosen an unnecessarily conservative assessment by setting the effect values below the lowest test concentration of 0.625 mg/L and using this value for classification and labelling. In Annex 10 - Detailed information on long term toxicity on daphnia (Daphnia magna) the registrants lay down that the evaluating approach of the DS is scientifically not justified. The significant decrease in reproduction in the lowest concentrations is due to a high consistent replication rate of the controls resulting in relative low standard deviations. The standard deviations show a strong overlapping between daphnids of the control and the subsequent test concentrations. If the values of the low concentrations were plotted against the controls, it becomes clear that the slightly reduced reproduction has no biological relevance, supporting the view of the registrants that effect values should be derived from the EC10 value. Considering this approach to be valid, there is no toxicity in the range of the water solubility and conclusively no classification is warranted.

Conclusion

In the available chronic daphnia study the lowest concentrations are statistically significant, but without biological relevance. The approach of the testing laboratory to determine an EC10 value was considered correct and in line with the regulatory

guidance. Considering this approach to be valid, there is no toxicity in the range of the water solubility and conclusively no classification is warranted.

Annex 1 – Comments to the CLH dossier

page	Citation from CLH report, proposal for Harmonized Classification and Labelling for Benzenamine, N-phenyl, reaction products with 2,4,4-trimethylpentene	Comments on the respective section of the CLH proposal
19	10.10.1 Adverse effects on sexual function and fertility	While the table heading is called "adverse effects […]" it should be pointed out that not all effects listed in the following table were indeed considered adverse. Therefore, this heading is misleading for the reader.
29	Estrous cycle: At the high dose level, a statistically significant increase of estrous cycles length (4.7 days vs. control 4.0 days) was noted as well as an increased mean percentage of days in diestrous stage (37% vs 26% in control).	Thus, the estrus cycle variability observed in this study is considered of limited biological relevance. Further details can be found in Annex 4 – Detailed information on data for proposed estrus cycle changes.
29	The mean number of implantation sites was significantly lower (-36%) in the high-dose group compared to the concurrent control group (8.8 vs 13.8 respectively). In the mid-dose group, while non-statistically significant, a decrease of 14% as compared to control group was already observed (Table 13). Considering the statistical significance at the high-dose (despite the low number of animal per group in a screening study) and the dose-response relationship, these changes are considered to be treatment-related and adverse in view of their magnitude.	While the finding in the high dose group was statistically significant and outside of the historical control data range of the laboratory, the mid dose group values were neither statistically significant nor were they outside of the normal biological variation observed for this rat strain in the laboratory. They should therefore be regarded as not biologically relevant and surely not adverse, as is suggested by the DS. Further information can be found in Annex 2 – Detailed information on effects on implantation sites and systemic toxicity.
29	The historical control range provided in the FSR is 9.0-13.2 pups delivered per litter, while the HCD range from the EOGRTS is 10.3-14.9 and 10.9-14.9 when considering a more appropriate timeframe (2015-2018).	Historical control data from one study protocol can only be adapted to another study protocol under the considerations of e.g., animal age, treatment protocol, group size, etc., and expert judgement on the study conduct is required in order to assess which data are adequate for consideration as historical control data. Expert judgement of the laboratory that conducted the study was that historical control data of OECD TG 443 studies are not adequate for assessment of data obtained in OECD TG 421/422 studies. Updated historical control data from OECD TG 421/422 studies were compiled which were conducted between 2015 and 2022 by the laboratory. These include 79 studies, with the following exposure routes: diet (11), gavage (58), drinking water (6), inhalation (4). More information can be found in the section "Historical control data".
30	In females, a significant decrease (-25%) in absolute ovarian weights was observed at the high-dose level, while not statistically significant, the relative weight was also decreased (-12%) (Table 14). Considering the statistical significance (despite the low number of animal per group in a screening study), the decrease in ovary weight is considered to be treatment-related and adverse at the high-dose level.	While the results in the OECD TG 421 study show a reduction in absolute ovary weights in the high dose group, the DS failed to communicate a concurrent reduction in body weights. Therefore, relative ovary weights were not significantly reduced and are considered of higher relevance for assessment. Further information can be found in Annex 3 – Detailed information on ovary parameters.
30	Table 15 : Absolute and Relative ovarian weights (Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene, OECD TG 421)	While this table does show the absolute and relative ovary weights of the respective study, it does omit the standard deviations, showing a high variability for this parameter. Therefore, the reader is unable to understand the scientific context of this information from the data provided. Further information can be found in Annex 3 – Detailed information on ovary parameters.

page	Citation from CLH report, proposal for Harmonized Classification and Labelling for Benzenamine, N-phenyl, reaction products with 2,4,4-trimethylpentene	Comments on the respective section of the CLH proposal
30	The main target organ was the liver. At the high-dose, increased absolute and relative liver weights were noted in males (+28% and +43% respectively) as well as increased relative liver weight in females (+30%) over concurrent controls), corroborated by histopathological findings (minimal to moderate centrilobular hypertrophy in all animals and minimal to slight fatty change in 6/10 males and 5/10 females, minimal single cell necrosis in 10/10 males) and changes in biochemical parameters related to liver functions (increased alkaline phosphatase activities, increased γ-glutamyl transferase activities, increased triglyceride and cholesterol values, decreased albumin and total protein). The same pattern of liver effects was observed at the mid-dose however at lower incidence and/or severity as compared to the high dose group. In the low-dose group, hepatic effects were limited to minimal centrilobular hypertrophy and minimal single cell necrosis/apoptosis in 2/10 males while in females an increase of alkaline phosphatase activities (+78%) was observed.	The information provided on histopathological findings in the liver is not correct. In the high dose group, focal necrosis was observed in 1/10 males, however no single cell necrosis was diagnosed.
30	Thyroid effects were observed from the low dose levels (minimal hypertrophy/hyperplasia of follicular cells of 3/10 males and 3/10 females in combination with altered colloid). In high-dose animals, thyroid effects consisted in a significantly increased relative weight (17%) in males, corroborated by hypertrophy/hyperplasia of follicular cells of 9/10 males (minimal to moderate) and of 6/10 females (minimal to mild) in combination with altered colloid. In males, thyroxine levels (T4) were significantly reduced from the mid-dose group and TSH level statistically increased in the high-dose group while no alteration of hormones levels was observed in dams and PND13 pups.	The information provided on thyroid organ weights is not correct. In the high dose group, relative thyroid weights of males were +32% as compared to concurrent controls (not 17% as stated by the DS).
31	The Extended One-Generation Reproductive Toxicity Study (OECD TG 443, GLP compliant), with a reliability of 1 reported, except for the developmental neurotoxicity part (auditory startle response, motor activity, morphometrics) presenting some limitations and triggering a reliability of 2 (Unpublished study report, 2021), is considered as the key study for effects on sexual function and fertility. Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene was administrated via diet to groups of 10 male and 25 female Wistar rats (P0 animals) at concentrations of 0 ppm, 200 ppm (eq. to 18 mg/kg bw/d in males/females), 600 ppm (eq. to 54 mg/kg bw/d in males/females) and 1800 ppm (eq. to 167/166 mg/kg bw/d in males/females). P0 animals were treated at least for 10 weeks prior to mating to produce a litter (F1 generation). Mating pairs were from the same dose group. Pups of the F1 litter were selected (F1 rearing animals) and assigned to 4 different cohorts (1A, 1B, 2A and 2B) which were subjected to specific post weaning examinations (1A and 1B dedicated to reproductive endpoints and 2A and 2B to developmental toxicity (DNT) endpoints). Cohort 1B (= P1 generation parental animals) was selected to produce F2 pups.	The perceived limitations of the DS for the DNT cohort are not considered an acceptable reasoning to change the validity of the study. The limitations listed by the DS will be addressed in the relevant sections below. The information on animal numbers treated provided by the DS is not correct; in the P0 generation and the F1 generation Cohort 1B, 25 animals per sex and dose group were treated. F1 Cohort 1A included 20 animals per sex and dose and F1 Cohorts 2A and 2B included 10 animals per sex and dose, each.

page	Citation from CLH report, proposal for Harmonized Classification and Labelling for Benzenamine, N-phenyl, reaction products with 2,4,4-trimethylpentene	Comments on the respective section of the CLH proposal
31	With regards to dose selection, the same dose levels as those tested in the OECD TG 421 study (DRF study) could have been tested in the EOGRTS since no overt toxicity (no mortality and no clinical signs) was observed up to 3000 ppm.	The dose levels were selected based on the effects observed in the previous OECD TG 421 study conducted with an extended pre-treatment period (see section 2 of this Annex). Considering the effects observed on the livers at the high dose level and the intention to continue the study through a second-generation pairing, the top dose level of 1800 ppm was chosen with the intention to induce some toxicity but avoid severe suffering of the animals (as is prescribed in the OECD testing guideline 443). Systemic toxicity was observed with decreased food consumption, decreased body weights and body weight gains in all dose groups, with dose-dependent differences regarding severity.
31	HCD have been provided, however several shortcomings limit their reliability. The collection period generally exceeds the recommended 5-years encompassing the year of the study. The protocol of the studies is not always clearly indicated (e.g. whether OECD TG 416 or TG 443 where followed, the route of administration not indicated for all studies). The studies included in HCD are different according to the parameters considered, which further limits the transparency and readability of those data. In view of the limitations of the provided HCD, they were not be given much weight compared to the concurrent control group, which anyway represent the most relevant comparator for determining treatment-related effects if the concurrent control is not aberrant.	new historical control data from similar studies were compiled which were conducted between 2017 and 2022. Further information can be found in Annex 2 – Detailed information on effects on implantation sites and systemic toxicity.
31	Estrous cycle: while estrous cycle data from P0 females and F1C1A females did not revealed any treatment-related effect, in estrous cycle data from P1 female, the mean estrous cycle duration was: 4.0 / 4.0 / 4.0 and 4.3** (**:p<=0.01) days in control, low-dose, mid-dose and high-dose groups respectively. The slightly prolonged average in the high-dose group is mainly driven by an increased mean number of days in diestrous stage (41% vs 33% in control).	In the F1 Cohort 1A females, estrus cycle duration of control and treated animals were comparable (3.9, 3.9, 4.0 and 4.0 days in control, low, mid, and high dose groups, respectively). In the F1 Cohort 1B (P1 females), a slightly but significantly prolonged estrus cycle duration was found in the high dose group (4.0, 4.0, 4.0, and 4.3** days (**p<=0.01), in control, low, mid, and high dose groups, respectively). Since these findings were not observed in the concurrent Cohort 1A or the parental generation and no impairment of fertility was observed, the biological significance of these findings is questionable. Additionally, a cycle duration of 4-5 days is considered normal in rats and the historical control data of the laboratory show cycle durations of 3.9-4.6 days. Further information is provided in Annex 4 – Detailed information on data for proposed estrus cycle changes.
31-32	In P1 mid-dose females the mean number of implantation sites was below (-9%) the concurrent control values (11.2 vs 12.3) without reaching statistically significance. As a consequence of the lower number of implants, the mean number of F2 pups delivered per dam (average litter size) was statistically significantly below (-10%) the concurrent control values (10.8 vs 12). For both parameters the values were close to the lowest value of the provided historical control data and when considering a more appropriate timeframe (2015-2018). Considering the low reliability of the HCD, and the clear dose-response relationship, the effects are considered treatment-related from the mid-dose in P1 females.	new historical control data from similar studies were compiled which were conducted between 2017 and 2022. Further information can be found in Annex 2 – Detailed information on effects on implantation sites and systemic toxicity.

page	Citation from CLH report, proposal for Harmonized Classification and Labelling for Benzenamine, N-phenyl, reaction products with 2,4,4-trimethylpentene	Comments on the respective section of the CLH proposal
33	In females, a statistically significant decrease in absolute ovarian weights was observed in high-dose P0 females (-13%) and in high-dose P1 females (-12%). No correlated histopathological findings were detected in P0 females (while histopathology was not performed in P1 females) (Table 16).	In the OECD TG 443 study, absolute ovary weights were significantly decreased in the high dose group for both the parental (P) as well as offspring (F1, Cohort 1B) generation. However, in F1 Cohort 1A, the absolute ovary weights were in the same range as the concurrent controls. The DS omits, however, that no significant or biologically relevant changes were observed in the relative organ weights. Therefore, the decreased absolute ovary weights are considered secondary to reduced body weights rather than substance-specific organ toxicity. This is supported by the lack of histopathological findings in the rat ovaries, thus demonstrating that the reductions in absolute ovary weights should be considered due to a general systemic toxicity including reduced body weights rather than an organ-specific reprotoxic effect. Further information can be found in Annex 3 – Detailed information on ovary parameters.
33	Table 17: Absolute and Relative ovarian weights (Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene, OECD TG 443)	This table – in contrast to the description in the title – only contains the absolute ovary weights observed for P0 and P1 (F1 Cohort 1B) animals. It omits the information on relative ovary weights, standard deviations are not provided for any of the parameters, and ovary weights for F1 Cohort 1A animals are not included at all. Tables containing full information on ovary weights can be found in Annex 3 – Detailed information on ovary parameters.
33	The delay in puberty onset was observed in both sexes of the high-dose group, which does not suggest an endocrine mode of action. Furthermore, a decreased body weight was noted at weaning in high-dose animals (-12%) while the weight at puberty onset is similar in all groups (Table 17 and Table 18). The delay in sexual maturation observed at the high dose level is therefore considered as a consequence of the delayed general development (lower pup weights).	This interpretation is supported by the authors of these comments, as well as by the Registrants of the substance under REACH. The apparent delay in the sexual maturation of both males and females of the high dose group is small, the timing well within the historical control range of the laboratory and associated with lower body weights of the affected offspring. The body weight at which puberty was reached was below that of the control animals, suggesting that the pup developmental delay due to reduced nutritional status of the dams also led to a delay in sexual maturation of the female offspring. The correlation between reduced pup weights and subsequent delays in development including sexual maturation have in the past been well established(Melching-Kollmuss et al., 2017). Further supporting this argumentation, the direction of change for the delay in vaginal opening and preputial separation is not consistent with the expected findings for estrogenic or anti-androgenic effects. Estrogens typically advance vaginal opening so that it is significantly earlier not delayed; therefore, the slight 1-day delay observed in both vaginal opening and preputial separation (same direction of change) is not consistent with a direct hormonal effect of the substance.

page	Citation from CLH report, proposal for Harmonized Classification and Labelling for Benzenamine, N-phenyl, reaction products with 2,4,4-trimethylpentene	Comments on the respective section of the CLH proposal
34	At the high-dose, increased absolute and relative liver weights were noted in P0 males (+15%/+19%), P0 females (+23%/+36%) and P1 females (+19%/+32%) as well as increased relative liver weight in P1 males (+14 %%) over concurrent controls. Histopathological correlates in P0 were: minimal to moderate centrilobular hypertrophy in all males and in 12/20 females, fatty change in 14/20 (minimal to severe) and diffuse hepatocellular hypertrophy in 6/20 females (slight to moderate) associated with changes in biochemical parameters related to liver functions (increased alkaline phosphatase activities, increased γ-glutamyl transferase activities, increased triglyceride and cholesterol values, decreased albumin and total protein). The same pattern of liver effects was observed at the mid-dose however at lower incidence and/or severity as compared to the high dose group. In the low-dose group, hepatic effects were limited to minimal centrilobular hypertrophy 4/20 P0 males and 3/20 P0 females as well as an increase of alkaline phosphatase activities in both sexes and a slight decreased of albumin level in females.	With regards to the effects described, several data points are missing which are required for a full understanding of the data available. For liver weights, data from F1 Cohort 1A were omitted. In the high dose group, absolute and relative liver weights in F1 Cohort 1A males were -1%/+9%** and for females +21%**/+31%** compared to concurrent controls. Histopathological finding of fatty change in P0 generation animals was in 14/20 males (the sex specification was missing in the text). In the low dose group of P0 animals, liver fatty change was observed in histopathological investigations in 2/20 males. In addition to the data described by the DS, histopathological investigations are also available for the F1 Cohort 1A animals showing similar findings with liver hepatocellular hypertrophy in 20/20 males and 18/20 females of the high dose group as well as liver fatty change in 10/20 males of the high dose group. This was supported by clinical chemistry findings in the mid and high dose groups. Low dose group females of Cohort 1A also showed significantly increased ALP activity, however no histopathological findings were reported. However, due to a concurrent decrease in albumin, these findings were considered treatment-related and adverse.
34	Thyroid effects were observed from the low-dose group (minimal hypertrophy/hyperplasia of follicular cells of 7/20 P0 males and 2/20 P0 females in combination with altered colloid). In high-dose animals, thyroid effects consisted in significantly increased absolute and relative weights in P0 males (+24%/+27%) and increased relative weight in P0 females (+21%), corroborated by hypertrophy/hyperplasia of follicular cells of 15/20 males (minimal to moderate) and of 12/20 females (minimal to slight) in combination with altered colloid. In P0 males, T4 were significantly reduced from the low-dose group while TSH levels were statistically increased from the low-dose group in P0 females.	Regarding the thyroid effects observed in the P0 generation, the occurrence of hypertrophy/hyperplasia of follicular cells was reported incorrectly by the DS; in females 16/20 (not 12/20) animals showed minimal to slight hypertrophy/hyperplasia of thyroid follicular cells. Thyroid altered colloid was found in 19/20 females and 17/20 males.
35	While not statistically significant, the mean number of implantation sites was decreased (-16%) in the high dose group (Table 19).	While the DS claims that there was a decrease of 16% in implantation sites in the high dose group, this was not considered relevant by the study director having considered all the data. Further, this effect was not statistically significant. Therefore, this study should not be further considered for an assessment of substance-specific effects on implantation sites. Further information can be found in Annex 2 – Detailed information on effects on implantation sites and systemic toxicity.
35	Significant increased relative liver weight was observed from the low-dose level in males and from the mid-dose in females, associated with histopathological findings (hepatocellular hypertrophy and hepatocellular vacuolation) and changes in biochemical parameters related to liver functions (increased alkaline phosphatase activities, triglyceride and cholesterol values, decreased albumin and total protein and bile acids) from the mid-dose level. Thyroid effects were observed from the low dose level (minimal to slight hypertrophy of follicular cells) in males. Decreased T4 and increased TSH levels were observed in all treated groups (both sexes, not always statistically significant and no clear dose-response relationship).	In addition to the description by the DS, also absolute liver weights were significantly increased from the low dose level in males. However, at the low dose level these changes were considered not toxicologically relevant due to a lack in concurrent changes in clinical chemistry and histopathology.

page	Citation from CLH report, proposal for Harmonized Classification and Labelling for Benzenamine, N-phenyl, reaction products with 2,4,4-trimethylpentene	Comments on the respective section of the CLH proposal
35	It is noteworthy that no effect was observed at histopathological examination of the spinal cord (cervical, thoracic and lumbar) and the sciatic nerve in male and female adults.	To put this assessment into perspective, it should be taken into consideration that during an OECD TG 422 study the fixation procedure of tissues is suboptimal for detection of these effects when compared to the neuropathological investigations performed for the DNT cohort (immersion fixation vs. perfusion fixation). In addition, the sectioning in which this was detected in the neurohistopathology of the OECD TG 443 study DNT module were longitudinal sections whereas an OECD TG 422 study only includes cross-sections of the spinal cord. Therefore, a very slight finding in the neuropathological investigation of the OECD TG 443 study DNT cohort might not have been detected in the OECD TG 422 study. Additionally, no findings in the thoracic cord were noted in generation F0 and F1A in the OECD TG 443 study. In these cohorts, the thoracic cord samples were immersion fixed and examined as cross sections comparable to the OECD TG 422 study.
35-38	Discussion of data obtained with Reaction products of diphenylamine with nonene, branched	The read-across performed by the DS is not considered reliable. For detailed discussion on read-across, please refer to the comments submitted for Reaction products of diphenylamine with nonene, branched. The registrants under REACH have previously removed the read-across for the endpoint fertility and are in the process of removing it also for all other endpoints. For further information, please refer to General note on the read-across proposed by the DS.
38	Lower mean number of implantation sites and consequently smaller litter sizes (main critical effect). The litter size (mean number of delivered pups) was significantly reduced by 34% in high-dose dose (3000 ppm eq. 271 mg/kg bw/d) dams of the OECD TG 421 study, by 20% and 18% in high-dose dose (1800 ppm eq. 166 mg/kg bw/d) P0 and P1 dams respectively and by 10% in mid-dose dose (600 ppm eq. 54 mg/kg bw/d) P1 dams of the EOGRTS. The mean number of implantation sites in high-dose (225 mg/kg bw/d) dams of the OECD TG 422 study was decreased by 16 % without reaching statistically significance. The litter size (mean number of delivered pups) was significantly reduced by 19% and 31% in the mid-dose (1500 ppm eq. 133 mg/kg bw/d) and high-dose dose (5000 ppm eq. 443 mg/kg bw/d) dams exposed to the analogue Reaction products of diphenylamine with nonene, branched. This effect was dose-related and observed from dose levels (mid-dose groups) where no effects on body-weight and food consumption were observed during the premating periods.	Implantation sites and subsequent litter size changes reached statistical significance in studies conducted with the substance only at the highest dose level and accompanied by clear systemic toxicity. The mean number of implantation sites, if statistically significantly reduced, was only slightly below the historical control range considered relevant for data evaluation by the testing laboratory in all studies with the substance. A more detailed discussion of the available data on implantation sites and subsequent impact on litter size is provided in Annex 2 – Detailed information on effects on implantation sites and systemic toxicity.

page	Citation from CLH report, proposal for Harmonized Classification and Labelling for Benzenamine, N-phenyl, reaction products with 2,4,4-trimethylpentene	Comments on the respective section of the CLH proposal
38	Decreased ovary weights. Absolute ovary weight was significantly reduced by 25% in high-dose dams exposed to Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene in the OECD TG 421 study and by 13% and 12% in P0 and P1 high-dose dams respectively in the EOGRTS. Dose-related and significant decrease in absolute ovarian weights was observed in dams exposed to the analogue Reaction products of diphenylamine with nonene, branched (by 18% and 40% at mid- and high dose level respectively) in the OECD TG 421 study. The relative weight was also impacted from the mid-dose level but statistically significantly only at the high dose level. While a statistically significant decrease (-14%) in absolute ovarian weight was also observed in the recovery group of this study (not mated females exposed to 5000 ppm during 10 weeks with a 2-week recovery period) no effect on ovary weight was noted in females up to 1000 mg/kg bw/d of the analogue by gavage in corn oil in a GLP-compliant 90-day toxicity study (Unpublished study report, 2013- Study 7). While this effect was consistently observed in functional reproductive studies, no histopathological correlate was reported in any of the studies.	A statistically significantly decreased absolute ovary weight is identified by the DS in the available OECD TG 421 and 443 studies with the substance (BASF SE, 2021, 2020a). However, the DS disregards the decreased body weights clearly impacting on organ weights – in line with this, relative ovary weight is not statistically significantly affected by the substance in any of the studies. Moreover, no histopathological correlate was observed in any of the studies with the substance. More details on this are provided in Annex 3 – Detailed information on ovary **.**
38	Effects on cyclicity. Increase of estrous cycles length associated with an increased mean percentage of days in diestrous stage was observed in high-dose dams exposed to Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene and Reaction products of diphenylamine with nonene, branched. Statistical significance was reached in high-dose females exposed to Benzenamine, N,phenyl-, reaction products with 2,4,4-trimethylpentene in the OECD TG 421 and in high-dose P1 females exposed in the EOGRTS.	Significantly increased estrus cycles length was found in the high dose groups of the OECD TG 421 study and one of three cohorts of the OECD TG 443 study performed with the substance. No changes in estrus cycle length were found in the other two cohorts of the OECD TG 443 study up to the highest dose tested. The increased estrus cycle lengths were associated with a slight increase in the mean percentage of days in diestrus stage in single animals. More details are provided in Annex 4 – Detailed information on data for proposed estrus cycle changes**.**
39	In the EOGRTS, a slight delay on puberty onset was observed in both sexes of the high-dose group while the weight at puberty onset is similar in all groups. A decreased body weight was noted at weaning in high-dose animals (-12%). The delay in sexual maturation observed at the high dose level is therefore considered as a consequence of the delayed general development (lower pup weights).	This interpretation is supported by the authors of these comments, as well as by the Registrants of the substance under REACH. For further information, see above.
40	At this dose level the systemic toxicity was not marked (i.e.: no lethality, no dramatic reduction in absolute body weight, no coma), body weights at the end of the premating period were decreased by 10% and liver and thyroid effects were observed. However, a direct correlation between systemic toxicity and reduced implantation sites could not be established based on the data available and lower numbers of implantation sites with subsequently smaller litter sizes were already observed at lower dose levels where body weight was not affected.	It should be clearly pointed out that while the definition of marked toxicity by the DS (“no lethality, no dramatic reduction in bodyweight and no coma”) might be the criteria to justify non-classification based on systemic toxicity, these are also the criteria to justify humane sacrifice during study conduct according to OECD Guidance document 19 (OECD, 2000). Therefore, it should not be the goal to achieve this level of toxicity during an OECD guideline study, given that the OECD guidelines specifically state that the high dose level should be chosen to induce some toxicity, but not severe suffering or death (OECD, 2018). Therefore, the assessment of these effects should not be as “marked” toxicity but would rather be considered “excessive” toxicity in the context of an OECD Guideline study. The effects observed on liver and thyroid at these dose levels are considered pronounced toxicity to maternal animals.
40	The consistency of the effects observed with the two structural analogues through the available studies further strengthens the causality between exposure to these two SDPAs and the observed reproductive outcomes.	In contrast to the assessment of the DS, the read-across is considered not reliable when using the RAAF criteria for assessment. For detailed discussion, please refer to the comments submitted for Reaction products of diphenylamine with nonene, branched.

page	Citation from CLH report, proposal for Harmonized Classification and Labelling for Benzenamine, N-phenyl, reaction products with 2,4,4-trimethylpentene	Comments on the respective section of the CLH proposal
49-51	Discussion of prenatal developmental toxicity data from Reaction products of diphenylamine with nonene, branched	Data on Reaction products of diphenylamine with nonene, branched are discussed in the comments submitted to the CLH proposal for this substance.
52	In the OECD TG 422 performed with Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene, the mean number of delivered F1 pups per dam (litter size) was significantly decreased (-28%) in the high-dose group resulting from both a decreased number of implantation sites and to an increase of post- implantation loss (14.1% vs 0% in controls). It should be noted that the post-implantation in control group is particularly low (0%). Table 26: Mean number of implantation sites and pups delivered pups and calculated postimplantation loss (Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene, OECD TG 422)	In the study report of the OECD TG 422 study, no information on post-implantation loss was provided. Therefore, the DS calculated the post-implantation loss numbers provided in the CLH proposal themselves. However, they used the group means as basis for calculation of post-implantation loss instead of the single animal values. Therefore, the numbers reported for post-implantation loss are not correct. Further information can be found in Annex 7 – Discussion of proposed information death of the developing organism
56	The ASR test presented some limitations: poor reporting of the apparatus used, statistical analysis not in line with the NAFTA guidance (i.e. no mention, or results presented for testing for interactions of sex, trial blocks and treatment) for maximal amplitude and latency as well as a complete absence of any statistical testing for habituation (a variable required under OECD 443). Furthermore, the lack of HCD and positive control increases the possibility the risk of false negative findings taking into account the low statistical power in DNT investigations.	The following is a detailed description of the process: On PND 24, the auditory startle response test was carried out in all animals of cohort 2A using the SR-LAB; STARTLE RESPONSE SYSTEM (San Diego Instruments, San Diego, CA, U.S.A.). For all animals, the examinations started in the morning at their respective test date. If, at a given test date, several animals were tested, the trials were conducted in a randomized sequence. Age-appropriately sized, tube shaped, transparent acrylic enclosures were used to accommodate the animals during the test. Each enclosure has an attached motion sensor and is mounted on a solid enclosure base. The entire unit is placed in a heavy wooden, plastic-laminated, isolation cabinet, which minimizes extraneous noise and vibrations. Each cabinet is equipped with internal light and fan, and contains, in a separate compartment, a complete sound generation system able to produce background noise and white noise stimuli, the level of both is adjustable. The response of the animal to the sound stimulus is issued by the motion sensor as voltage which is automatically recorded over the entire response window. The highest voltage during the response window is labeled “maximum amplitude”, the time in milliseconds from the start of the response to the maximum amplitude is labeled “latency to the peak of the response”. The animals were given a 5-minute acclimation period in the enclosure with a 70 dBA background noise. Then the startle response was recorded in 50 trials at a startle stimulus sound level of 120 dBA with a 5 - 10 second variable interval between the trials. Response was recorded for 50 milliseconds. Measurement was carried out with the light and ventilator switched on in the isolation cabinets; no food or water was provided during the test. Data (maximum amplitude, latency to the peak of the response) were analyzed in 5 blocks of 10 trials each. All records were printed, signed and are part of the raw data.

page	Citation from CLH report, proposal for Harmonized Classification and Labelling for Benzenamine, N-phenyl, reaction products with 2,4,4-trimethylpentene	Comments on the respective section of the CLH proposal
56	Therefore, the effects from the mid-dose level on the mean maximal amplitude in males as well as habituation (proxy calculation mean Block1 minus mean Block 5) from the mid-dose in males and in males and females combined are considered biologically relevant considering their magnitude in the absence of appropriate statistical analysis (testing for interactions of sex, trial blocks and treatment) and positive controls (Table 29).	The calculations conducted by the DS contain several mistakes, which distort the available data. For the comparison of the max. amplitude, only the mean values were taken into consideration, without assessment whether these values might be adequate. The control group contains one animal with an extreme max. amplitude compared to the rest of the group, therefore the mean values have a very large standard deviation. For a parameter with high intragroup variability such as auditory startle response, the median values would be more appropriate, or the mean values must exclude such outliers. Both scientifically appropriate approaches show no difference between control and high dose groups. Habituation is considered as the relative difference in maximum amplitude with which the animals react to auditory startle impulses between the first experiments and the last experiments. However, simple reduction of means is not appropriate for assessment. Rather, the auditory startle responses for each block should be evaluated and compared in curves. If this is done in the scientifically appropriate approach, no differences are observed between control and high dose animals. For further information on this part, please refer to Annex 6 – Detailed information on proposed neurological functional deficits.
56	There was a slight but statistically significant brain length reduction (-3.2%) in high dose C2A males.	The DS omits that high dose males also had smaller brain weights, which in turn explain the smaller measurements and are due to reduced animal size and weights. Correlating morphometrics with brain weights results in no difference between control and high dose animals. For further information, please refer to Annex 5 – Detailed information on neurohistopathology and neuromorphometrics.
56	However, the corpus callosum width was non-significantly increased in the high dose group by 17 and 16%, in males and females, respectively (low- and mid-dose levels not analysed). This is a rather large and biologically significant changes in the size of a brain region. Furthermore, a two way anova (sex, treatment) statistical analysis of the data performed by the DS showed that the effect is actually statistically significant. It is noteworthy that the corpus callosum is the principal inter-hemispheric myelinated tract (white matter) and histopathological findings linked to myelin degeneration in the cord white matter were observed in C2A animals.	Corpus callosum width was slightly increased in both males and females, however this increase was within the historical control data of the laboratory and the coefficient of variation observed for this parameter (see Annex 5 – Detailed information on neurohistopathology and neuromorphometrics). Assessment of the data per sex did not reveal a statistical significance. Only if both sexes are assessed together with a statistical test assuming normal distribution (which the DS has not shown is appropriate), the data are statistically significant. While the DS mentions that histopathological findings were made in the white matter, it should clearly be pointed out that no histopathological correlate or functional impairment were observed for the corpus callosum. Therefore, this change is considered not toxicologically relevant. For further information, please refer to Annex 5 – Detailed information on neurohistopathology and neuromorphometrics.
57	While not discussed in the study report and not graded, axonal degenerations were also slightly increased in other areas (2 vs 0 tibial nerve degeneration in high-dose males; lumbar cord axonal degeneration and sciatic nerve degeneration in 2 high-dose females vs 0 in controls).	The study report clearly states that “All other findings occurred either individually or were biologically equally distributed over control and treatment groups. They were considered to be incidental or spontaneous in origin and without any relation to treatment.” It can be found in the literature that axonal degeneration is commonly observed in all age groups of rats as an occasional spontaneous finding (Blankenship et al., 2016; Kaufmann et al., 2012). For further information, see Annex 5 – Detailed information on neurohistopathology and neuromorphometrics.

page	Citation from CLH report, proposal for Harmonized Classification and Labelling for Benzenamine, N-phenyl, reaction products with 2,4,4-trimethylpentene	Comments on the respective section of the CLH proposal
57	In the absence of effects in C2B males, the study author interpreted the axonal degeneration in thoracic cord of C2A males to be a chronic toxic effect rather than a developmental effect. However, exposure during the developmental period could have contributed to the delayed effects observed in C2A animals on PND77 even if not observed at an earlier time point (PND22). According to RAC note (RAC/62/2022/05) addressing developmental neurotoxicity and neurotoxicity under the current CLP hazard classes, adverse effects on the nervous system investigated or detected at any point in the life span of the organism exposed during the developmental period, covering both prenatal and postnatal development until sexual maturation (determined by preputial separation and vaginal opening), should be addressed under developmental toxicity (DNT), […]	A discussion on the evidence of these effects as developmental toxic vs. occurring after repeated exposure as well as a comment on the assessment of these data under the CLP can be found in Annex 5 – Detailed information on neurohistopathology and neuromorphometrics
57-58	T4 and TSH values in PND13 male and female pups were not affected by treatment in the OECD TG 421 performed with the substance or its analogue. In the EORGTS, T4 and TSH values in PND4 male and female pups were not statistically significantly changed (Table 31). However, in high-dose group, hormone values of only 2 pups of each sex could be measured (due to small litter size) and therefore were not included in the statistical analysis. In both sexes, T4 values from the mid-dose group were decreased as compared to the concurrent controls and outside the HCD range. TSH values in male and female PND22 pups from the mid-dose group were significantly increased. While TSH values of all test groups were included in the HCD ranges, the clear dose-response relationship strongly support a treatment related finding.	None of the studies available with the substance shows significant effects on thyroid hormones in pups prior to weaning. The DS suggests that a relevant effect might have occurred in the OECD TG 443 study even in the absence of statistical significance. As the values were low to begin with, every minimal decrease seems large in relative numbers, but without corroborating findings such as statistical significance or increases in TSH, which were clearly not observed, no toxicological relevance can be attributed. On PND 22, statistically significant increases in TSH were found in both males and females, however all values were within the HCD of the laboratory and therefore within the biological variation commonly observed for rats of this age and strain. No concurrent change in T4 values was found. Therefore, these findings must be considered of limited toxicological relevance. Further details can be found in Annex 5 – Detailed information on neurohistopathology and neuromorphometrics.

page	Citation from CLH report, proposal for Harmonized Classification and Labelling for Benzenamine, N-phenyl, reaction products with 2,4,4-trimethylpentene	Comments on the respective section of the CLH proposal
58	T4 and TSH measurements were not performed in F2 pups which is not in line with OECD TG 443 requirements.	The OECD TG 443 does not specifically require T4 and TSH measurements in F2 pups. It doesn´t even specifically require it for F1 PND 4 pups. See the excerpt of the relevant paragraphs below: 55. Systemic effects should also be monitored in F1 animals. Fasted blood samples from a defined site are taken from ten randomly selected cohort 1A males and females per dose group at termination, stored under appropriate conditions and subjected to standard clinical biochemistry, including the assessment of serum levels for thyroid hormones (T4 and TSH), haematology (total and differential leukocyte plus erythrocyte counts) and urinalysis assessments. 56. The surplus pups at PND 4 are subject to gross necropsy and consideration given to measuring serum thyroid hormone (T4) concentrations. If necessary, neonatal (PND 4) blood can be pooled by litters for biochemical/thyroid hormone analyses. Blood is also collected for T4 and TSH analysis from weanlings subject to gross necropsy on PND 22 (F1 pups not selected for cohorts). On the contrary, it provides the option of terminate the F2 litters already at PND 4 which definitely precludes blood sampling on PND 22: 53. Cohort 1B animals can be maintained on treatment beyond PND 90 and bred to obtain a F2 generation if necessary. Males and females of the same dose group should be cohabited (avoiding the pairing of siblings) for up to two weeks, beginning on or after PND 90, but not exceeding PND 120. Procedures should be similar to those for the P animals. However, based on a weight of evidence, it may suffice to terminate the litters on PND 4 rather than follow them to weaning or beyond. The request for exactly equal investigations in both F1 and F2 offspring and thus thyroid hormone measurements in the F2 was only recently (March 2023) issued in ECHA´s report of the EOGRTS review project. However, this was not yet published at the time when the study was conducted. Further information is provided in Annex 5 – Detailed information on neurohistopathology and neuromorphometrics
58-59	In respect to the observed effects related to nervous system development in this EOGRTS study, a recent review has highlighted the crucial role of THs in myelinisation process in both humans and rodents (Pagnin, 2021). In rodent models, developmental hypothyroidism interferes with neuronal migration, differentiation, and myelination and thyroid hormone regulates genes that control formation of the corpus callosum and neuronal migration (Goodman and Gilber, 2007). Hypothyroxinemia during critical windows may lead to cognitive and hearing deficits (Noyes, 2019).	In contrast to this statement by the DS, no functional impairments were observed during any of the investigations conducted, when analyzed with the appropriate scientific approach.

page	Citation from CLH report, proposal for Harmonized Classification and Labelling for Benzenamine, N-phenyl, reaction products with 2,4,4-trimethylpentene	Comments on the respective section of the CLH proposal
59-60	Death of the developing organism: Post-implantation loss and foetal viability were not affected by treatment with the analogue Reaction products of diphenylamine with nonene, branched, in both rat and rabbit PNDTSs. In the main study in rabbits, increased number of abortions (4 vs. 2 in control) was observed at the high dose level (1 abortion out of 5 does was also noted in the mid and high doses of the range finding study). While a direct effect cannot be excluded, the severe drop in food consumption GD 7-23 (up to -59% in comparison to the control group) may partly be involved, as supported by studies on caloric restriction during pregnancy in rabbit (Matsuzawa, 1981; Cappon, 2005; Matsuoka, 2006; Lopez-Tello, 2019). In the generational studies performed with Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene and its analogue, the decreased litter size at birth observed in all studies results from the decreased number of implantation sites; this effect is addressed in the section dedicated to effects to sexual function and fertility. Post-implantation loss and foetal viability were not affected in any of the generational studies except in the OECD TG 422 where the mean post-implantation loss in the high-dose dams was increased compared to controls (14% vs 0%). It is noteworthy that the value in controls of the study was particularly low. Furthermore, the viability index was significantly reduced in this OECD TG 422 at the high-dose level (88.7% at 225 mg/kg bw/d versus 100% in controls). However, no treatment-related effect was observed on live birth, viability and lactation indices in the OECD TG 421 up to 3000 ppm (eq. to 271 mg/kg bw/d) and in the EOGRTS up to 1800 ppm (eq. to 166 mg/kg bw/d) performed with Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene or in the OECD TG 421 up to 5000 ppm (eq. to 443 mg/kg bw/d) performed with its analogue Reaction products of diphenylamine with nonene, branched. → Based on the available data, there is slight evidence that Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene induces death of the developing organism based on reduction of pre- and postnatal viability in the high-dose animals of the OECD TG 422 but these effects were not reproduced in the other generational studies performed with the substance and its analogue. Post-implantation viability was not affected in the PNDTSs in rats and rabbits with the analogue. The increased number of abortions in rabbits provide slight evidence that Reaction products of diphenylamine with nonene, branched could induce death of the developing organism, however these abortions may be partly related to the severe drop in food consumption observed at this dose level, as demonstrated in published studies on caloric restriction during pregnancy in rabbit.	The DS concludes that there is slight evidence that the substance induces death of the developing organism based on reduction of pre- and postnatal viability in the high dose animals of the OECD TG 422 study. Indeed, the post-implantation loss calculated by the DS from the implantation site and litter size data of this study shows higher values for all treated groups as compared to the control group. No frame of reference (historical control data) is available for this study, but the values are within the range of the naturally observed variations of the same strain in the other studies available. In addition, an increased postnatal mortality was observed in the high dose group of the OECD TG 422 study with the substance, which was largely attributed to one female, however as two other litters were affected as well with one pup each, a relationship with treatment could not be excluded. No substance-induced pre- or postnatal mortality was observed in any other study available with the substance. Therefore, these effects presumed relevant by the DS are of questionable biological relevance, could not be reproduced in subsequent studies and should therefore not be considered for classification purposes. Further information on data regarding this section can be found in Annex 7 – Discussion of proposed information death of the developing organism. The data presented on Reaction products of diphenylamine with nonene, branched are not discussed in this document as the read-across approach taken by the DS is not considered reliable (for further information please refer to General note on the read-across proposed by the DS). A detailed discussion of these findings can be found in the comments on the CLH proposal for Reaction products of diphenylamine with nonene, branched.

page	Citation from CLH report, proposal for Harmonized Classification and Labelling for Benzenamine, N-phenyl, reaction products with 2,4,4-trimethylpentene	Comments on the respective section of the CLH proposal
60	Structural abnormality: the analogue Reaction products of diphenylamine with nonene, branched was not teratogenic in rat. In rabbit, an increase of external malformations was observed at the top dose (four fetuses in one single litter with multiple external malformations). The clustered appearance limited to one litter with similar spectrum of findings (i.e. craniofacial malformations consisting in domed head, cleft palate and small tongue in all those four fetuses) suggest rather a genetic origin than a treatment-related effect. Statistically significant increases of two skeletal variations (i.e. irregular ossification of interparietal and unossified talus) were also observed at this dose-level as well as a general delay in ossification. In the high-dose C2A animals of the EOGRTS, neurohistopathological findings (increased incidence of axonal degeneration in the thoracic cord 9/10 males as well as slight increased incidence of axonal degeneration in other area in males and females) and neuromorphometric changes (decreased brain length in males and increased corpus callosum width in both males and females) were observed. The increased of nipple number on PND13 observed in high-dose pups in the OECD TG 421 and in high-dose F2 pups on PND13 in the EOGRTS performed with Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene is considered to be rather a consequence of a general delay of pup development than a specific effect on hormonal homeostasis in the absence of alteration in other sensitive endpoints related to antiandrogenic potential. → Based on the available data, there is clear evidence that Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene induces abnormalities in the central nervous system in animals exposed during the developmental period. In the PNDTS in rabbits performed with the analogue, delay in ossification was noted in the presence of maternal toxicity. Regarding the cluster of four fetuses with multiple common malformations from a single litter of the high-dose group, a genetic origin is considered more likely than a treatment-related effect.	The DS concludes that there is clear evidence for structural abnormalities in the central nervous system of animals exposed during the developmental period based on neurohistopathological and neuromorphometric findings in the high dose group of the OECD TG 443 study performed with the substance. While males of the developmental neurotoxicity (DNT) cohort investigated at postnatal day (PND) 77 did show increased incidences of axonal degeneration in the thoracic cord, also the concurrent controls showed an unusually high background of this parameter. Additionally, this finding was not observed in females and no clear increase of severity was observed as dose response. Additional investigations in pups of the DNT cohort at PND 22 could not confirm this observation, either. Therefore, it was concluded by the laboratory that this finding is due to neurotoxicity following repeated exposure rather than developmental neurotoxicity. In the OECD TG 422 study with the substance (WIL Research Europe, 2014), no histopathological effect on spinal cord was reported, however based on the technical processing conducted in a neuropathological investigation, such a finding is much more likely detected in the DNT of an OECD TG 443 as compared to a standard repeated dose toxicity study. The neuromorphometric changes observed were not considered statistically significant in the study report, however the DS conducted its own statistical analysis of the data showing statistical significance in a two-way ANOVA (sex, treatment). Of note, no other brain morphometric parameter investigated was impacted. However, investigation of the historical control data available on neuromorphometric parameters showed the findings were both within the historical control data of the laboratory as well as within the variation coefficient for this parameter and can therefore not be regarded as biologically relevant. Further information can be found in Annex 5 – Detailed information on neurohistopathology and neuromorphometrics. The read-across approach taken by the DS is not considered reliable (for further information please refer to General note on the read-across proposed by the DS).

page	Citation from CLH report, proposal for Harmonized Classification and Labelling for Benzenamine, N-phenyl, reaction products with 2,4,4-trimethylpentene	Comments on the respective section of the CLH proposal
61	Altered growth: In rabbits exposed to the analogue Reaction products of diphenylamine with nonene, branched, a significant decrease in fetus weight (-12%) was observed at the highest dose (100 mg/kg bw/d) associated with delays of ossification. At this dose level, does consumed 31% less food than the concurrent control does during the treatment period (GD6-28), showed marked reduced defecation and had slight reduced corrected body weight. Postnatal growth was also altered from PND7 up to weaning in the high-dose groups of the generational studies performed with Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene EOGRTS, OECD TG 421) or its analogue (OECD TG 421). At these dose levels, effects on body weight of similar magnitude were observed in females at the end of the lactation period. → Based on the available data, there is evidence that both Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene and its analogue alter growth of the developing organism at dose levels also affecting maternal/parental weight.	Evidence for effects of the substance on growth of the developing organism is claimed based on reduction of postnatal weights in the OECD TG 421 study. Importantly, the reductions in fetal postnatal weights were limited to the mid and high dose group and only occurred together with pronounced maternal toxicity (reduced body weight and food consumption during gestation). It is well-known that maternal toxicity can impact several developmental parameters, including fetal weights as secondary unspecific mechanism. For further information on the data available for Reaction products of diphenylamine with nonene, branched please refer to Annex 8 – Detailed discussion of proposed effects on pup growth.
61	Functional deficiency: In the EOGRTS, despite limitations of the auditory startle response test, effects from the mid-dose level on mean maximal amplitude in males as well as decreased habituation from the mid-dose in males and in males and females combined are considered biologically relevant in the absence of appropriate statistical analysis (testing for interactions of sex, trial blocks and treatment) and positive controls. → Based on the available, there is some evidence that Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene induces functional deficiency. However, the deficiencies of the test limit the reliability on the effects observed in auditory startle response. Such parameters were not investigated for its analogue.	The DS claims that there is some evidence that the substance induces functional neurological deficiency based on changes in the ASR investigated as part of the OECD TG 443 DNT module. In the analysis of this data package, the DS has made several grave mistakes which distort the data reported in the CLH dossier and do not reflect the data available in the study report. Adequate scientific data analysis led to the conclusion that no decrease in mean max. amplitudes or habituation response was observed between control and high dose animals. Further information can be found in Annex 6 – Detailed information on proposed neurological functional deficits. Based on these data as well as on the lack of findings in the functional observational battery, it can be concluded that the exposure of animals with the test substance during developmental stages does not lead to a functional impairment.
64-75	Assessment of the reliability of read-across	The read-across to Reaction products of diphenylamine with nonene, branched, is considered not reliable for the endpoints assessed due to lack of supporting data and differences in data quality and quantity between both substances. Further information on why the read-across is not supported can be found in the comments submitted on the CLH proposal for Reaction products of diphenylamine with nonene, branched

page	Citation from CLH report, proposal for Harmonized Classification and Labelling for Benzenamine, N-phenyl, reaction products with 2,4,4-trimethylpentene	Comments on the respective section of the CLH proposal
80	Experimental BCF results are not available for the registered UVCB substance. QSAR estimations provided by the registrant to estimate BCF values of relevant constituents using the BCFBAF EPISuite and the BCF base-line model of Catalogic are presented in the table below.	Bioaccumulation potential has been assessed by the Registrants using experimental data, with support of QSAR modelling. Consideration of QSAR estimations and experimental data, bioaccumulation for several constituents of the UVCB substance is expected. However, all data on the main components as well as a related substance display BCF values below 2000. The relevant data for the assessment of bioaccumulation in fish, which are considered conservative , represent the data from the bioaccumulation study. The DS has evaluated the available bioaccumulation study in fish for the constituent Mono-Nonyl diphenylamine (EC 248-295-7) and reached the conclusion that the BCFKin in fish is 2219 L/kg, newly deviated, is more relevant than the BCFSS of 1730 L/kg. The registrants disagree with this evaluation and have found several inconsistencies if the study performed in 2000 would be evaluated according to the criteria after the revision of the OECD TG 305 in 2012. In addition, the registrants consider that the shortcomings in the performed fish study contribute to an unrealistic worst-case evaluation which cannot be used to robustly assess bioaccumulation. Since data are close to the threshold of 2000, a valid evaluation of the substance either to be bioaccumulative (“B”) or not is crucial with regard to the PBT assessment. Therefore, the registrants strongly recommend to re-perform the bioaccumulation study in fish.
81	Considering the measured log Kow≥4 for C4C4DPA constituent and the estimated log Kow≥4 for the main constituents of the UVCB substance, it is therefore concluded that the substance has a potential for bioaccumulation in aquatic species. The BCF model predictions support this conclusion.	Some constituents of the UVCB are considered to be bioaccumulative, but no constituent is expected to fulfil the PBT criteria for B nor vB.
85	According to CLP guidance, although measured concentrations are preferred, classification may be based on studies where nominal concentrations are the only valid data available. Although a calculated EL10 is available, the DS used the NOELR in the application of the classification criteria.	The data provide a valid EC/EL10 value which should be considered for classification in preference to the NOELR chosen by the DS. Daphnids were exposed to nominal concentrations based on Water Accommodated Fractions (WAFs). In the lowest concentration of 0.625 mg/L loading rate the reproduction decreased by 4 % compared to the control, a value which was statistically significant. Therefore, no NOELR could be derived. However, the data were sufficient to derive an EL10-value for a loading rate of 1.68 mg/L. According to the Guidance on the Application of the CLP Criteria (Jan 2024), preference should be given to the use of the EC10 over the NOEC for the classification of substances, similarly in OECD 2006: Series on Testing and Assessment Number 54, Current approaches in the statistical analysis of ecotoxicity data: a guidance to application.

page	Citation from CLH report, proposal for Harmonized Classification and Labelling for Benzenamine, N-phenyl, reaction products with 2,4,4-trimethylpentene	Comments on the respective section of the CLH proposal
		Biological tests aim not only to derive a level of statistical significance but also an assessment of biological relevance. This is consistent with the regulatory guidance and scientific understanding which is why EC10 is preferred over NOEC (European Chemicals Agency., 2024). While the EC10 is based on actual measured values, the NOEC is an artificial and random value in the test design chosen by the testing laboratory. Due to the low statistical spreading (CV: 9.5%) in reproduction of the control daphnia,the first two following concentrations of 0.625 and 1.25 mg/l (nominal) are already significant. Both the standard deviation and the coefficient of variance show a strong overlap of values. If values are displayed in a diagram, there are practically no differences between the control points and the points of the corresponding two subsequent concentrations, with strong overlaps in the standard deviations.

Annex 2 – Detailed information on effects on implantation sites and systemic toxicity

In the current CLH proposal, the DS claims that the substance Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene impacts the mean number of implantation sites and consequently litter sizes:

> ***"Lower mean number of implantation sites and consequently smaller litter sizes (main critical effect).*** *The litter size (mean number of delivered pups) was significantly reduced by 34% in high-dose dose (3000 ppm eq. 271 mg/kg bw/d) dams of the OECD TG 421 study, by 20% and 18% in high-dose dose (1800 ppm eq. 166 mg/kg bw/d) P0 and P1 dams respectively and by 10% in middose dose (600 ppm eq. 54 mg/kg bw/d) P1 dams of the EOGRTS. The mean number of implantation sites in high-dose (225 mg/kg bw/d) dams of the OECD TG 422 study was decreased by 16 % without reaching statistically significance.*
> *The litter size (mean number of delivered pups) was significantly reduced by 19% and 31% in the mid-dose (1500 ppm eq. 133 mg/kg bw/d) and high-dose dose (5000 ppm eq. 443 mg/kg bw/d) dams exposed to the analogue Reaction products of diphenylamine with nonene, branched.*
> *This effect was dose-related and observed from dose levels (mid-dose groups) where no effects on body-weight and food consumption were observed during the premating periods."*

This assessment is not adequately reflecting the data available and relevant for the decision on classification of the substance. The following information reviews the available data/studies and provides more details to support the findings compiled in the summary.

1. OECD TG 422

In an OECD TG 422 study (WIL Research Europe, 2014), Wistar rats were treated with the test substance at doses of 0, 25, 75 or 225 mg/kg bw/d via gavage in corn oil for two weeks prior to mating, throughout mating, gestation and lactation until post-natal day (PND) 4. Systemic toxicity with changes in organ weights, clinical pathology and histopathology were observed for the liver at the mid and high dose levels.

Effects on toxicity to reproduction

No effects on fertility parameters were observed up to the highest dose tested. The high dose group had a lower litter size (mean pup number) compared to controls (Table 1). This was attributed to an increase in postnatal loss and a correspondingly lower viability index, mainly attributable to one female. For further information on this part, please refer to Annex 7 – Discussion of proposed information death of the developing organism. While the DS claims that there was a decrease of 16% in implantation sites in the high dose group, this was not statistically significant and not considered relevant by the study director having considered all the data. Therefore, this study should not be further considered for an assessment of substance-specific effects on implantation sites.

Based on the data available, the study director concluded that no reproductive toxicity was observed up to the highest dose tested. For the postnatal loss observed in the high dose group, a correlation with treatment could not be excluded. Therefore, the NOAEL for fertility in this study was 225 mg/kg bw/d and the developmental NOAEL was 75 mg/kg bw/d.

Table 1. Implantation sites and litter size observed in OECD TG 422 study (WIL Research Europe, 2014)

Dose level	[mg/kg bw/d]	0	25	75	225
Pregnant	N	10	10	10	9
Implantation sites	mean	10.9	11.5	10.5	9.2
	SD	2.4	2.2	1.7	2.3
Litter size	mean	10.9	10.5	9.4	7.9*
	SD	2.6	2.5	2.0	1.8

Steel-test, *p ≤ 0.05

Historical control data from this laboratory were available (in the time frame of the commenting period) only for the time period 2015-2017 for OECD TG 421 and 422 studies (Table 2). Based on this information, the implantation sites found in the high dose group (total of 83 implantation sites in this group) is within the historical control data of the laboratory and thus should be regarded as biologically not relevant.

Table 2: Historical control data from 48 OECD 421 and 422 studies for the years 2015-2017.

	Implantation sites per dam	Total number of implantation sites per group	Total number of offspring born per group
Mean	12.6	114	106
SD	2.7	15	15
Min.	1	80	73
Max.	18	142	136

Target organ toxicity: liver and thyroid

The liver was identified as target organ for systemic effects in this study. Males were more strongly affected than females, however since only females are relevant for the discussion on implantation sites, only the data related to effects on livers of parental females are shown in Table 3. Liver weight increases were observed starting at the mid dose level (75 mg/kg bw/d), however these were not statistically significant. Based on the organ weight changes together with findings in clinical pathology, and concurrent incidences of histopathological findings, the effects observed at this dose level were considered treatment-related and toxicologically relevant. At the high dose level, liver weight increases in females had not reached statistical significance either, however clinical pathology findings were stronger and in combination with histopathological findings of minimal to slight hepatocellular hypertrophy and up to moderate hepatocyte vacuolation, the effects on the liver at this dose level were considered treatment-related and adverse.

Table 3: effects observed on liver parameters in dams in OECD TG 422 study (WIL Research Europe, 2014)

Dose [mg/ kg bw/d]	0	25	75	225
liver weights [% of control, abs/rel]	100/100	104†/100†	117†/117†	105†/111†
Clinical chemistry	---	---	ALP ↑ Bilirubin ↑ Total protein ↓ Albumin ↓	ALP ↑ Bilirubin ↑ Cholesterol ↑ Total protein ↓ Phosphate ↓ Albumin ↓
Histopathology	---	---	Hepatocellular hypertrophy • Grade 1 (1/5F)	Hepatocellular hypertrophy • Grade 1 (2/6F) • Grade 2 (4/6F)
	---	---	Hepatocyte vacuolation • Grade 1 (2/5F)	Hepatocyte vacuolation • Grade 1 (3/6F) • Grade 2 (1/6F) • Grade 3 (2/6F)

--- No treatment-related findings
* Effects are considered treatment-related, but not adverse
† Effects were not statistically significant
ALP Alkaline phosphatase

Thyroid organ weights in females were unchanged throughout the course of the study (Table 4). Decreased levels of thyroxine (T4) were observed in all treated dose levels, however based on the lack of a dose-response and in the absence of concurrent histopathological changes, the toxicological relevance of these findings was considered questionable by the study director. Thyroid stimulating hormone (TSH) was increased in all treated females, however the changes did reach statistical significance only in the low and mid dose levels. In the absence of adverse findings observed in the thyroids during microscopic examination, the study director considered the increased TSH values as not toxicologically relevant.

Overall, the liver was a clear target organ with adaptive but toxicologically relevant changes in the mid dose group and adverse findings in the high dose group. Thyroid hormone changes occurred in dams of all dose levels without histopathological correlates and were therefore of unclear toxicological relevance. The parental NOAEL for systemic toxicity was 25 mg/kg bw/d.

Table 4: effects observed on thyroid parameters in dams in OECD TG 422 study with Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene (WIL Research Europe, 2014)

Dose [mg/kg bw/d]	0	25	75	225
thyroid weights [% of control, abs/rel]	---	100†/100†	94†/100†	88†/100†
thyroid hormones	---	total T4 ↓ TSH ↑*	total T4 ↓ TSH ↑*	total T4 ↓ TSH ↑†*
Histopathology	---	---	---	Follicular cell hypertrophy: • Grade 1 (1/5F)

--- No treatment-related findings
* Effects are considered treatment-related, but not adverse
† Effects were not statistically significant

2. OECD TG 421

An OECD TG 421 study with extended study design was conducted as a dose-range finder for the following OECD TG 443 study (BASF SE, 2020a). Wistar rats (10 animals per sex and group) were treated with doses of 0, 300, 1000 or 3000 ppm Benzenamine, N-phenyl-, reaction products with 2,4,4- trimethylpentene in diet for ten weeks prior to mating and throughout mating, females were further treated during gestation and 13 days of lactation. Systemic toxicity was observed in decreased food consumption and body weights in the mid and high dose groups. Target organ toxicity was observed for the liver and thyroid with increased organ weights, histopathological changes and clinical pathology changes observed in all dose groups to a varying degree. Further information on systemic toxicity is provided in the section below.

Effects on toxicity to reproduction

Mating, fertility, gestation, and live birth indices as well as post-implantation loss were without treatment-related changes in all dose groups. Increased estrus cycle length was observed in females of the high dose groups, which is further discussed in Annex 4 – Detailed information on data for proposed estrus cycle changes. In the high dose group, a reduced mean number of implantation sites (8.8 vs. 13.8 in control) and a subsequently reduced litter size was reported (Table 5). This finding was statistically significant and outside of the historical control data of the laboratory (updated historical control data below). The DS mentions that the mid dose group already showed lower implantation sites, however it should be clearly pointed out that these values are neither statistically significant nor biologically relevant as they are within the range of biological variation usually observed for this rat strain. The decrease in implantation sites observed occurred only in the context of systemic toxicity (reduced food consumption, reduced body weights, pronounced target organ toxicity in liver and thyroid). It was previously shown that these effects on dams can contribute to toxicity to reproduction, and based on the data available, a mode of action is proposed with a primary action of the substance on the target organ liver, followed by secondary effects on thyroid and thyroid hormones which then in turn result in perturbations of fertility. A detailed description of the postulated mode of action as well as a summary of relevant literature are provided in section 4 of this Annex (Mode of action for effects on fertility).

Table 5: Litter numbers, implantation sites and litter size in OECD TG 421 study (BASF SE, 2020a)

Concentration	[ppm]	0	300	1000	3000
Approx. dose	[mg/kg bw/d]		28	95	271
Pregnant dams	N	9	6	8	9
Number of litters	N	9	6	8	9
Implantation sites	N	124	84	95	79
	Mean	13.8	14.0	11.9	8.8**
	SD	2.0	1.1	1.4	1.6
Pups delivered	N	110	84	92	72
	Mean	12.2	14.0	11.5	8.0**
	SD	3.2	1.1	1.6	2.0

Wilcoxon with Bonferroni-Holm (one-sided-), ** $p \leq 0.01$

Based on the data available, the NOAEL for fertility was set at 1000 ppm (95 mg/kg bw/d).

Historical control data

One key argument of the DS regarding reliability of the available data was that the historical control data (HCD) provided in the study reports was unreliable:

"However, the range of HCD for the number of implantation sites from 2008 to 2018 period provided in the study report of the OECD TG 443 (Unpublished study report, 2021) performed with Benzenamine, N-phenyl-, reaction products with 2,4,4- 30rimethylpentane by the same laboratory, was 11.1 – 15.3 sites and 11.2-15.3 when considering a more appropriate timeframe (2015-2018). When considering the later HCD, the value of the concurrent control is well within the HCD range while the values of the mid-dose and high-dose groups are outside […]"

It should be pointed out that historical control data from one study protocol can only be adapted to another study protocol under the considerations of e.g., animal age, treatment protocol, group size, etc., and expert judgement on the study conduct is required to assess which data are adequate for consideration as historical control data. The expert judgement of the laboratory that conducted the studies was that HCD from OECD TG 443 studies are not appropriate for assessment of data obtained from OECD TG 421 studies. Given that study reports are finalized as soon as possible, only data from past years can be considered as historical control data at the time of report generation. The study in-life phase was in 2019, therefore new historical control data from OECD TG 421/422 studies were compiled which were conducted between 2015 and 2022. These include 79 studies, with the following exposure routes: diet (11), gavage (58), drinking water (6), inhalation (4). The results are presented in Table 6.

Table 6: Historical control data from 79 OECD TG 421/422 studies performed at the testing laboratory, compiled on March 6, 2024

	Runtime	Litters	Implantation sites [mean]	Pups delivered [mean]	Post-implantation loss [mean%]
Mean of means			12.7	12.0	6.4
S.D.			0.90	0.83	3.90
Min.	2015	8	9.8	9.0	0.9
Max.	2022	10	14.7	14.0	16.8

S.D., standard deviation

Target organ toxicity: liver and thyroid

The liver was identified as target organ for systemic effects in this study. Since only females are relevant for the discussion on implantation sites, only the data related to effects on livers of parental females are shown in Table 7. Liver weight increases were observed starting at the mid dose level (1000 ppm, approx. 95 mg/kg bw/d in females), however only relative liver weights were statistically significantly increased at the mid and high dose level. Alkaline phosphatase (ALP) levels were significantly increased in a dose-related fashion in all treated groups. Similarly, triglycerides were increased, and total bile acid decreased in a dose-related incidence and severity. At the high dose level, γ-glutamyltransferase (GGT) was also significantly increased. This indication of liver damage was corroborated by histopathological findings of minimal hepatocellular hypertrophy at the mid dose level, which increased in incidence and severity up to moderate in the high dose group and was accompanied by periportal liver fatty change in several animals.

Table 7: (adverse) effects observed on liver parameters in dams in OECD TG 421 study (BASF SE, 2020a)

Concentration [ppm]	0	300	1000	3000
Dose [mg/ kg bw/d]	0	28	95	271
liver weights [% dev. from control, abs/rel]	---	♀ 99†/99†	♀ 110†/116	♀ 114†/130
Clinical chemistry	---	ALP ↑ Triglycerides ↑†	ALP ↑ Total bile acid ↓ Triglycerides ↑	ALP ↑ GGT ↑ Total bile acid ↓ Triglycerides ↑ Cholesterol ↑ Albumin ↓ total bilirubin ↑†
Histo-pathology	---	---	Hepatocyte hypertrophy, centrilobular • Grade 1 (8/10F)	Hepatocyte hypertrophy, diffuse, centrilobular accenduated • Grade 1 (3/10F) • Grade 2 (5/10F) • Grade 3 (2/10F)
	---	---	---	Liver fatty change, periportal • Grade 1 (3/10F) • Grade 2 (2/10F)

--- No treatment-related findings
* Effects are considered treatment-related, but not adverse
† Effects were not statistically significant

Relative thyroid weights were statistically significantly increased at the mid dose level; the increased organ weights at the high dose level were not statistically significant (Table 8). Histopathological analysis showed follicular cell hypertrophy with increased incidences as compared to controls and historical controls from the low dose level. In the high dose group, the hypertrophy/hyperplasia of follicular cells was accompanied by altered colloid with flaky appearance. These changes were regarded as related to treatment with the test substance.

Table 8: effects observed on thyroid parameters in dams in OECD TG 421 study (BASF SE, 2020a)

Concentration [ppm]	0	300	1000	3000
Dose [mg/ kg bw/d]	0	28	95	271
thyroid weights [% dev. from control, abs/rel]	---	99†/103†	113†/124	95†/111†
thyroid hormones	---	---	---	---
Histopathology	---	Follicular cell hypertrophy*: • Grade 1 (2/10F) • Grade 2 (1/10F)	Follicular cell hypertrophy*: • Grade 1 (2/10F)	Follicular cell hypertrophy: • Grade 1 (5/10F) • Grade 2 (1/10F)
	---	Thyroid altered colloid*: • Grade 1 (3/10F)	Thyroid altered colloid*: • Grade 1 (2/10F)	Thyroid altered colloid: • Grade 1 (6/10F)

--- No treatment-related findings
* Effects are considered treatment-related, but not adverse
† Effects were not statistically significant

Based on the data available, the NOAEL for systemic toxicity in females was 300 ppm (28 mg/kg bw/d).

3. OECD TG 443

An extended one-generation reproduction toxicity study was performed according to OECD TG 443 and under GLP (BASF SE, 2021). Wistar rats (25 per sex and dose) were treated with doses of 0, 200, 600 or 1800 ppm Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene via diet. F0 animals were treated for at least 10 weeks prior to mating to produce a litter (F1 generation). Pups of the F1 litter were selected (F1 rearing animals) and assigned to 4 different cohorts (1A, 1B, 2A and 2B) which were subjected to specific postweaning examinations. Cohort 1B (F1 generation parental animals) was selected to produce F2 pups. Cohort 1A animals were analysed for specific reproductive parameters and Cohort 2A and 2B animals were used for assessment of developmental neurotoxicity (see Annex 5 – Detailed information on neurohistopathology and neuromorphometrics and Annex 6 – Detailed information on proposed neurological functional deficits).

The dose levels were selected based on the effects observed in the previous OECD TG 421 study conducted with an extended pre-treatment period (see section 2 of this Annex). Considering the effects observed on the liver at the high dose level and the intention to continue the study through a second-generation pairing, the top dose level of 1800 ppm was chosen with the aim to induce some toxicity but avoid severe suffering of the animals (as is prescribed in the OECD Testing Guideline 443). Systemic toxicity was observed with decreased food consumption, decreased body weights and body weight gains in all dose groups, with dose-dependent differences regarding severity.

Effects on fertility

A slight but statistically significant reduction in numbers of implantation sites were observed in F0 and F1 parental animals of the high dose group (Table 9). These values were outside of the historical control data of the laboratory (updated range: 11.14 – 13.7).

Concurrently, the mean number of pups per dam in the high dose group was also found to be statistically significantly reduced. In the F0 generation of the OECD TG 443 study, no effects on litter size and implantation sites were observed in the mid dose group. Similarly, no statistically significant effect on implantation sites was observed in the mid dose group of the F1 generation of the OECD TG 443 study, while the litter size was slightly but statistically significantly reduced (Table 9). This finding was considered not biologically relevant as it was well within the historical control data of the laboratory (updated HCD 10.7-12.8).

Table 9: Litter numbers, implantation sites and litter size from OECD TG 443 study (BASF SE, 2021)

Concentration	[ppm]	0	200	600	1800
Approx. dose	[mg/kg bw/d]		18	54	166
F0 parental animals					
Pregnant females	N	24	24	25	25
Litters	N	24	23	25	25
Implantation sites	N	302	276	298	267
	Mean	12.1	11.0	11.9	10.7**
	SD	3.0	3.4	2.1	1.5
Pups delivered	N	288	261	288	239
	Mean	12.0	11.3	11.5	9.6**
	SD	1.9	1.3	2.1	2.2
F1 parental animals (Cohort 1B)					
Pregnant females	N	23	24	24	25
Litters	N	23	24	24	25
Implantation sites	N	283	282	270	255
	Mean	12.3	11.8	11.2	10.2**
	SD	2.3	2.2	1.9	1.9
Pups delivered	N	277	270	260	244
	Mean	12.0	11.2	10.8*	9.8**
	SD	2.2	2.3	1.8	1.9

Wilcoxon with Bonferroni-Holm (one-sided-), ** $p \leq 0.01$

On a single animal basis, the control and high dose groups show a significant overlap for the parameter implantation sites (Figure 1) for both generations, demonstrating the magnitude of the effect is actually much smaller than the mean values would suggest.

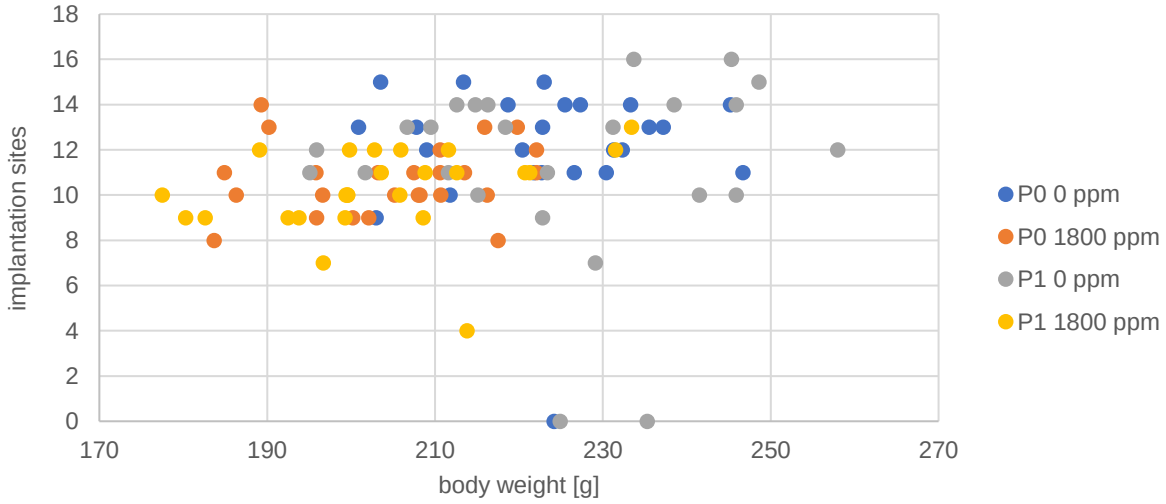


Figure 1: graphic depiction of single animal data on body weights and implantation sites per dam in F0 (P0) and F1 Cohort 1B (P1) groups of OECD TG 443 (BASF SE, 2021)

Because the effect on implantation sites is subtle, it is difficult to determine which animals are affected and thus correlation of this borderline effect on a single animal basis with coarse parameters such as body weight or food consumption is challenging. The large overlap between groups and the uncertainty which animals are affected to a biologically relevant extent introduces a large degree of uncertainty to the interpretation. While the effect is statistically significant at the top dose, the overall magnitude and significance across the animals in the groups is not definitive.

The DS suggests that a dose-response exists with regards to the implantation sites. Based on the OECD TG 421 and OECD TG 443 studies available, no clear dose-response could be derived. Table 5 and Table 9 show that the high dose groups are affected for each of the studies and generations, however no dose-response can be observed e.g., for the F0 generation in the OECD TG 443. Clearly, working in a biological system some variation of biological parameters must be anticipated, even if during study conduct the utmost care is taken to ensure animals are treated similarly. Therefore, scientific data analysis relies on the trifecta of statistical significance, biological relevance, and dose-response to indicate relevant effects in biological systems. While the findings in the high dose groups fulfil the requirements "statistical significance" and "biological relevance", none of the other dose groups were found to present with these criteria fulfilled. Therefore, a dose-response including other dose levels than the high dose group as described by the DS is not supported by scientific analysis of the data available.

Based on the data available, the NOAEL for fertility and reproductive performance was set at 600 ppm (approx. 54 mg/kg bw/d) for F0 and F1 parental rats.

Historical control data

One key argument of the DS regarding reliability of the available data was that the historical control data (HCD) provided in the study reports was unreliable:

"HCD have been provided, however several shortcomings limit their reliability. The collection period generally exceeds the recommended 5-years encompassing the year of the study. The protocol of the studies is not always clearly indicated (e.g. whether OECD TG 416 or TG 443 where followed, the route of administration not indicated for all studies). The studies included in HCD are different according to the parameters considered, which further limits the transparency and readability of those data. In view of the limitations of the provided HCD, they were not be given much weight compared to the concurrent control group, which anyway represent the most relevant comparator for determining treatment-related effects if the concurrent control is not aberrant."

As described above, a biological system is known to vary to a certain extent. Therefore, historical control data are considered important for assessment of biological relevance of an observed effect. Given that study reports are finalized as soon as possible after finishing of the analysis phase, only data from past years can be considered as historical control data at the time of report generation. The study in-life phase was in 2019, therefore new historical control data from similar studies were compiled which were conducted between 2017 and 2022. These include 16 studies, with the following exposure routes: diet (8), gavage (7), drinking water (1). The results are presented in Table 10.

Table 10: Historical control data from 16 OECD TG 443/416 studies performed at the testing laboratory, compiled on March 27, 2024

	Runtime	Litters	Implantation sites [mean]	Pups delivered [mean]
Mean of means		24	12.8	12.0
S.D.			0.59	0.60
Min.	2018	21	11.4	10.7
Max.	2022	25	13.7	12.8

Target organ toxicity: liver and thyroid

The DS states that at the top dose level in the study, the systemic toxicity was not marked, however they define marked toxicity as *"no lethality, no dramatic reduction in bodyweight and no*

coma". It should be clearly pointed out that while those might be the criteria to justify non-classification based on systemic toxicity, these are also the criteria to justify humane sacrifice during study conduct according to OECD Guidance document 19 (OECD, 2000). Therefore, it should not be the goal to achieve this level of toxicity during an OECD guideline study, given that the OECD testing guidelines specifically state that the high dose level should be chosen to induce some toxicity, but not severe suffering or death (OECD, 2018). Therefore, the assessment of these effects should not be regarded as "marked" toxicity but would rather be considered "excessive" toxicity in the context of an OECD Guideline study.

Table 11: effects observed on liver parameters in F0 dams in OECD TG 443 study (BASF SE, 2021)

Concentration [ppm]	0	200	600	1800
Dose [mg/ kg bw/d]	0	18	54	166
F0 parental animals				
liver weights [% of control, abs/rel]		100†/102†	107/108	123/136
Clinical chemistry		ALP ↑*	ALP ↑ Triglycerides ↑	ALP ↑ Triglycerides ↑ Albumin ↓ Globulin ↑ GGT ↑ Cholesterol ↑
Histopathology		Hepatocellular hypertrophy, centrilobular* • Grade 1 (3/20F)	Hepatocellular hypertrophy, centrilobular • Grade 1 (4/20F) • Grade 2 (1/20F)	Hepatocellular hypertrophy, centrilobular • Grade 1 (3/20F) • Grade 2 (5/20F) • Grade 3 (4/20F) Hepatocellular hypertrophy, diffuse • Grade 2 (5/20F) • Grade 3 (1/20F)
F1 Cohort 1A [rearing animals]				
liver weights [% of control, abs/rel]		100†/99†	105†/106	121/131
Clinical chemistry		ALP ↑* Albumin ↓*	ALP ↑ Albumin ↓	ALP ↑ Triglycerides ↑ Albumin ↓ Globulin ↑ GGT ↑ Cholesterol ↑ Total bile acid ↓
Histopathology		---	Hepatocellular hypertrophy, centrilobular Grade 1 (4/20F)	Hepatocellular hypertrophy, centrilobular • Grade 1 (9/20F) • Grade 2 (6/20F) • Grade 3 (3/20F)
F1 Cohort 1B [F1 parental animals]				
liver weights [% dev. from control, abs/rel]	---	101†/100†	107†/107†	119/132

--- No treatment-related findings
* Effects are considered treatment-related, but not adverse
† Effects were not statistically significant

The liver was identified as target organ for systemic effects in this study as well. Since only females are relevant for the discussion on implantation sites, only the data related to effects on livers of parental females are shown in Table 11. Increased liver weights were observed from the mid dose level (600 ppm, approx. 54 mg/kg bw/d) in F0 parental animals and F1 Cohort 1A animals. In combination with changes in clinical chemistry and hepatocellular

hypertrophy observed in histopathological investigations, these findings were considered treatment-related and adverse. In Cohort 1B parental animals (F1), liver weights were increased at the highest dose level tested (1800 ppm, approx. 166 mg/kg bw/d). No further evaluations for target organ toxicity were performed in these animals, however the liver weight changes were considered treatment-related and adverse. Based on comparison with the data generated in F0 and F1 Cohort 1A animals, the effects on liver seem to be consistent over all groups.

In the thyroid glands of F0 females of all test groups hypertrophy/hyperplasia of follicular cells and altered colloid were observed. This corresponded to increased organ weights in the high dose group (Table 12). In combination with the changes in hormone levels observed in clinical pathology, these findings were regarded to be treatment-related and adverse in all test groups. In F1 Cohort 1A females, follicular cell hypertrophy/hyperplasia was observed from the mid dose group and altered colloid was observed in the high dose group. In combination with clinical pathology findings in these dose groups, these findings were considered treatment-related and adverse.

Table 12: effects observed on thyroid parameters in female animals in OECD TG 443 study (BASF SE, 2021)

Concentration [ppm]	0	200	600	1800
Dose [mg/ kg bw/d]	0	18	54	166
F0 parental animals				
Thyroid weights [% of control, abs/rel]		96†/110†	111†/112†	108†/121
Thyroid hormones		TSH ↑	TSH ↑	TSH ↑
Histopathology		Follicular cell hypertrophy: • Grade 1 (2/20F)	Follicular cell hypertrophy: • Grade 1 (7/20F)	Follicular cell hypertrophy: • Grade 1 (12/20F) • Grade 2 (4/20F)
	Thyroid altered colloid: • Grade 1 (3/20F)	Thyroid altered colloid: • Grade 1 (7/20F)	Thyroid altered colloid: • Grade 1 (8/20F) • Grade 2 (2/20F)	Thyroid altered colloid: • Grade 1 (1/20F) • Grade 2 (12/20F) • Grade 3 (6/20F)
F1 Cohort 1A [rearing animals]				
Thyroid weights [% of control, abs/rel]		101†/100†	113/114†	113/122
Thyroid hormones			total T4 ↓, TSH ↑†	TSH ↑
Histopathology			Follicular cell hypertrophy: • Grade 1 (3/20F)	Follicular cell hypertrophy: • Grade 1 (6/20F) • Grade 2 (2/20F)
				Thyroid altered colloid: • Grade 1 (2/20F) • Grade 3 (1/20F)

--- No treatment-related findings
* Effects are considered treatment-related, but not adverse
† Effects were not statistically significant

Based on the data available, the NOAEL for systemic toxicity was set at < 200 ppm (< 18 mg/kg bw/d) in female rats.

4. Mode of action for effects on fertility

Possible causes for a reduction in implantation sites specific to reproductive toxicity include changes in estrus cycle, follicle counts, sperm parameters, mating behaviour and sexual organ functionality. No effects were observed on sperm parameters and mating behaviour in the available studies. A discussion on estrus cycle changes can be found in Annex 4 – Detailed information on data for proposed estrus cycle changes. Information on differential ovarian follicle count as measure for sexual organ functionality is included in Annex 3 – Detailed information on ovary parameters. No relevant changes could be found in any of the parameters listed above, which greatly reduces the likelihood of a specific mechanism.

Other factors which can impact implantation include maternal stress, either through systemic toxicity or other mechanisms. In the repeated dose-type studies conducted with Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene in rats, the observed systemic toxicity shows a consistent pattern with the liver being the primary target organ. Moreover, changes in thyroid organ weight and histopathology (hypertrophy) as well as thyroid stimulating hormone (TSH) and thyroxine (T4) were seen. The pattern of liver and thyroid effects is fully in line with those expected for changes in thyroid parameters commonly seen in rats in response to chemically induced hepatic metabolism.

This mode of action is well described in several peer-reviewed publications (McClain, 1989a; Meek et al., 2003; Zoeller et al., 2007). Due to increased metabolic capacity of the liver subsequent to xenobiotic / test substance exposure, conjugation of T4 with mainly glucuronic acid is also increased as a bystander effect. This in turn leads to lower levels of predominantly circulating T4. Since thyroid hormones are part of a feedback loop in the thyroid – pituitary axis, this drop in T4 (and most likely also T3) leads to increased production and release of TSH in order to compensate for the increased loss of T3/T4. Responding to increased TSH levels, the thyroid gland then increases production capacity for the thyroid hormones, which is achieved by cellular hypertrophy to accommodate for the additional requirement in capacity. Several pharmaceuticals and xenobiotics have been shown to act via this pathway, the most prominent being phenobarbital. The hallmarks of this mode of action are consistent with the findings for both liver and thyroid parameters of the registered substance: Hepatocellular hypertrophy, increased liver weight, decreased T4 levels, increased TSH levels and thyroid follicular hypertrophy accompanied by increased thyroid weight. Therefore, the thyroid changes are considered secondary to increased hepatocellular metabolic capacity.

Beside this general toxicity, subtle changes in specific female fertility parameters, i.e., reduced number of implantation sites leading to lower litter sizes, as well as inconsistent effects on estrous cycle and ovary weights occurred in the studies available with Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene.

It is well-known that an imbalance of thyroid hormones affects female fertility, and a balanced level of thyroid hormones is crucial for successful pregnancy. Thyroid hormones are known to regulate the secretion of the most relevant reproductive hormones, i.e., estradiol and progesterone, which play a crucial role in maturation of oocytes and are essential for embryo implantation. The molecular basis of thyroid hormone action during implantation and early pregnancy has become clearer in the last 10-20 years and is summarized in Colicchia et al. and Silva et al., describing the relevance of thyroid hormone function before, during and past the implantation window (Colicchia et al., 2014; Silva et al., 2018).

In rats, the impact of hypothyroidism on female fertility can be studied by experimental induction of hypothyroidism by treatment with 6-propyl-2-thiouracil (PTU), with induced hypothyroidism confirmed by significantly increased levels of TSH and decreased levels of T3

and T4. As expected, the induced hypothyroidism was found to affect the levels of circulating reproductive hormones such as estradiol, progesterone and prolactin (Hapon et al., 2010). Hypothyroidism did not influence ovulation rate or the number of corpora lutea at term, but a diminished number of implantation sites and consequently pups per litter were observed (Hapon et al., 2010; Rinaldini et al., 2021). Treatment of hypothyroid rats with T3 was found to re-establish the number of implanted embryos to normal, further substantiating the importance of thyroid hormones for successful implantation (Rinaldini et al., 2021). For successful implantation of the embryo in the uterine epithelium, uterine receptivity is essential. In hypothyroid rats, decreased expression of uterine-receptivity factors (homeobox A10 and osteopontin) as well as changes in prostaglandin signaling were observed (Kowalczyk-Zieba et al., 2021).

The postulated mode of action leading to the observed reduced number of implantation sites and litter sizes comprises the following key events:

Table 13. Listing of key events identified for the suspected mode of action.

Key Event 1	Induction of hepatic enzymes
Key Event 2	Conjugating circulating thyroid hormones leading to lower levels of bioavailable T3 & T4, triggering an increased compensatory production of TSH resulting in cellular hypertrophy in the thyroid
Key Event 3	Change in uterine receptivity crucial for early implantation, no impact on the number of corpora lutea or fetal survival
Adverse effect	Lower number of implantation sites and consequently pups per litter

Overall concordance of the related findings regarding key events 1-3 and the associated adverse outcome in the available toxicological studies with the substance is assessed in the following.

Table 14. Concordance of dose-response relationships together with temporal association of effects observed in Wistar rats.

Dose [mg/kg bw/d]	***Key Event 1 (time observed and severity)***	***Key Event 2 (time observed and severity)***	***Key Event 3 (time observed and severity)***	***Adverse effect (time observed and severity)***
	Liver changes indicating liver enzyme induction	***Changes in TSH and/or thyroid hormones / thyroid hypertrophy***	***Uterine receptivity***	***Lower number of implantation sites***
18	+ (17 weeks, F0) - (13 weeks, F1)	+ (17 weeks, F0) - (13 weeks, F1)		- (17 weeks, F0) - (13 weeks, F1)
25	- (53 days)	- (53 days)		- (53 days)
28	+ (16 weeks)	+ (16 weeks)		- (16 weeks)
54	+ (17 weeks, F0) + (13 weeks, F1)	+ (17 weeks, F0) + (13 weeks, F1)		- (17 weeks, F0) - (13 weeks, F1)
75	+ (53 days)	- (53 days)		- (53 days)
95	+ (16 weeks)	+ (16 weeks)		- (16 weeks)
166	++ (17 weeks, F0) ++ (13 weeks, F1)	++ (17 weeks, F0) ++ (13 weeks, F1)		+ (17 weeks, F0) + (13 weeks, F1)
225	++ (53 days)	- (53 days)		- (53 days)
271	++ (16 weeks)	++ (16 weeks)		+ (16 weeks)

- indicates no effect; +, ++ and +++ indicates the effect size (severity)

As described in detail in the paragraphs above and summarized in Table 14, slight, non-adverse liver changes indicating adaptive liver enzyme induction were observed starting from 75 mg/kg/day level for short(er) exposure duration (i.e., 53 days) and already from 18 or 26 mg/kg/day under subchronic exposure conditions in parental animals. In the F1 generation liver changes are found only starting from 54 mg/kg/day following 13 weeks exposure. Stronger liver effects were seen starting at 166 mg/kg/day after subchronic exposure duration both in F0 and F1 animals. Appearance of thyroid findings match the occurrence of liver findings and were not observed in the absence of liver findings.

A biologically relevant reduction in the number of implantation sites was not observed after 53 days of exposure at any dose level up to 225 mg/kg/day; moreover significant reductions in implantation sites were observed after subchronic exposure duration only at dose levels starting from 166 mg/kg/day in F0 animals (with implantation sites of F1 animals being slightly below the HCD already at 54 mg/kg/day, but not statistically significant and still associated with liver and thyroid findings at this dose level). As can be seen in the table, the number of implantation sites was not reduced in the absence of liver and thyroid findings and appeared only at sufficiently strong liver and thyroid findings.

Unfortunately, reproductive hormones and other factors affecting key event 3 are not assessed in regulatory OECD TG studies, thus no data on this key event are available for the substance. However, it is plausible to consider that a change in these parameters can be expected within this context, as shown in the literature (see above).

In conclusion, the slightly reduced mean number of implantation sites and associated slight reduction in litter size can be considered secondary to the liver and thyroid effects as described above. Based on the observed pattern and concordance of effects throughout the available studies, the postulated mode of action is considered highly likely.

Human relevance of the mode of action

The observed thyroid changes are consistent in both type of change as well as occurrence with a secondary effect due to liver enzyme induction. This mode of action is well described in several peer-reviewed publications (McClain, 1989b; Meek and et al., 2003; Zoeller and et al., 2007). As described above, the thyroid changes are considered secondary to increased hepatocellular metabolic capacity.

The rat shows a particularly high sensitivity to chemicals which act via this mode of action, whereas humans are far less sensitive. The reason for this are widely accepted species differences in thyroid hormone regulation (McClain, 1989b; Meek and et al., 2003). These comprise e.g., the lack of the thyroxine-binding protein in the adult rat in combination with shorter half-lives of T4 and T3. An overview is provided in Table 15.

Experimental evidence underscores the increased susceptibility of the rat to disturbance of the thyroid homeostasis (Lewandowski et al., 2004; Takayama et al., 1986; US-EPA, 1998).

Altogether, the effects observed on thyroid hormone changes are considered secondary to liver enzyme induction and due to species differences, these findings are regarded as not relevant for humans.

Table 15: Comparison of the thyroid system in humans and rodents (adapted from Choksi et al., 2003(Choksi et al., 2003), and Lewandowski et al., 2004(Lewandowski et al., 2004))

Parameter	**Human**	**Rat**
Half life of T4	5-9 days	0.5-1 day
Half life of T3	1 day	0.25 days
High-affinity thyroxine binding globulin	Present	Absent
Primary serum-binding protein	Thyroxine-binding globulin	Albumin
T4 production (rate/kg bw)	1x	10x
Serum T3	147 ng/dL	25-100 ng/dL
Serum T4	7.2 µg/dL	3-7 µg/dL
Serum TSH	0.05-0.5 ng/mL	0.6-3.4 ng/mL
Sex difference in serum TSH levels	No difference	Adult males > adult females
Follicular cell morphology	Follicular height is equal in males and females	Follicular height in males is greater than in females

Annex 3 – Detailed information on ovary parameters

In the CLH proposal, the DS claims that the substance Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene causes changes in ovary weights:

> ***"Decreased ovary weights**. Absolute ovary weight was significantly reduced by 25% in high-dose dams exposed to Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene in the OECD TG 421 study and by 13% and 12% in P0 and P1 high-dose dams respectively in the EOGRTS.*
> *Dose-related and significant decrease in absolute ovarian weights was observed in dams exposed to the analogue Reaction products of diphenylamine with nonene, branched (by 18% and 40% at mid- and high dose level respectively) in the OECD TG 421 study. The relative weight was also impacted from the mid-dose level but statistically significantly only at the high dose level. While a statistically significant decrease (-14%) in absolute ovarian weight was also observed in the recovery group of this study (not mated females exposed to 5000 ppm during 10 weeks with a 2-week recovery period) no effect on ovary weight was noted in females up to 1000 mg/kg bw/d of the analogue by gavage in corn oil in a GLP-compliant 90-day toxicity study (Unpublished study report, 2013**- Study 7**).*
> *While this effect was consistently observed in functional reproductive studies, no histopathological correlate was reported in any of the studies."*

1. <u>OECD TG 421</u>

An OECD TG 421 study with extended study design was conducted as a dose-range finder for the following OECD TG 443 study (BASF SE, 2020a). Wistar rats (10 animals per sex and group) were treated with doses of 0, 300, 1000 or 3000 ppm Benzenamine, N-phenyl-, reaction products with 2,4,4- trimethylpentene in diet for ten weeks prior to mating and throughout mating, females were further treated during gestation and 13 days of lactation, at which time ovary weights were determined for parental females (Table 16).

Absolute ovary weights were significantly lower in high dose females, however also body weights of treated animals were affected. To put this parameter into perspective, relative ovary weights should also be analyzed as these take body weights into consideration. Importantly, relative ovary weights were not significantly changed and in the same range as the concurrent controls. Histopathological analysis of the ovaries did not identify any changes.

Table 16: Absolute and relative ovary weights observed in OECD TG 421 study (BASF SE, 2020a)

Concentration in diet [ppm]		0 ppm	300 ppm	1000 ppm	3000 ppm
Absolute ovary weights [mg]	Mean	107.0	110.7	104.7	80.3**
	Deviation (% of control)	100	103	98	75
	SD	21.94	16.62	14.41	14.80
	n	10	10	10	10
Relative ovary weights [%]	mean	0.045	0.049	0.048	0.04
	Deviation (% of control)	100	108	108	88
	SD	0.008	0.006	0.005	0.007
	n	10	10	10	10

Kruskal-Wallis and Wilcoxon test, two-sided; * $p \leq 0.05$, ** $p \leq 0.01$

2. OECD TG 443

An extended one-generation reproduction toxicity study was performed according to OECD TG Guideline 443 and under GLP (BASF SE, 2021). Wistar rats (25 per sex and dose) were treated with doses of 0, 200, 600 or 1800 ppm Benzenamine, N-phenyl-, reaction products with 2,4,4- trimethylpentene via diet. F0 animals were treated for at least 10 weeks prior to mating to produce a litter (F1 generation). Pups of the F1 litter were selected (F1 rearing animals) and assigned to 4 different cohorts (1A, 1B, 2A and 2B) which were subjected to specific postweaning examinations. Cohort 1B (F1 generation parental animals) was selected to produce F2 pups; Cohort 1A animals were analysed for specific reproductive parameters.

Absolute ovary weights were significantly decreased in the high dose group for both the parental (P0) as well as offspring (F1, Cohort 1B) generation. However, in Cohort 1A females, the absolute ovary weights were in the same range as the concurrent controls (Table 17).

Table 17: Absolute and relative ovary weights of all female animals investigated in the OECD TG 443 study

Concentration in diet [ppm]		0	200	600	1800
Approx. dose [mg/kg bw/d]		0	18	54	166
F0 [P0 parental animals]					
Absolute ovary weights [mg]	Mean	118.8	118.76	117.24	103.92**
	Deviation (% of control)	100	100	99	87
	SD	16.79	16.48	17.24	16.16
	n	25	25	25	25
Relative ovary weights [%]	mean	0.049	0.05	0.049	0.048
	Deviation (% of control)	100	102	100	97
	SD	0.006	0.006	0.007	0.007
	n	25	25	25	25
F1 generation, Cohort 1A [rearing animals]					
Absolute ovary weights [mg]	Mean	82.15	91.85	93.6	85.6
	Deviation (% of control)	100	112	114	104
	SD	13.97	17.04	19.93	17.53
	n	20	20	20	20
Relative ovary weights [%]	mean	0.041	0.045	0.047	0.046
	Deviation (% of control)	100	110	114	112
	SD	0.007	0.008	0.008	0.008
	n	20	20	20	20
F1 generation, cohort 1B [P1 parental animals]					
Absolute ovary weights [mg]	Mean	111.6	112.08	114.6	98.08**
	Deviation (% of control)	100	100	103	88
	SD	14.3	22.88	16.06	18.43
	n	25	25	25	25
Relative ovary weights [%]	mean	0.047	0.046	0.048	0.046
	Deviation (% of control)	100	99	103	97
	SD	0.005	0.009	0.007	0.007
	n	25	25	25	25

Kruskal-Wallis and Wilcoxon test, two-sided; * $p \leq 0.05$, ** $p \leq 0.01$

In addition to the ovary weights, also the body weights of the treated animals, particularly in the high dose group, were affected. Since treatment was conducted for all (F1) or most (P0) of the animal's life span, the relative organ weights correlating the animals' body weights with the respective organ weights are a crucial indicator to distinguish organ-specific effects versus

general growth delays (Ghasemi et al., 2021). No significant or biologically relevant changes were observed in the relative organ weights (Table 17).

Therefore, the decreased absolute ovary weights are considered secondary to reduced body weights rather than substance-specific organ toxicity. This is supported by the lack of histopathological findings in the rat ovaries, thus demonstrating that the reductions in absolute ovary weights should be considered as general systemic toxicity including reduced body weights rather than an organ-specific reproductive toxicity effect.

Changes isolated to absolute organ weights, without histopathological correlates or functional impairment should not be considered an adverse substance-specific effect (Sellers et al., 2007).

To further support a lack of functional impairment on the ovaries, the data obtained from differential ovarian follicle count (DOFC) performed in the OECD TG 443 study should be considered. Table 18 shows the DOFC in control and high dose animals of F1 generation Cohort 1A at sacrifice. No difference with respect to primordial and/or growing follicles can be seen.

Table 18: Differential ovarian follicle count of control and high dose groups in F1 Cohort 1A observed in OECD TG 443 study

Concentration [ppm]		0	1800
Approx. Dose [mg/kg bw/d]		0	166
Number of animals	N	20	20
Primordial	N Mean	7157 358	6878 344
Growing	N Mean	608 30	661 33
Primordial + growing	N Mean	7765 388	7539 377

Overall, the changes to ovary weights were only observed in absolute organ weights in the high dose group together with reduced body weights. Relative ovary weights were not significantly affected, and no changes were found in DOFC analysis. Therefore, the ovary weight changes proposed by the DS are considered to be of no toxicological significance and should be disregarded for classification purposes.

Annex 4 – Detailed information on data for proposed estrus cycle changes

In the present CLH proposal, the DS claims that the substance Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene affects the estrus cyclicity on the basis of the OECD TG 421 and 443 study:

> *"Increase of estrous cycles length associated with an increased mean percentage of days in diestrous stage was observed in high-dose dams exposed to Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene and Reaction products of diphenylamine with nonene, branched. Statistical significance was reached in high-dose females exposed to Benzenamine, Nphenyl-, reaction products with 2,4,4-trimethylpentene in the OECD TG 421 and in high-dose P1 females exposed in the EOGRTS."*

A disruption in cycling is characterized by persistent estrus, diestrus, or an extended duration of irregular cycles (Goldman et al., 2007).The OECD TG 421 and 443 studies conducted with Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene do not indicate a substance-specific, treatment-related adverse effect on the estrus cyclicity of rat as discussed in detail in this Annex.

1. OECD TG 421

An OECD TG 421 study with extended study design was conducted as a dose-range finder for the following OECD TG 443 study (BASF SE, 2020a). Wistar rats (10 animals per sex and group) were treated with doses of 0, 300, 1000 or 3000 ppm Benzenamine, N-phenyl-, reaction products with 2,4,4- trimethylpentene in diet for ten weeks prior to mating and throughout mating, females were further treated during gestation and 13 days of lactation.

Increased estrus cycle lengths were observed in the high dose group (3000 ppm), only. This was mostly due to a longer time of the animals in diestrus (7.8 vs. 5.6 days in controls). Within the high dose group, three of the ten animals in the group showed one cycle with three days in diestrus and one animal showed one long diestrus phase (exceeding three days) with four days in diestrus. Single estrus cycles with a three-day diestrus detection are considered normal for rats (see below). The one animal with one four-day diestrus cycled normally (1-2 days in diestrus) for the remainder of the observation time. Thus, the estrus cycle variability observed in this study is considered of limited biological relevance. Furthermore, no changes in mating or fertility index were observed, thus supporting the lack of functional impairment of fertility in female rats.

2. OECD TG 443

An extended one-generation reproduction toxicity study was performed according to OECD TG Guideline 443 and under GLP (BASF SE, 2021). Wistar rats (25 per sex and dose) were treated with doses of 0, 200, 600 or 1800 ppm Benzenamine, N-phenyl-, reaction products with 2,4,4- trimethylpentene via diet. F0 animals were treated for at least 10 weeks prior to mating to produce a litter (F1 generation). Pups of the F1 litter were selected (F1 rearing animals) and assigned to 4 different cohorts (1A, 1B, 2A and 2B) which were subjected to specific postweaning examinations. Cohort 1B (F1 generation parental animals) was selected to produce F2 pups; Cohort 1A animals were analyzed for specific reproductive parameters.

No changes in estrus cycle length were observed in the F0 generation (4.1 vs. 4.1 days in concurrent controls). In the F1 Cohort 1A females, estrus cycle duration of control and treated animals were comparable (3.9, 3.9, 4.0 and 4.0 days in control, low, mid, and high dose groups, respectively). In the F1 Cohort 1B (P1 females), a slightly but significantly prolonged estrus cycle duration was found in the high dose group (4.0, 4.0, 4.0, and 4.3** days (**p ≤ 0.01), in control, low, mid, and high dose groups, respectively). In this group, 3/25 animals showed one single cycle with three days diestrus duration, and two females showed two cycles with three days diestrus duration. Three further females were detected with four days diestrus duration, only one of which showed multiple cycles with long diestrus. All other females were found with normal cycles (1-2 days in diestrus) during the observation period. Since these findings were not observed in the concurrent Cohort 1A or the parental generation, and no impairment of fertility was observed, the biological significance of these findings is questionable.

Generally, the rat estrus cycle is divided into proestrus, estrus, metestrus and diestrus. Average duration of diestrus is around 55-57 hours (Chaitra et al., 2020; Cora et al., 2015; Paccola et al., 2013; Westwood, 2008), thus with daily staging during study conduct, a 1-3 day-long detection of diestrus is considered normal. Therefore, a cycle duration of 4-5 days is considered normal in rats and together with the historical control data of the laboratory showing cycle durations of 3.9-4.6 days, the slightly longer estrus cycle observed in the F1 Cohort 1B is considered of limited toxicological relevance.

Thyroid hormone imbalances have been shown to interfere with estrus cyclicity (Wei et al., 2018). Experiments conducted on phenobarbital-treated rats showed a dose-related increase of irregular estrus cycles, which could be reversed by administration of T4 (Li et al., 2011). Based on a similar mechanism of action observed with Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene (secondary hypothyroidism due to liver enzyme induction, see Annex 2, section 4 Mode of action for effects on fertility), single inconsistent findings may be considered attributable to a disturbance of thyroid hormone balance rather than a substance-specific effect. Finally, no impairment of mating or fertility index was observed in any of the studies conducted, thus supporting the lack of functional impairment of fertility in female rats following treatment with Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene.

Further information on the mode of action proposed for Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene can be found in Annex 2, section 4 Mode of action for effects on fertility.

Annex 5 – Detailed information on neurohistopathology and neuromorphometrics

In the CLH dossier, the DS makes the following claims regarding structural effects on offspring development based on the OECD TG 443 study performed in rats with substance Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene:

> *"In the high-dose C2A animals of the EOGRTS, neurohistopathological findings (increased incidence of axonal degeneration in the thoracic cord 9/10 males as well as slight increased incidence of axonal degeneration in other area in males and females) and neuromorphometric changes (decreased brain length in males and increased corpus callosum width in both males and females) were observed. […]* ***Based on the available data, there is clear evidence that Substance [2] induces abnormalities in the central nervous system in animals exposed during the developmental period. Such parameters were not investigated for substance [1]."***

1. OECD TG 443

An extended one-generation reproduction toxicity study was performed according to OECD TG 443 and under GLP (BASF SE, 2021). Wistar rats (25 per sex and dose) were treated with doses of 0, 200, 600 or 1800 ppm Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene via diet. F0 animals were treated for at least 10 weeks prior to mating to produce a litter (F1 generation). Pups of the F1 litter were selected (F1 rearing animals) and assigned to 4 different cohorts (1A, 1B, 2A and 2B) which were subjected to specific postweaning examinations. Cohort 1B (F1 generation parental animals) was selected to produce F2 pups; Cohort 1A animals were analyzed for specific reproductive parameters. Cohorts 2A (C2A) and 2B (C2B) were employed to investigate developmental neurotoxicity.

Neurohistopathology

As noted by the DS, a slightly higher incidence of axonal degeneration in the thoracic cord was observed in 9 of 10 F1 adult C2A male rats of the highest exposure group (i.e. 1800 ppm) tested with the substance Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene. However, it is our opinion that the CLH dossier fails to put this finding into proper context from a weight-of-evidence perspective. Furthermore, the claim that a slight increase in the incidence of axonal degeneration occurred in "other" areas in males and females does not accurately reflect the data or report. Other than the thoracic cord, there are no data to support the presence of axonal degeneration in any other areas in male or female rats at any dose level in the OECD TG 443 study with the substance. Therefore, the following paragraphs will focus on the CLH claims that the substance induces axonal degeneration in the thoracic cord and neuromorphometric changes in the brain.

There are several direct lines of evidence that indicate that the slightly higher incidence of axonal degeneration in the thoracic cord of the high dose C2A males are either an age- or rat-specific artifact, or at worst, a secondary response to chronic high dose exposure. There is no evidence in this study of a developmental origin for this finding or that this change produced functional effects on the central nervous system. Firstly, the slightly higher incidence of this finding in high dose C2A males occurred against a high background of axonal degeneration in the thoracic cord of both males and females across all dose groups including healthy concurrent control animals. As shown in Table 19, 20% of C2A control males and 40% of C2A

control females exhibited this same finding, with an equal incidence and severity in control and high dose females. There is no evidence of a dose-response pattern in C2A females, and no clear indication that severity worsened with increasing dose for either sex. The severity reported by the study pathologist across all dose groups including controls was graded as "minimal" (lowest severity grade), with only a single study animal showing a slightly higher grade of "slight to mild" at the highest dose tested of 1800 ppm, which is a dose that also produced overt signs of systemic toxicity in both F0 and F1 animals.

While the DS states that axonal degenerations were also slightly increased in other areas "(2 vs 0 tibial nerve degeneration in high dose males; lumbar cord axonal degeneration and sciatic nerve degeneration in 2 high dose females vs 0 in controls)" (CLH dossier, p.57), it should also be pointed out that axonal degeneration is commonly observed in all age groups of rats as an occasional spontaneous finding (Blankenship et al., 2016; Kaufmann et al., 2012). These findings mentioned by the DS were either made in single animals or are comparable to control. Based on the common observation of axonal degeneration in rats as incidental finding in this type of investigations, no toxicological weight can be given to single appearances of these findings without corroborating information.

Table 19. Incidence and severity of axonal degeneration in the thoracic cord of F1 generation Cohort 2A (PND 77) male and female rats following chronic exposure with Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene in an OECD TG 443 study (BASF SE, 2021)

Thoracic cord	Males				Females			
Concentration in diet [ppm]	0	200	600	1800	0	200	600	1800
No. of animals	10	10	10	10	10	10	10	10
Degeneration, axonal	2	5	4	9	4	3	1	4
Grade 1 (minimal)	2	5	4	8	4	3	1	4
Grade 2 (slight, mild)				1				
Gitter cells present	1	2	1	5	2	1	0	0

In addition to the high background rate and minimal severity observed in adult animals across all dose groups, including healthy controls, this finding did not correlate with any functional neuromuscular or neurobehavioral deficits in these same C2A animals. The results of the functional observations batteries (FOBs) and neurobehavioral testing in these animals were normal (for further information, see also Annex 6 – Detailed information on proposed neurological functional deficits).

One weakness of the OECD TG 443 study is its inability to distinguish between chronic neurotoxicity effects of exposure and latent neurodevelopmental effects in Cohort 2 animals. Unlike the OECD TG 426 test guideline for developmental neurotoxicity (DNT) studies, exposure to the test substance in an OECD TG 443 study continues beyond PND 21 until the time of scheduled necropsy at 12 weeks of age for C2A (adult) animals. Importantly, neurohistopathological evaluation of the C2B adolescent animals showed no evidence of axonal degeneration in control or high dose animals earlier in development on PND 22, suggesting that the presence of this finding in C2A adult animals, including controls, was either age-related or the result of chronic exposure.

The DS provides the argument that the neurohistopathological changes observed are likely developmental, since a previous OECD TG 422 study performed with the substance showed no histopathological findings on spinal cord and sciatic nerve. However, it should be taken into consideration that during an OECD TG 422 study the fixation procedure of tissues is suboptimal for detection of these effects when compared to the neuropathological investigations performed for the DNT cohort of the OECD TG 443 (immersion fixation vs. perfusion fixation). In addition, the sectioning in which this finding was detected in the

neurohistopathology of the OECD TG 443 study DNT module were longitudinal sections whereas an OECD TG 422 study only includes cross-sections of the spinal cord. Therefore, a very slight finding in the neuropathological investigation of the OECD TG 443 study DNT cohort might not have been detected in the OECD TG 422 study. Additionally, no findings in the thoracic cord were noted in generation F0 and F1A in the OECD TG 443 study. In these cohorts, the thoracic cord samples were immersion fixed and examined as cross sections comparable to the OECD TG 422 study.

Neuromorphometrics

The DS claims that neuromorphometric changes were observed in high dose F1 animals in the OECD TG 443 study with the substance. Specifically, the DS noted a decrease in brain length in C2A males and an increase in the width of the corpus callosum in both C2A males and females.

Figure 2 shows brain length and width data for individual C2A adult males across all dose groups. Mean brain length x width measurements for C2A adult males were 21.28 mm x 15.81 mm for control males versus 20.59 mm x 15.56 mm for high dose males, which represents mean differences in brain size of less than 1 mm, i.e., 0.69 mm difference for length (-3.2%) and 0.25 mm difference for width; the changes observed for brain width did not reach statistical significance. As shown in Figure 2, part of this difference may be attributed to two control males (No. 602 and 604) with higher-than-average brain length and width measurements (as compared to HCD, Table 20). Individual brain length and width measurements for all C2A males in the high dose group fell within the range of values in the concurrent control group. In addition, the absolute brain weights of C2A males were 95% of controls, whereas relative brain weights were 102%. The reduced brain length of -3.2% was in the same range as the reduced absolute brain weights and thus is more likely to be a consequence of the animals being smaller rather than a substance-specific effect. In the absence of any corresponding effects on brain width, neurobehavioral or FOB endpoints in males or females (see also Annex 6 – Detailed information on proposed neurological functional deficits), these data do not support the claim of a neuromorphometric effect on brain length in C2A adult males. This is especially true given that literature measuring differences in brain size found no direct correlation with overall cognitive performance (Schoenemann et al., 2000).

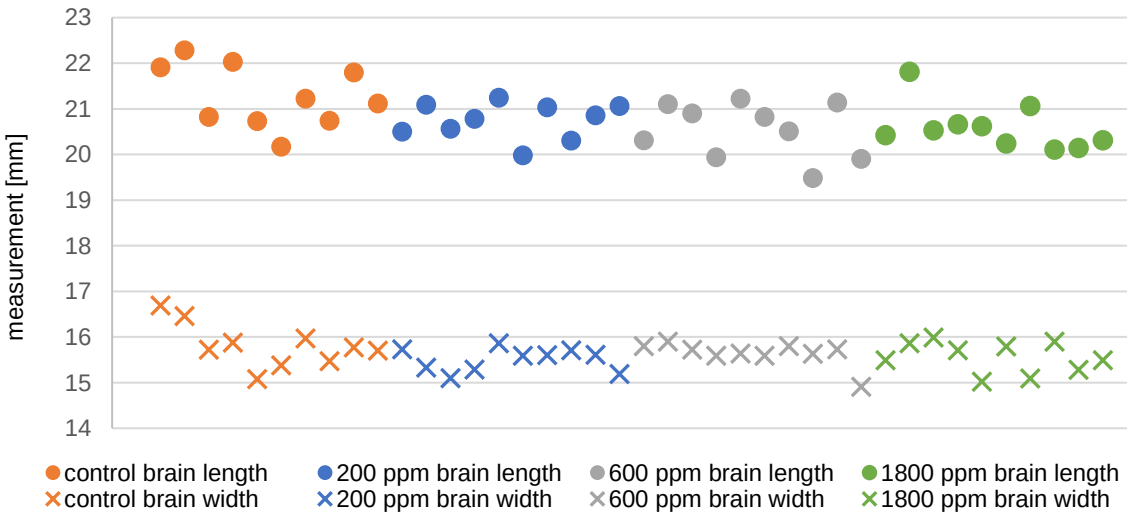


Figure 2: Individual animal brain length and width of F1 generation Cohort 2A (adult) male rats in OECD TG 443 study (BASF SE, 2021).

Table 20: Historical control data compiled for neuromorphometrics and brain weights in male animals based on five OECD TG 443 studies with DNT cohort performed between 2014 and 2024 (analysis at PND 77)

	Brain			
	Length [mm]	Width [mm]	Absolute weights	Relative weights
Mean	21.116	15.859	2.048	0.650
Min.	20.760	15.580	1.955	0.636
Max.	21.700	16.600	2.132	0.670

The DS claims that brain morphometric analysis shows an increased width of the corpus callosum in high dose C2A adult males and females. Mean width of the male and female corpus callosum in the high dose group was 0.69 mm and 0.65 mm, respectively, compared to 0.59 mm and 0.56 mm in the concurrent control males and females. Of note, there is considerable variability across individual animals. Using statistical analysis suitable also for non-normally distributed data, no statistical significance could be found for either sex. Only if males and females are considered together in a statistical test assuming normal distribution is statistical significance attained.

Historical control data was compiled on corpus callosum width based on five OECD TG 443 studies with DNT cohorts conducted between 2014 and 2024 (Table 21). Three studies were conducted via diet, one via gavage and one via drinking water. The HCD are depicted for each, males and females, given that there is generally a sex-dependent difference in measurements. Based on the historical control data available, the values reported for high dose animals are well within the normal biological variation observed for this parameter in this rat strain. For the control group males, the values were unusually low with 8/10 animals below the HCD for this parameter. In addition, the group mean value for Corpus callosum width for control males was below the HCD of the laboratory (Table 21). In females, 4/10 animals of the control group were below the HCD range of the laboratory, while the group mean value was just within.

In addition, a coefficient of variation was determined for this parameter by the laboratory, to provide further orientation for normal biological variation. Based on the variations commonly observed for this parameter, the change of high dose group as compared to controls of +16% in males and females is within the commonly observed biological variation range for this parameter.

Overall, the changes presumed by the DS are of limited biological relevance.

Table 21: Historical control data compiled for neuromorphometrics (corpus callosum width) based on five OECD TG 443 studies with DNT cohort performed between 2014 and 2024 (analysis at PND 77)

	Corpus callosum width [mm]		Corpus callosum coefficient of variation [%]	
	Females	Males	Females	Males
Mean	0.68	0.74	18.89	16.53
Min.	0.54	0.68	12.56	13.44
Max.	0.76	0.8	23.89	21.6

Moreover, the direction of change (increase in width) does not correlate with other parameters that would be considered of toxicological concern. The width of the corpus callosum is one of 10 morphometric measurements collected on the brains of the C2A animals in an OECD TG 443 study. There was no indication of corresponding effects in any of the other nine brain regions analyzed by morphometry in these same C2A males or females. An additional assessment of the data available with the statistical method conducted by the DS (two-way ANOVA accounting for treatment and sex) did not show statistically significant changes in any other neuromorphometric parameter (Table 22). The DS found it noteworthy that “the corpus callosum is the principal inter-hemispheric myelinated tract (white matter) and

histopathological findings linked to myelin degeneration in the cord white matter were observed in C2A animals." However, neurohistopathological findings in the study were limited to the thoracic part of the spinal cord, no relevant findings were obtained in other sections. Many studies have been published on the detrimental effects of a smaller or missing corpus callosum in rats and humans, but there is no evidence in studies of rats that shows a small increase in the corpus callosum without other physiological effects causes any differences in overall development. In addition, studies that reviewed a thickened corpus callosum in humans indicate that clinical manifestation of this change was never seen in isolation but "invariably associated with additional brain abnormalities" and "part of a neurogenetic syndrome in most cases" (Lerman-Sagie et al., 2009). Therefore, a slight increase in one of 10 measured brain morphometric parameters (i.e., width of corpus callosum), in the absence of corresponding effects on any of the other nine brain morphometric parameters or neurobehavioral or FOB endpoints in males or females (see Annex 6 – Detailed information on proposed neurological functional deficits), does not support the DS claims of a neuromorphometric effect in C2A animals. Contrary to the DS claims that a few isolated findings represent *"clear evidence that Substance [2] induces abnormalities in the central nervous system in animals exposed during the developmental period"*, the weight-of-evidence presented here demonstrates otherwise; the findings were either present in healthy control animals at unusually high background rates, are known to occur spontaneously, showed minimal severity, lacked clear dose-response patterns, represented small and sometimes non-significant differences, were not consistently seen at other developmental timepoints examined, and/or occurred in isolation with no correlation to other neuromorphometric parameters or functional signs of neurotoxicity or neurobehavioral effects.

Table 22: Analysis of neuromorphometric parameters obtained in OECD TG 443 study (BASF SE, 2021) using two-way ANOVA (sex,treatment) to compare control and high dose groups.

Neuromorphometric parameter	p-value
Base of lobus vermis cerebelli	0.65044
Corpus callosum width	0.03675
Frontal cortex left	0.94775
Fronal cortex right	0.99080
Hippocampus left	0.82395
Hippocampus right	0.23663
Nucleus caudatus width left	0.11406
Nucleus caudatus width right	0.73170
Parietal cortex left	0.21128
Parietal cortex right	0.20159

Discussion of proposed developmental neurotoxicity due to thyroid imbalances

None of the studies available with the substance show significant effects on thyroid hormones in pups prior to weaning. The DS suggests that a relevant effect might have occurred in the OECD TG 443 study even in the absence of statistical significance. Looking at the absolute values of thyroid hormones, this must be clearly contradicted. Well-known within the toxicological community, thyroid hormone values show a large intragroup variation to begin with, making data interpretation very challenging (Beekhuijzen et al., 2019; European Chemicals Agency., 2023; Li et al., 2019). Given that control males were below the HCD values already, a decrease from 16 to 14.1 nmol/L must be regarded as minimal, especially looking at the standard deviation of above 5 for controls. As the values were low to begin with, every minimal decrease seems large in relative numbers, but without corroborating findings such as statistical significance or increases in TSH, which were clearly not observed, no toxicological relevance can be attributed. On PND 22, statistically significant increases in TSH were found in both males and females, however all values were within the HCD of the

laboratory and therefore within the biological variation commonly observed for rats of this age and strain. No concurrent change in T4 values was found. Therefore, these findings must be considered of limited toxicological relevance.

The DS mentions that *"T4 and TSH measurements were not performed in F2 pups which is not in line with OECD TG 443 requirements"* (CLH dossier, p. 58). However, it should be clearly pointed out that the OECD TG 443 guideline does not specify thyroid hormone measurements in the F2 pups. While it is commonly done following ECHA's recommendations on OECD TG 443 studies (European Chemicals Agency., 2023), this publication was not available at the time of study conduct and could therefore not be taken into consideration. Therefore, the statement by the DS is incorrect and the study was conducted in full compliance with the OECD Guideline 443.

2. Assessment of available data in the context of CLP

RAC-62 guidance (RAC-62, 2022) on addressing developmental neurotoxicity and neurotoxicity under the current CLP hazard classes concludes the following:

"Adverse effects on the nervous system investigated or detected at any point in the life span of the organism exposed during the developmental period, covering both prenatal and postnatal development until sexual maturation (determined by preputial separation and vaginal opening), should be addressed under developmental toxicity (DNT), even if the exposure had also continued after sexual maturation."

This guidance references CLP 3.7.1.4 in further defining developmental toxicity as "any effect which interferes with normal development." Such effects can manifest at any point in the life span of the organism and includes death of the developing organism, structural malformations, altered growth, and functional deficiency. While we do not disagree with the scientific rationale for these guidance criteria, we do question the broad application of such criteria particularly for categorizing specific findings in an EOGRTS as Category 1B or Category 2. The sole basis for this guidance according to RAC-62 is due to the fact that:

"It is generally not possible to distinguish the precise origin or timing of the toxicological insult when the exposure has continued after the developmental period."

This inability to precisely distinguish between developmental toxicity versus chronic toxicity in adult animals within the scope of a given study design seems to be a reasonable basis for further investigation, not a basis for automatic classification of any observable change as a definitive developmental toxicant (i.e., Category 1B). At worst, such a finding in F1 C2A adult animals may warrant ultraconservative classification as a suspected developmental finding (i.e., Category 2) until further data become available for more precise classification. However, the DS has taken the extreme approach of broadly applying this RAC guidance to automatically categorize an already questionable isolated finding without functional impairment (axonal degeneration) in F1 C2A adult males of the OECD TG 443 study as a definitive Category 1B development neurotoxicant with little consideration for the scientific evidence. As previously noted, axonal degeneration was observed in F1 C2A males and females across all dose groups including 20-40% of healthy control animals, and the severity was the same across all groups with no evidence of a dose response pattern, i.e., minimal to slight (lowest severity grades). Even more importantly, there is absolutely no evidence that this change "interfered with normal development" in any of these animals, which is the foremost property of a developmental toxicant. Specifically, this change did not manifest in any of the classic conditions associated with developmental toxicity, including death of the developing organism, structural malformations, altered growth or functional deficiency. No

functional deficits were observed in any of these same animals during the functional observational batteries or neurobehavioral assessments. This specific change has also not been confirmed or reproduced in other studies to date. Based on the weight of evidence, there is currently no sound scientific basis for classifying this specific change, which was observed in both treated and control animals with minimal severity and not associated with any functional deficits, as a definitive Category 1B developmental toxicant as proposed by the DS, especially without consideration of further studies or investigation to better distinguish the origin of this change in treated and control animals.

CLP Annex I, section 3.7.1.4 specifies that "[…] for pragmatic purposes of classification, developmental toxicity essentially means adverse effects induced during pregnancy, or as a result of parental exposure." Therefore, in the spirit of the legislation, a clear distinction should be made between effects arising from exposure during development as compared to subsequent adult life.

RAC further notes that a similar approach for considering findings in F1 C2A animals as evidence for developmental neurotoxicity should be broadly applied to other target organ toxicities investigated at any point in the lifespan of the offspring. In the broadest sense of this application, all findings observed in any of F1 or F2 cohorts in an EOGRTS would by default be considered developmental toxicity since it is not possible to distinguish the precise origin or timing of the toxicological insult within the scope of the OECD TG 443. While we understand the importance of not missing potential signals of developmental toxicity, we must agree as scientists that not all observed changes in F1 adult animals should by default be grounds for classification as a definitive Category 1B developmental toxicant, especially on the sole basis of being unable to distinguish the origin within the scope of the study design. At worst, follow-up studies or investigation would be warranted for more accurate hazard identification and classification. This is particularly true in this case where the finding in question was of minimal/slight severity across all groups, including an atypically high incidence of the same change in control animals, and not associated with any functional or neurobehavioral deficits.

Annex 6 – Detailed information on proposed neurological functional deficits

In the CLH dossier, neurological functional impairment was reported as effects on the auditory startle response as well as decreased habituation by the DS:

> *In the EOGRTS, despite limitations of the auditory startle response test, effects from the mid-dose level on mean maximal amplitude in males as well as decreased habituation from the mid-dose in males and in males and females combined are considered biologically relevant in the absence of appropriate statistical analysis (testing for interactions of sex, trial blocks and treatment) and positive controls.*
> *→ Based on the available,* ***there is some evidence that Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene induces functional deficiency****. However, the deficiencies of the test limit the reliability on the effects observed in auditory startle response. Such parameters were not investigated for its analogue.*

1. OECD TG 443

An extended one-generation reproduction toxicity study was performed according to OECD TG 443 and under GLP (BASF SE, 2021). Wistar rats (25 per sex and dose) were treated with doses of 0, 200, 600 or 1800 ppm Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene via diet. F0 animals were treated for at least 10 weeks prior to mating to produce a litter (F1 generation). Pups of the F1 litter were selected (F1 rearing animals) and assigned to 4 different cohorts (1A, 1B, 2A and 2B) which were subjected to specific postweaning examinations. Cohort 1B (= F1 generation parental animals) was selected to produce F2 pups; Cohort 1A animals were analysed for specific reproductive parameters. Cohorts 2A (C2A) and 2B (C2B) were employed to investigate developmental neurotoxicity.

Auditory startle response

As part of the developmental neurotoxicity investigations of the OECD TG 443 DNT module conducted with the substance, an auditory startle response test was carried out on post-natal day (PND) 24 in all animals of cohort 2A. The DS had the following comments on the ASR:

"The ASR test presented some limitations: poor reporting of the apparatus used, statistical analysis not in line with the NAFTA guidance (i.e. no mention, or results presented for testing for interactions of sex, trial blocks and treatment) for maximal amplitude and latency as well as a complete absence of any statistical testing for habituation (a variable required under OECD 443). Furthermore, the lack of HCD and positive control increases the possibility the risk of false negative findings taking into account the low statistical power in DNT investigations."

The following is a detailed description of the process:

On PND 24, the auditory startle response test was carried out in all animals of cohort 2A using the SR-LAB; STARTLE RESPONSE SYSTEM (San Diego Instruments, San Diego, CA, U.S.A.). For all animals, the examinations started in the morning at their respective test date. If, at a given test date, several animals were tested, the trials were conducted in a randomized sequence. Age-appropriately sized, tube shaped, transparent acrylic enclosures were used to accommodate the animals during the test. Each enclosure has an attached motion sensor and is mounted on a solid enclosure base. The entire unit is placed in a heavy wooden, plastic-laminated, isolation cabinet, which minimizes extraneous noise and vibrations. Each cabinet

is equipped with internal light and fan, and contains, in a separate compartment, a complete sound generation system able to produce background noise and white noise stimuli, the level of both is adjustable. The response of the animal to the sound stimulus is issued by the motion sensor as voltage which is automatically recorded over the entire response window. The highest voltage during the response window is labeled “maximum amplitude”, the time in milliseconds from the start of the response to the maximum amplitude is labeled “latency to the peak of the response”.

The animals were given a 5-minute acclimation period in the enclosure with a 70 dBA background noise. Then the startle response was recorded in 50 trials at a startle stimulus sound level of 120 dBA with a 5 - 10 second variable interval between the trials. Response was recorded for 50 milliseconds. Measurement was carried out with the light and ventilator switched on in the isolation cabinets; no food or water was provided during the test. Data (maximum amplitude, latency to the peak of the response) were analyzed in 5 blocks of 10 trials each.

All records were printed, signed and are part of the raw data.

No influence of the test substance on auditory startle habituation (maximum amplitude and latency) was observed in any male or female animal in all treated groups. The DS claims to have observed a decrease in the auditory startle response (ASR) relative to concurrent control animals. This perceived relative decrease in ASR of mid and high dose group animals stems from one animal in the control group, which showed an extremely high maximum amplitude compared to the other animals (Figure 3). Therefore, this one animal is considered an outlier. Similarly, three animals in the low dose group showed a mean maximum amplitude that was more than twice as high as those of the rest of the group. With data as variable as this, analysis of mean values can be error-prone due to outliers. Therefore, data can be either (a) reanalyzed excluding outliers or (b) analyzed using a median instead of a mean value for the group. If the data are analyzed without these outliers, the mid and high dose group males are within the same range as the concurrent controls (Table 23). Similarly, the median was calculated in addition to the mean value and also this calculation resulted in no relevant change of control vs. high dose group values (high dose group as 99.1% of control, Table 23).

Table 23: Mean maximum amplitude of startle response in PND24 males (Cohort 2A) in OECD TG 443 study (blocks 1-5) (BASF SE, 2021)

Concentration in diet [ppm]		0	200	600	1800
Full dataset	N	10	10	10	10
	Mean	398.5	477.1	350.8	324.7
	SD	227.4	253.6	85.1	63
	% of control [based on mean]	100	119.7	88.03	81.5
	Median	316.5	381	359	313.5
	% of control [based on median]	100	120.4	113.4	99.1
Without outliers	Mean	330.1	329.7	350.8	324.7
	SD	75.2	66.4	85.1	63
	N	9	7	10	10
	% of control	100	99.9	106.3	98.4

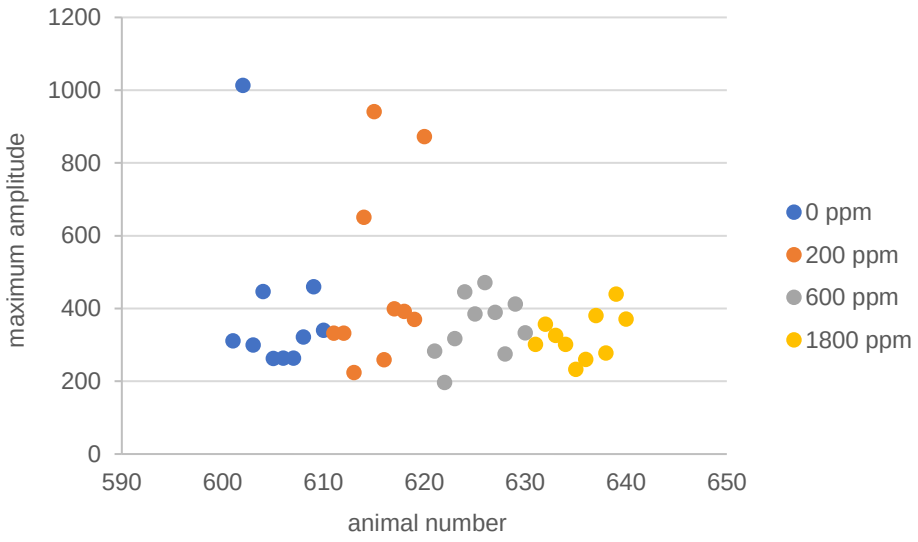


Figure 3: Individual animal data for auditory startle response in males at PND24 (mean maximum amplitude values) observed in OECD TG 443 (BASF SE, 2021)

In addition, historical control data were compiled for mean maximum amplitude parameter of auditory startle response. Data from eight studies containing ASR investigations were available between 2014 and 2022, and the control values observed were collected and are presented in Table 24. The mean maximum amplitudes of all groups were well within the historical control data of the laboratory and any changes are therefore considered not biologically relevant.

Table 24: Historical control data for maximum amplitude of ASR, obtained from 8 studies between 2014 and 2022.

	Block 1-5	
	♂	♀
Mean max. amplitude	443.5	404.9
SD	87.4	66.4
Min.	292.1	295.9
Max.	535.1	513.0

Habituation

Further, the DS analyzed the habituation response of the animals. Habituation of rats to auditory stimuli can be used as an approximation of learning. Generally, it should be emphasized that due to the low animal numbers and high individuality of the response, these results can only be considered indicative and definitive learning and/or memory experiments are required to investigate effects of a substance on learning/memory. While the DS calculated the difference between mean values of block 1 and 5 and then normalized the values to the concurrent control, looking at the single animal values, shows that this is difficult due to the spread of values and high standard deviations. Therefore, the startle response data were re-analyzed with mean values for each block and plotting those for the control and high dose groups (blue and orange icons in Figure 4, respectively). Based on these data, no relevant change in habituation can be identified based on the data available. Statistical analysis of the data (Kruskal-Wallis and Wilcoxon (two-sided)) did not provide any statistical significance of the data.

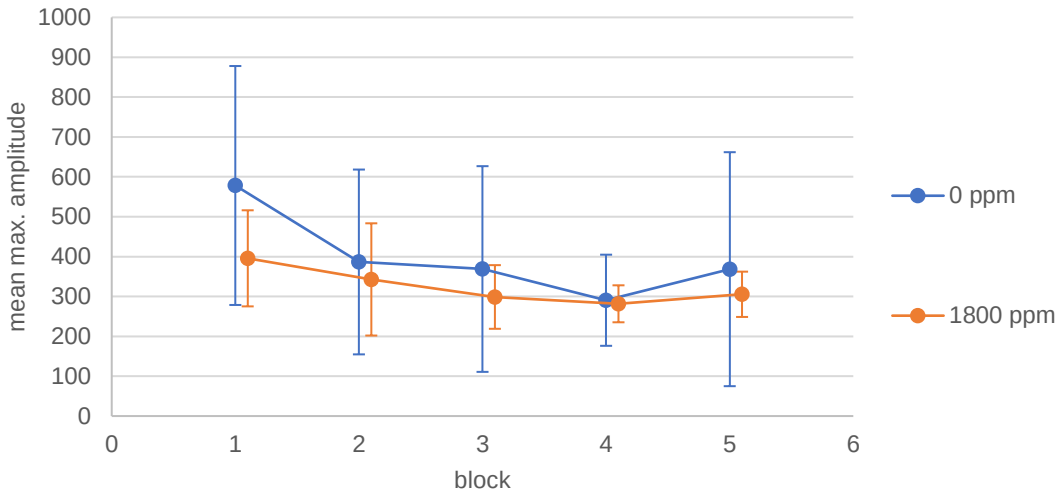


Figure 4: Mean max. amplitude of control (blue) and high dose (orange) C2A males over five blocks in response to auditory startle as observed in DNT module of OECD TG 443 study (BASF SE, 2021).

Based on these data as well as the lack of findings in the functional observational battery, it can be concluded that exposure of animals with the substance Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene during developmental stages does not lead to a functional impairment.

Annex 7 – Discussion of proposed information death of the developing organism

In the present CLH proposal, the DS claims that there is slight evidence that Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene induces death of the developing organism:

> *"In the generational studies performed with Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene and its analogue, the decreased litter size at birth observed in all studies results from the decreased number of implantation sites; this effect is addressed in the section dedicated to effects to sexual function and fertility. Post-implantation loss and foetal viability were not affected in any of the generational studies except in the OECD TG 422 where the mean post-implantation loss in the high-dose dams was increased compared to controls (14% vs 0%). It is noteworthy that the value in controls of the study was particularly low.*
> *Furthermore, the viability index was significantly reduced in this OECD TG 422 at the high-doselevel (88.7% at 225 mg/kg bw/d versus 100% in controls). However, no treatment-related effect was observed on live birth, viability and lactation indices in the OECD TG 421 up to 3000 ppm (eq. to 271 mg/kg bw/d) and in the EOGRTS up to 1800 ppm (eq. to 166 mg/kg bw/d) performed with Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene or in the OECD TG 421 up to 5000 ppm (eq. to 443 mg/kg bw/d) performed with its analogue Reaction products of diphenylamine with nonene, branched.*
> *→ Based on the available data,* ***there is slight evidence that Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene induces death of the developing organism*** *based on reduction of pre- and postnatal viability in the high-dose animals of the OECD TG 422 but these effects were not reproduced in the other generational studies performed with the substance and its analogue."*

1. OECD TG 422

In an OECD TG 422 study (WIL Research Europe, 2014), Wistar rats were treated with test substance at doses of 0, 25, 75 or 225 mg/kg bw/d via gavage in corn oil for two weeks prior to mating, throughout mating, gestation and lactation until post-natal day (PND) 4. No effects on fertility parameters were observed up to the highest dose tested.

Post-implantation loss was not reported as a finding in the study report. Analysis of single animal data provided mean post-implantation losses of 2.0, 10.0, 11.1, and 12.7%, respectively for control, low, mid and high dose groups (Table 26). For the calculation, post-implantation loss was in a first step calculated for each single animal and all animals without implantation sites (non-pregnant animals) were excluded. Several females were reported with higher litter numbers than implantation sites, which was explained in the study report due to resorption of implantation sites prior to staining. These females were defined as having 0% post-implantation loss. Then the mean values and standard deviations for post-implantation loss were calculated. The apparent differences to the values provided in the CLH proposal probably stem from the DS calculating the post-implantation losses from the group mean values and not from single animal values, which is the scientifically correct approach. No historical control data for this animal strain and laboratory are available in the study report, however historical control data from the performing laboratory are available for the years 2015-2017 and specify post-implantation survival index (Table 25). Based on these data, the post-implantation loss observed is within the commonly observed biological variation for this rat strain and study type. Comparing historical control data as well as mean post-implantation loss observed in the other available studies with the substance (derived from the same rat

strain; described in the following and Table 26) further support a lack of biological relevance. Additionally, other studies with this substance did not report increased post-implantation loss, thus supporting the lack of biological relevance of these values.

Table 25: Historical control data available for post-implantation survival index, collected in 48 OECD TG 421 and 422 studies between 2015 and 2017.

	Post-implantation survival index (%)
Mean	93
SD	5
Min	79
Max	108

2. OECD TG 421

An OECD TG 421 study with extended study design was conducted as a dose-range finder for the following OECD TG 443 study (BASF SE, 2020a). Wistar rats (10 animals per sex and group) were treated with doses of 0, 300, 1000 or 3000 ppm Benzenamine, N-phenyl-, reaction products with 2,4,4- trimethylpentene in diet for ten weeks prior to mating and throughout mating, females were further treated during gestation and 13 days of lactation.

The post-implantation loss reported was 11.5% in the control group, 0.0% in low dose group (300 ppm), 4.3% in mid dose group (1000 ppm) and 9.7% in the high dose group (3000 ppm). Thus, treated animals showed post-implantation loss in the same range of concurrent controls and the values reflected the normal range of biological variation inherent in the strain of rats used for this study as all respective values were within the range of the historical control data (Table 26).

3. OECD TG 443

An extended one-generation reproduction toxicity study was performed according to OECD TG 443 and under GLP (BASF SE, 2021). Wistar rats (25 per sex and dose) were treated with doses of 0, 200, 600 or 1800 ppm Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene via diet. F0 animals were treated for at least 10 weeks prior to mating to produce a litter (F1 generation). Pups of the F1 litter were selected (F1 rearing animals) and assigned to 4 different cohorts (1A, 1B, 2A and 2B) which were subjected to specific postweaning examinations. Cohort 1B (F1 generation parental animals) was selected to produce F2 pups; Cohort 1A animals were analyzed for specific reproductive parameters.

There were no indications for test substance-induced intrauterine embryo-/fetolethality since the post-implantation loss in the F0 generation did not show any statistically significant differences between the groups with 4.8% in controls, 9.0% in low dose, 3.4% in mid dose, and 10.2% in high dose group. In the P1 generation (F1 Cohort 1B), the post-implantation loss was 2.0% in controls, 4.5% in the low dose, 3.3% in the mid dose and 4.0% in the high dose groups. All values were within the normal range of biological variation observed for this rat strain (see historical control data, Table 26).

Table 26: Post-implantation loss observed in three different studies with Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentane (OECD TG 422, OECD TG 421, OECD TG 443) (BASF SE, 2021, 2020a; WIL Research Europe, 2014)

Dose levels		Control	Low dose	Mid dose	High dose	HCD
OECD TG 422	Mean [%]	2.0	10.0	11.1	12.7	n.a.
	SD	4.0	10.6	13.5	13.4	
	N	9	10	10	9	

OECD TG 421	Mean [%]	11.5	0.0	4.3	9.7	0.9 – 16.8
	SD	20.7	0.0	4.6	8.4	
	N	9	6	8	9	
OECD TG 443 (F0)	Mean [%]	4.8	9.0	3.4	10.2	2.4 – 17.7
	SD	8.7	19.9	5.9	19.0	
	N	24	24	25	25	
OECD TG 443 (F1)	Mean [%]	2.0	4.5	3.3	4.0	
	SD	4.0	6.2	6.9	7.5	
	N	23	24	24	25	

Annex 8 – Detailed discussion of proposed effects on pup growth

In the present CLH proposal, the DS concludes that:

> *"In rabbits exposed to Reaction products of diphenylamine with nonene, branched, a significant decrease in fetus weight (-12%) was observed at the highest dose (100 mg/kg bw/d) associated with delays of ossification. At this dose level, does consumed 31% less food than the concurrent control does during the treatment period (GD6-28), showed marked reduced defecation and had slight reduced corrected body weight.*
> *Postnatal growth was also altered from PND7 up to weaning in the high-dose groups of the generational studies performed with Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene EOGRTS, OECD TG 421) or its analogue (OECD TG 421). At these dose levels, effects on body weight of similar magnitude were observed in females at the end of the lactation period.*
> → *Based on the available data, there is evidence that both Reaction products of diphenylamine with nonene, branched and its analogue* ***alter growth of the developing organism*** *at dose levels also affecting maternal/parental weight."*

The authors would like to point out, that several data brought forward on this section of the document were derived from studies conducted with Reaction products of diphenylamine with nonene, branched. While the read-across is not considered reliable, the data are included as prenatal developmental toxicity studies on Benzenamine, N-phenyl-, reaction products with 2,4,4- trimethylpentane are not currently available.

1. <u>OECD TG 414 (rabbit, EC 701-385-4)</u>

The substance Reaction products of Benzenamine, N-phenyl with nonene (branched) was tested for its potential to induce prenatal developmental toxicity in New Zealand White rabbits. The test substance was administered as an aqueous suspension to groups of 25 inseminated female New Zealand White rabbits orally by gavage in doses of 0, 10, 30 and 100 mg/kg bw/d in 0.5% Carboxymethyl cellulose (CMC) on gestation days (GD) 6 through 28.

Clinical symptoms were found in mid and high dose does as increased incidences of no or reduced defecation together with reduced food consumption. The mean food consumption in the high dose group was significantly reduced during GD7-23 (up to -59% compared to controls). Overall, the high dose animals did consume 31% less food than the concurrent control animals during the treatment period.

The mean body weights (BW) and the average body weight gain (BWC) of the high dose rabbits were distinctly reduced. Overall, the high dose rabbits lost weight (-24.0 g vs. +104.4 g in control) during the treatment period (GD6-28, Figure 5).

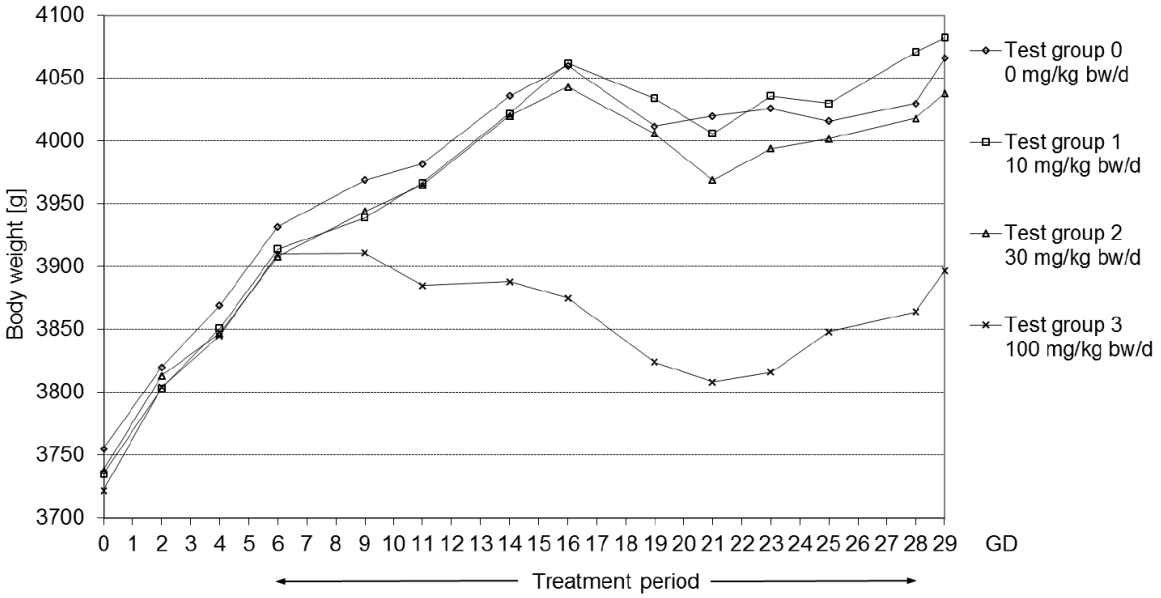


Figure 5: Mean body weights of pregnant animals observed in OECD TG 414 study with Reaction products of diphenylamine with nonene, branched in rabbits (BASF SE, 2019a)

The mean fetal weight of the high dose group was statistically significantly lower than control in male fetuses and when both sexes were combined (-12% in comparison to the concurrent control). The mean weight of the female high dose fetuses was also slightly lower but without attaining statistical significance. The mean fetal weights of low and mid dose groups were not influenced by the test substance and did not show any biologically relevant differences in comparison to the control group (Table 27).

Table 27: Rabbit fetal weights observed in OECD TG 414 study

Dose [mg/kg bw/d]	0	10	30	100
All viable fetuses				
Mean	38.3	38.6	36.3	33.6*
SD	5.23	6.72	5.07	5.93
N	20	24	24	20
Male Fetuses				
Mean	39.1	39.2	36.0	33.9*
SD	5.76	5.48	5.31	5.91
N	20	23	23	20
Female Fetuses				
Mean	37.6	37.3	36.0	33.6
SD	5.15	6.83	5.49	6.30
N	20	23	24	20

Statistics: Dunnett-test (two-sided), * $p \leq 0.05$

Reduced rabbit fetal weights were observed concurrent with severe maternal toxicity only. For the rabbit, the literature describes that in case of severe reduction of feed intake for does, the fetal weights are affected as well (Beyer et al., 2011; Cappon et al., 2005; Danielsson, 2013; Tyl, 2012). These findings are in quality and quantity comparable to the effects observed in the prenatal developmental toxicity study in rabbits.

Based on the correlation of maternal toxicity and reduced fetal weights as well as the literature available on this topic, this finding is considered secondary to maternal toxicity and not a substance-specific developmental effect.

2. OECD TG 421 (EC 701-385-4)

Reaction products of Benzenamine, N-phenyl with nonene (branched) was administered via diet to groups of 10 male and 10 female Wistar rats at concentrations of 0, 500, 1500 and 5000 ppm (BASF SE, 2020b). The duration of treatment covered a 10-week premating followed by a 2-week mating period in both sexes. Females were further treated throughout gestation as well as up to 13 days of lactation period.

Clinical examinations showed reduced food consumption in high dose females throughout the premating period, the gestation phase and during lactation (-11.6%, -20.1% and -23.5% vs. controls, respectively). In the mid dose group, the food consumption was reduced during gestation (-8.8% vs. control).

Body weights of high dose females were reduced during the premating phase, at the end of premating on study day 70, the decrease in body weight was -10.8% as compared to the concurrent controls (Table 28).

During gestation, a decreased body weight was observed for high dose and mid dose females (-18.5% and -8.8% on GD20 as compared to controls). During lactation period, decreased body weight in female animals of the high dose group (-16.9% as compared to controls) and in the mid dose group (-7.1%) was observed from lactation days 1 to 13.

Table 28: Data on body weights, body weight change and food consumption in OECD TG 421 study (BASF SE, 2020b)

		Main study groups							
Concentration in diet	[ppm]	0		500		1500		5000	
Approx. dose	[mg/kg bw/d]	0		40		122		397	
Body weights ♂ d91 / ♀ d70	Mean (♂ / ♀) SD Dev. vs. control [%]	412.9 29.1	217.9 11.9	403.1 15.5 -2.4	224.6 9.3 3.1	396.0 28.4 -4.1	212.3 13.4 -2.5	363.4** 21.7 -12.0	194.3** 14.5 -10.8
Body weights gestation (GD20)	Mean SD Dev. vs. control [%]	343.4 23.4		339.4 20.2 -1.2		313.3* 22.0 -8.8		279.8** 20.2 -18.5	
Body weights lactation (LD13)	Mean SD Dev. vs. control [%]	294.3 16.2		290.4 16.0 -1.3		273.5* 18.8 -7.1		244.6** 17.0 -16.9	
Body weight change ♂ d 0-91 / ♀ d 0-70	Mean (♂ / ♀) SD	296.2 28.6	122.1 8.5	285.9 13.4	128.9 9.9	279.7 23.0	116.1 112.	247.3** 19.9	97.8** 12.5
Body weight change gestation (GD 0-20)	Mean SD	118.8 12.4		107.3 12.5		95.7** 14.3		82.0** 16.6	
Body weight change lactation (LD 1-13)	Mean SD	36.5 9.6		39.0 7.1		31.4 7.7		32.2 11.2	
Food consumption d0 – d70	Mean (♂ / ♀) SD Dev. vs. control [%]	21.0 0.4	15.7 0.5	21.4 1.0 1.9	15.9 0.4 0.9	21.3 1.4 1.1	15.9 1.0 0.9	19.7 1.1 -6.4	13.9** 0.3 -11.6
Food consumption gestation (GD0-20)	Mean SD Dev. vs. control [%]	21.0 1.2		20.1 1.0 -4.1		19.2** 1.0 -8.8		16.8** 1.6 -20.1	
Food consumption lactation (LD1-13)	Mean SD Dev. vs. control [%]	47.2 3.7		46.0 3.5 -2.6		43.9 3.7 -7.1		36.1** 3.3 -23.5	

Dunnett test (two-sided), * $p \leq 0.05$, ** $p \leq 0.01$

Mean body weights of the male and female pups of test group 3 (5000 ppm) were statistically significantly reduced from PND 7 (-15.0% in male pups, -15.4% in female pups, and -15.7% both sexes combined) onwards until scheduled sacrifice on PND 13 (-18.8% in male pups, -18.4% in female pups, and -18.8% below control, both sexes combined, Table 29).

Decreases in pup weights were observed only in dose groups that also showed significant maternal toxicity. In the high dose group, body weight reductions in dams were of similar magnitude as compared to those in pups (-16.9% vs. -18.8%, respectively). No other growth delays were observed for the offspring.

Maternal toxicity is well-known to influence developmental parameters, including pup weights (Beyer et al., 2011; Danielsson, 2013; Tyl, 2012). Based on the concurrent incidence, the similar magnitude and the well-established causal relationship, the reduction in pup weights are considered secondary to unspecific maternal toxicity and not a direct substance-specific developmental toxicity.

Table 29: pup weights on post-natal day 13 observed in OECD TG 421 study (BASF SE, 2019a)

Concentration in diet	[ppm]	0	500	1500	5000
Males PND 13	Mean	33.0	31.8	30.4	26.8**
	SD	2.7	2.4	3.0	2.0
	Dev. Vs. control [%]		-3.7	-7.8	-18.8
Females PND 13	Mean	32.2	30.9	29.9	26.2**
	SD	2.2	2.3	2.4	2.1
	Dev. Vs. control [%]		-3.8	-7.2	-18.4
Males + Females PND 13	Mean	32.5	31.3	30.2	26.4**
	SD	2.3	2.3	2.7	2.1
	Dev. Vs. control [%]		-3.7	-7.3	-18.8

Dunnett test (two-sided), *p ≤ 0.05, ** p ≤ 0.01

3. <u>OECD TG 421</u>

Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene was administered via diet to groups of 10 male and 10 female Wistar rats (F0 animals) at concentrations of 0, 300 ppm, 1000 and 3000 ppm. The duration of treatment covered a 10-week premating period and a 2-week mating period. Females were further treated throughout gestation as well as 13 days of lactation.

Table 30. Data on body weights, body weight change and food consumption in OECD TG 421 study (BASF SE, 2020a)

		Main study groups							
Concentration in diet	[ppm]	0		500		1500		5000	
Approx. dose	[mg/kg bw/d]	0		40		122		397	
Body weights ♂ d 84 / ♀ d 70	Mean (♂ / ♀)	401.2	219.2	384.3	211.2	395.7	204.1*	365.0*	198.6**
	SD	13.1	13.4	36.4	14.4	29.4	11.7	27.0	10.6
	Dev. vs. control [%]			-4.2	-3.7	-1.4	-6.9	-9.0	-9.4
Body weights gestation (GD20)	Mean	337.3		340.3		309.8		281.0**	
	SD	23.6		14.7		14.7		17.7	
	Dev. vs. control [%]			0.9		-8.2		-16.7	
Body weights lactation (LD13)	Mean	289.2		292.7		265.2**		246.0**	
	SD	10.0		10.1		10.3		14.9	
	Dev. vs. control [%]			1.2		-8.3		-14.9	
Body weight change ♂ d 0-84 / ♀ d 0-70	Mean (♂ / ♀)	312.8	142.4	297.6	134.0	307.9	126.9*	276.2**	122.0**
	SD	11.8	11.3	34.1	12.8	28.1	9.4	23.7	11.4
Body weight change gestation (GD 0-20)	Mean	119.9		120.2		102.0*		81.9**	
	SD	17.6		10.3		7.2		9.6	
Body weight change lactation (LD 1-13)	Mean	36.0		41.4		35.3		31.2	
	SD	14.0		5.6		10.7		8.9	
Food consumption d0 – d70	Mean (♂ / ♀)	21.7	15.9	21.2	15.5	21.5	15.9	20.4	15.2
	SD	0.8	1.1	1.3	1.3	0.4	1.5	0.7	0.7
	Dev. vs. control [%]			-2.6	-2.0	-0.8	0.2	-6.1	-4.2
Food consumption gestation (GD0-20)	Mean	21.7		20.9		19.2**		16.7**	
	SD	1.7		1.3		1.1		1.0	
	Dev. vs. control [%]			-3.6		-11.8		-22.9	
Food consumption lactation (LD1-13)	Mean	46.5		46.9		44.5		37.3**	
	SD	4.2		2.8		1.5		2.8	
	Dev. vs. control [%]			0.9		-4.4		-19.7	

Dunnett test (two-sided), *p ≤ 0.05, ** p ≤ 0.01

When compared to controls, food consumption was significantly decreased in mid and high dose females during gestation (-11.8%, -22.9%, respectively, Table 30) and in high dose females during lactation (-19.7%). Mid and high dose females presented with decreased body weights at the end of pre-mating (-6.9% and -9.4%, respectively). At the end of gestation, a decreased body weight was observed in females of the mid and high dose groups (-8.2

and -16.7%, respectively). At the end of lactation, body weights of females were also reduced for mid and high dose groups (-8.3% and -14.9%, respectively).

Pup weights were also affected; in the high dose group, the weights of male and female pups were reduced on postnatal day 7 (males -18.1%, females -17.2%, and both genders -17.7%) and 13 (males -26.6%, females -25.9%, and both genders -26.4%).

In the mid dose group, the pup weights were decreased in male and female animals on postnatal day 13 (males -8.7%, females -8.1%, and both genders -8.4%). However, in the mid dose group the body weight values in males (30.6g), females (29.8g) and both gender (30.2g) were within the historical control ranges (male pups 28.5-34.0g, females pups 27.7-33.3g, and both genders 28.1-33.6g), therefore the alterations in pup body weight in this test group were considered as treatment-related but not adverse (Table 31).

Table 31: Pup weights on postnatal day 13 observed in OECD TG 421 study (BASF SE, 2020a)

Concentration in diet	[ppm]	0	500	1500	5000
Males PND 13	Mean	33.5	31.7	30.6*	24.6**
	SD	2.1	2.3	1.2	2.2
	Dev. Vs. control [%]		-5.2	-8.7	-26.6
Females PND 13	Mean	32.5	30.8	29.8*	24.0**
	SD	2.1	1.3	1.2	2.2
	Dev. Vs. control [%]		-5.2	-8.1	-25.9
Males + Females PND 13	Mean	33.0	31.3	30.2*	24.3**
	SD	2.1	1.7	1.2	2.2
	Dev. Vs. control [%]		-5.2	-8.4	-26.4

Dunnett test (two-sided), *p ≤ 0.05, ** p ≤ 0.01

Decreases in pup weights were observed only in dose groups that also showed significant maternal toxicity. No other growth delays were observed for the offspring.

Maternal toxicity is well-known to influence developmental parameters, including pup weights (Beyer et al., 2011; Danielsson, 2013; Tyl, 2012). Based on the concurrent incidence, the similar magnitude and the well-established causal relationship, the reduction in pup weights are considered secondary to unspecific maternal toxicity and not a direct substance-specific developmental toxicity effect.

Annex 9 – Detailed information on bioaccumulation in fish

In the present CLH proposal, the DS claims that

> ***"Measured partition coefficient and bioaccumulation test data***
> *No experimental data evaluating the bioaccumulative properties are available for the UVCB substance "Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene".*
> *One partition coefficient study performed according to the OECD test guideline 123 (Unpublished study report, 2019d) is available in the registration dossier. The full study report has not been reviewed by the DS. The study was performed using one constituent (C4C4DPA) of the UVCB substance. This study is a GLP study and is considered reliable without restriction according to the registrant. The method used to determine the octanol-water partition coefficient was a slow-stirring method. The analytical method showed recovery rates of 82% of the test item demonstrating a sufficient accuracy of the test item. The partition coefficient of test substance was determined to be log Kow= 6.7 at 23°C and pH 6.7.*
> *Considering the measured log Kow≥4 for C4C4DPA constituent and the estimated log Kow≥4 for the main constituents of the UVCB substance, it is therefore concluded that the substance has a potential for bioaccumulation in aquatic species. The BCF model predictions support this conclusion."*

For the substance itself there are no experimental test data on bioaccumulation available. However, rather than use the screening information such as log Know values for constituents as the DS has done, the registrant considered it more valid to read across for this endpoint to a study on a similar substance, performed on the constituent Mono-Nonyl diphenylamine (EC 248-295-7) and to then support the experimental results with QSAR modelling on the individual constituents.

Bioaccumulation potential has therefore been assessed via a weight of evidence approach, using both QSAR modelling and experimental data - an OECD TG 305 bioaccumulation study in fish performed with the test substance Mono-Nonyl diphenylamine (EC 248-295-7) at the MITI institute in 2000. Considering the available experimental data and the QSAR data of the different models, the BCF based on experimental data is used to conservatively assess the bioaccumulation potential and used for the chemical safety assessment. This value as calculated by the study directors is 1730 L/kg. The position of the Registrants is that the substance is not bioaccumulative, as per Annex XIII of regulation 1907/2006/EC and the Guidance on information requirements and chemical safety assessment Chapter R.11 (PBT assessment, May 2008).

However, the experimental study was performed to meet the requirements of the Japanese authorities but was not designed to meet the requirements of the updated OECD TG 305 guideline or the criteria required to evaluate the 'B' criterion of the PBT assessment under REACH. There are therefore a number of shortcomings in the study, as recognized by the DS, however despite this the DS has reassessed this data against new criteria and has produced an unrealistic inaccurate evaluation which cannot be used to robustly assess bioaccumulation. The DS has evaluated the available bioaccumulation study in fish for the constituent Mono-Nonyl diphenylamine (EC 248-295-7) and reached the conclusion that the BCF in the fish is not 1720 but 2219 L/kg. The registrants disagree with this evaluation and have found several inconsistencies, but in short there are many issues when reassessing a study under new criteria and there is inherent uncertainty in these estimations as described in

OECD TG 305.. In addition, the registrants consider that the shortcomings in the performed fish study contribute to an unrealistic worst-case evaluation which cannot be used to robustly assess bioaccumulation. Finally the data are close to the threshold of 2000, and the multiple uncertainties would best be dealt with by re-performing the bioaccumulation study in fish as per the current testing proposal.

The following information describes the weight of evidence in more detail and particularly the assessment of the fish BCF study.

1. QSAR prediction of bioaccumulation

Two QSAR predictions are provided in the REACH dossier. Of particular note is the prediction using CATALOGIC v5.11.19 BCF base-line model v02.09. This model follows the OECD principles for the Validation, for Regulatory Purpose, of (Q)SAR Models. The prediction falls within the applicability domain of the model and indicates that this substance is not bioaccumulative. Using the latest version of the model (CATALOGIC v5.16.1.10, BCF-baseline model v.06.13), the substance is also not considered to be bioaccumulative. The model predicts a log BCF of 2.86 (BCF = 724) (Figure 6) and is within the domain range of the model and follows the OECD principles for QSAR predictions. This model applies mitigating factors to account for factors such as metabolism and steric hinderances, and as such, provides a robust prediction of realistic bioaccumulation. Therefore, this prediction indicates that the substance is unlikely to be bioaccumulative. This is also supported by a range of substances located in the training set for the model, which are similar to this substance, and also do not meet the criteria for bioaccumulation. As such, based on these data, there is an argument that this substance is not bioaccumulative. Given the borderline nature of the experimental data (the 2000 MITI study) and the limitations of this study described above, these results further demonstrate the need to clarify the bioaccumulative properties of this substance. As such, a new bioaccumulation study, conducted in accordance with the updated OECD 305 guideline, is considered justified.

Substance	**Log BCF**
Target substance (MNDPA; EC248-295-7)	2.86 (predicted)
Analogues identified in the training set	**Experimental Log BCF**
Diphenylamine	2.18
N,N-Bis(octylphenyl)amine	1.00
N,N'-Diphenyl-p-phenylenediamine	3.12
1-(N-Phenylamino)naphthalene	3.23
2-(N-Phenylamino)naphthalene	2.23
p,p'-dioctyldiphenylamine	0.000
4,4'-bis(alpha,alpha-dimethylbenzyl)diphenylamine	2.00

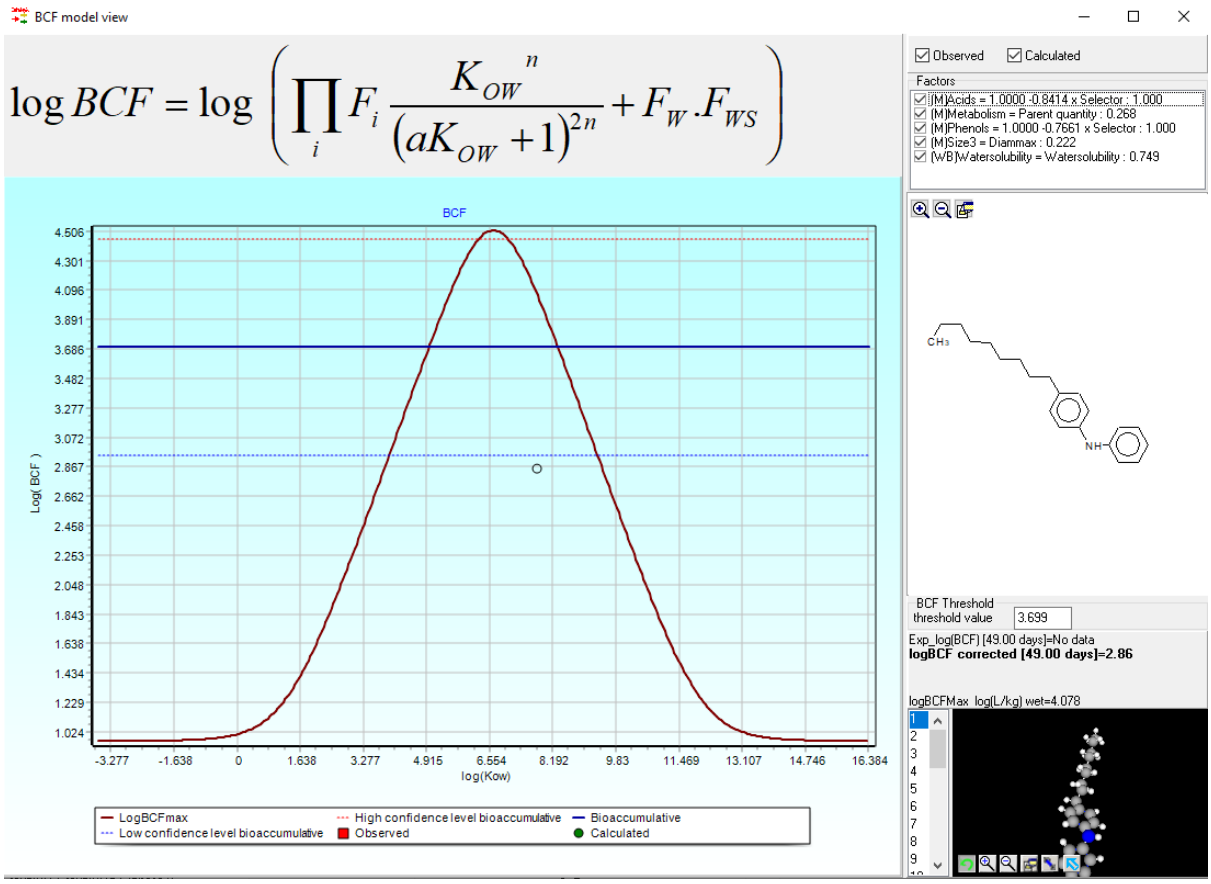


Figure 6: Predicted BCF of MNDPA

2. Hinderance of bioaccumulation based on molecular parameters

In further consideration of bioaccumulation, the European Chemical Agency (ECHA) considers D_{MAX} as a potential indicator of the likelihood of bioaccumulation. The D_{MAX} for MNDPA was calculated as 1.69 nm. The ECHA guidance Chapter R.11: PBT/vPvB assessment states "*From one study of a diverse set of substances it appeared that for compounds with a* $DMAX_{aver}$ *larger than 1.7 nm the BCF value will be less than 2000*". In this case provided by ECHA, the indicator value of 1.7 nm. However, this same guidance acknowledges that this is not an absolute cut-off and different models will calculate slightly different values. Therefore, it can be considered that MNDPA meets this threshold.

The threshold of 1.7 nm was based on the work of Dimitrov *et al.* (Dimitrov et al., 2003, 2002), cited in Environment Agency (2009). In this analysis, it was determined that there was an inverse relationship between D_{MAX} and log BCF, with the BCF decreasing exponentially as D_{MAX} increases. Dimitrov *et al.* (2003) determined that a D_{MAX} of 14.7 nm, a BCF of 5500 would be calculated, which would meet the criteria for very bioaccumulative (vB) under REACH. However, molecules with a D_{MAX} >1.5 nm had a log BCF of <3.3 (<2000), and therefore, do not meet the criteria for bioaccumulative (B) under REACH. As such, based on a D_{MAX} value of 1.69 nm, MNDPA can be considered unlikely to bioaccumulate.

This lack of bioaccumulation above a D_{MAX} of 1.5 nm can be explained by comparison with the cell membrane architecture; this threshold for maximum diameter is approximately the same as half the thickness of one of the two lipid layers that constituent a cell membrane (*Calculation of molecular dimensions related to indicators for low bioaccumulation potential*, 2009; Dimitrov et al., 2003, 2002). Therefore, at a D_{MAX} of >1.5 nm (15 Å), molecular size is such that passive

diffusion across the cell membrane does not readily occur, and therefore, substances above this threshold are not bioavailable. As such, the adoption of a threshold of 1.7 nm (17 Å) as a 'limit' for bioavailability in the scientific literature is likely to be highly conservative for the assessment of bioaccumulation potential.

A 2009 study by Nichols et al. further supports that MNDPA will poorly diffuse across cell membranes and therefore will not bioaccumulate. As shown in Figure 7 (taken from (Nichols et al., 2009)), the probability of a substance to cross the cell membrane decreases as the D_{MAX} increases. In this study, it was suggested that the probability of a substance crossing cell membranes decreases by 50 % when the D_{MAX} is 1.33 nm (Nichols et al., 2009). Based on this graph, it would appear that at a D_{MAX} of 1.69 nm, the probability of a substance crossing the cell membrane is ~10 %.

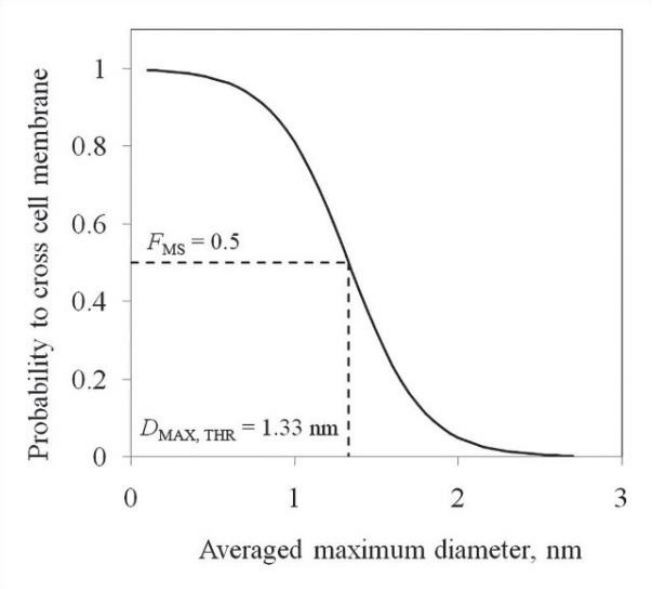


Figure 7 Relationship between the probability that a substance will cross a cell membrane and $D_{MA}X$ (Nichols et al., 2009)

A recent study on the fish intestinal membrane provides further support to indicate that MNDPA cannot readily pass through cell membranes (Miyata et al., 2022). This study assessed the absorption of nine chemicals across the intestinal tract of *Cyprinus carpio* using an everted gut sac method. The first four chemicals had $D_{MAXaver}$ values ranging from 1.16 to 1.51 nm, and the remaining five chemicals had $D_{MAXaver}$ values ranging from 1.72 to 3.49 nm. These results are summarized in Table 32.

For those chemicals where the $D_{MAXaver}$ were $\leq$1.51 nm ($\leq$15.1 Å), concentrations were higher in the inner solution than in the outer solution, indicating significant migration across the membrane, and therefore, the substance can be considered bioavailable. However, for chemicals with $D_{MAXaver}$ >1.51 nm (>15.1 Å), there was barely any passage through the intestinal membrane. Whilst there may be some differences in absorption characteristics between the fish intestine and gills, this study provides further supporting information on the use of $D_{MAXaver}$ for the assessment of bioaccumulation.

Table 32 Absorption of nine chemicals across a fish intestinal tract in an everted gut sac assay (Miyata et al., 2022)

Test chemical	$D_{MAXaver}$ (nm)	Concentration in:	
		Outer solution (mg/l)	Inner solution (mg/l)
1	1.16	122 ± 5.8	3350 ± 639
2	1.21	24.8 ± 6.2	305 ± 120
3	1.4	50.4 ± 5.4	303 ± 134
4	1.51	57.6 ± 4.0	140 ± 49
5	1.72	85.8 ± 4.2	<0.06 (limit of quantification)
6	2.33	25.9 ± 1.8	<0.1 (limit of quantification)
7	2.09	63.7 ± 5.1	<0.08 (limit of quantification)
8	3.49	97.7 ± 1.5	<0.1 (limit of quantification)
9	2.38	75.7 ± 5.1	0.3 ± 0.2

Therefore, based on a review of the literature information, it can be concluded that the D_{MAX} of MNDPA is sufficiently high that bioaccumulation will not occur, further indicating shortcomings with the 2000 MITI study. As such, it is considered appropriate to conduct a new OECD TG 305 study to address these shortcomings.

3. Experimental data and issues

A bioaccumulation study in fish is available for Mono-Nonyl diphenylamine (EC 248-295-7, MNDPA). Different modelling approaches identified the constituent MNDPA as the most critical component with regard to bioaccumulation for the registered UVCB ”Reaction products of diphenylamine with nonene, branched (EC 701-385-4)”. The bioaccumulation study in fish was performed in 2000 according to the protocol for New Chemical Substances (1974, amended 1998), which prescribes the procedure of testing new chemical substances as required by the Chemical Substances Control Law of Japan. The applied protocol is similar to the OECD TG 305 protocol of 2000. The OECD TG 305 protocol was revised and updated in 2012 and several significant modifications were implemented e.g. determination of a BCF kinetic (BCF k).

Comparing the old guideline protocol with the latest OECD TG 305 protocol, it becomes obvious that the modifications and deviations in the OECD 305 guideline lead to shortcomings/weak points (normalization to 5% fat, growth correction, depuration phase until 95% of the mass (TS) is removed) if a study performed in 2000 is evaluated according to the today-standards of the OECD 305 guideline.

The individual modifications and resulting shortcomings might seem small, but when considering the borderline nature of the results, they can result in the test substance being either non-bioaccumulative or bioaccumulative, based on the type of analysis used. This must be taken into consideration, as the registrant believes that these shortcomings clearly contribute to a worst-case evaluation which is not scientifically justified given the inherent uncertainty in the data.

Water solubility and solubilizer

The water solubility was determined in two independent studies. In the first study from 2002, the solubility ranged from 9-12 µg/L at 20°C in pure water. However, no details on the test substance characteristics (i.e., chemical name, CAS no., batch No., purity, etc.) were provided. As such, this study is considered of limited use in chemical assessment.

In a second, more recent study from 2010, the maximum water solubility at 20 °C for the whole UVCB (CAS 36878-20-3; including Mono-Nonyl diphenylamine) was determined to be 5 µg/L.

The results from this latter study providing also a full description of the test material, questions the maximum solubility of the selected target compound (Mono-Nonyl diphenylamine) in the MITI BCF study. The tested concentrations in the bioaccumulation study were 10 and 100 µg/L, far above the maximum solubility of the test material. This was due to the use of solubilizers to increase the concentration of test material in solution. Although the use of solubilizer is not excluded, the OECD TG 305 protocol does not recommend the use of solubilizers and suggests that their concentrations should be kept to a minimum. Notably, the concentrations of solubilisers in this study exceeded the concentrations of test substance in solution (<25 mg/L 2-methoxyethanol and 0.4-4 mg/L HCO-30), and neither solubiliser are listed in the current OCED 305 guideline. It should also be noted that in bioaccumulation studies performed for REACH registration, solubilizers are generally not recommended.

As the concentration of test substance in fish directly corresponds to the concentration in water, an artificial increase of the chemical concentration above the water solubility (as has been achieved here with the use of solubilizers) will also artificially increase the concentration of the chemical in fish. Therefore, a valid BCF value can only be determined, if the test concentration is in the range of the water solubility under test conditions. In the existing bioaccumulation study, this cannot be the case; the use of solubilizers has artificially increased the availability of test substance in water, which in turn may result in an artificially high bioaccumulation factor.

Exposure via the water phase vs dietary exposure

According to OECD TG 305, exposure via the water phase is appropriate for substance with a log Kow in the range of 1.5 to 6. For substance with a higher log Kow, exposure via the water phase should only be considered, if stable and dissolved concentrations can be achieved within the solubility limits. The highest concentration tested in the 2000 bioaccumulation study was 100 µg/L, which is demonstrably far above the limit of water solubility. The lower concentration used in this study (10 µg/L) was also likely above the limit of water solubility. Due to the limited details available, it is not clear if a stable concentration could be maintained without the use of a solubilizer.

The tested substance has an estimated log Kow of 7.6, a dietary exposure study based on the protocol of OECD TG 305 can be considered more appropriate. However, it is recognized that dietary bioaccumulation studies can provide significant challenges, notably in calculating a precise dose received by each fish, and with regards to the release of test substance from spiked food into the water phase, which may alter the type of exposure that takes place. As such, we would seek advice from the Regulator on the design of this study.

As an alternative, although not explicitly required under OECD TG 305, discussions with internal and external CROs indicate that from a practical point, tests with substances having a water solubility in the range of 10 µg/L and a high log Kow value should be performed with radiolabeled material in order to monitor the stable concentrations as well as uptake and loss rates of the substance in fish and to allow for a mass balance analysis. Therefore, it could be considered that running an OECD TG 305 study in the water-phase using ^{14}C-labelled test material, without the use of solubilisers may be appropriate. This would address the concerns with the 2000 bioaccumulation study highlighted above and will provide a definitive BCF value for this substance.

Kinetic BCF

The derivation of the kinetic BCF was introduced by the update of the guidance in 2012, however, was not required when the study was performed in 2000. Therefore, the study was designed to meet the requirements of the Japanese authorities, e.g. $BCF_{whole\ body}$. The registrants understand the demands of the authorities to provide a kinetic BCF value, a BCF value at steady state as well as BCF values normalized to 5% fat according to the current criteria of the OECD 305 guideline. These requirements cannot be reliability met using the study performed in 2000. However, an adjusted test design, considering the monitoring of individual test fish, exposure and depuration duration, the sampling intervals and the number of test fish is required:

Sixty fish were treated at each exposure level, however, only two fish were analyzed to determine the highest accumulated organs/tissues. The use of two fish is neither statistically significant nor representative of current practices, and therefore cannot be used to draw reliable conclusions of the bioaccumulation potential of a test substance.

Furthermore, to calculate an accurate kinetic constant, k2, the depuration phase should last until 95% of the mass is removed, or for a maximum of 56 days. The depuration phase was not studied completely in this Japanese testing. 82% of the substance had been removed within 42 days of depuration time. However, a partially k2 cannot be reliably extrapolated to the whole-study k2, as the k value fluctuated over the study period. For example, the k2 is about 0.037 and 0.05 in the first 10 days and second 10 days of depuration. About 50% difference in the two periods.

Based on ECHA's guidance (R.7c), the BCF_k can be calculated when the first-order kinetics apply. However, due to the limitations of the MITI study, it cannot be determined if this substance follows first-order kinetics. Therefore, it is clearly inappropriate to make such extrapolations based on this 2000 study and the re-calculated BCF value proposed by the DS should not be considered reliable or relevant for classification.

BCF values and growth correction

According to the revised protocol of OECD TG 305 from 2012, a kinetic and a steady state BCF should be normalized to 5 % fat, since investigations on aquatic organisms have shown a significant positive correlation between the accumulation of a chemical and the lipid content of organisms. Furthermore, the individual length of fish should be monitored in order to correct the BCF values for growth effects. However, in the MITI study length and weight measurements of individual fish or the lipid content of a representative fish sample were not provided, and thus any lipid and /or growth correction based on the data provided is not possible. Therefore, the BCF values may be considered with high uncertainty as they were not lipid- and growth normalized and thus neglecting the lipid content of the used test organism and a possible dilution due to possible/likely fish growth. In addition, the determination of the test substance in a mixed sample of two pooled fish did not allow to correct values according to the fish weight accounting for the high variation of the BCF values.

4. Conclusions

The bioaccumulation study in fish with the test substance Mono-Nonyl diphenylamine (EC248-295-7) at the MITI institute was performed in 2000 to meet the requirements of the Japanese authorities. The performed study was not designed to meet the requirements of the updated

OECD 305 guideline or the criteria required to evaluate the 'B' criterion of the PBT assessment under REACH.

When considered against the current criteria, the BCF study reveals several shortcomings for deriving a revised BCF, as performed by the DS. The additional information requested by ECHA and/or evaluating member states simply cannot be determined from this study with sufficient robustness to draw conclusions on the bioaccumulation potential of the tested substance. Many relevant data like the individual fish parameter (e.g., changes in weight and length or the lipid content) were not monitored but these have a great impact on the resulting BCF values. This is a particular concern for this substance, as the original analysis resulted in a BCF that was close to, but below the threshold for bioaccumulation under REACH. Re-analysis using current requirements and potentially flawed assumptions results in a BCF that is close to, but marginally above the threshold for bioaccumulation but which may be inaccurate and introduces more uncertainty. Furthermore, robust QSAR analysis, which included the identification of a range of similar substances indicates that the bioaccumulation criteria are not met. This is supported by analysis of the molecular parameter, D_{MAX}, which indicates that uptake, and subsequent bioaccumulation of MNDPA is likely to be hindered. As such, there is significant uncertainty regarding the use of the 2000 MITI study for PBT assessment.

The registrants consider that the shortcomings in the performed fish study to derive a new BCF contribute to an unrealistic inaccurate evaluation by the DS which cannot be used to robustly assess bioaccumulation.To overcome these shortcomings, the registrants suggest that a new OECD TG 305 study is performed to meet the requirements and criteria of the current guideline.

Annex 10 - Detailed information on long term toxicity on daphnia (Daphnia magna)

In the present CLH proposal, the DS claims that:

> *"The reported ELR10 of 1.68 mg/L (95% CI: 0.767-3.68) and NOELR of <0.625 mg/L is therefore based on the most sensitive endpoint (reproduction). No analytical monitoring was conducted. The solutions were visually inspected for the presence of any undissolved test substance. All test solutions were colorless clear and no undissolved test substance was visible. According to CLP guidance, although measured concentrations are preferred, classification may be based on studies where nominal concentrations are the only valid data available. Although a calculated EL10 is available, the DS used the NOELR in the application of the classification criteria. Considering a conservative approach in the absence of measured concentrations and in regards to the effects observed for substance [1] the NOELR <0.625 mg/L was kept for the application of the criteria for classification. Thus, based on NOELR of <0,625 mg/L from this test it can be concluded that the UVCB substance exert chronic toxicity to aquatic invertebrates. However, the results reported with the nominal loading rates of the substance probably under-estimates the chronic toxicity of the substance."*

The Registrants of this substance disagree with this proposal and consider that it is unnecessarily conservative; it adopts a very simplistic approach to dealing with uncertainty which does not properly consider the data. The registrants therefore fully support the testing laboratory's derivation of the EC 10 value from the different effect values based on nominal exposure rates, as recommended in the OECD TG 23. Considering the EC10 of 1.68 mg/L based on nominal values is valid, there is no toxicity within the water solubility of the substance and consequently no classification and labelling is required.

1. <u>Choice of toxicity value, and hazard level</u>

The data provide a valid EC/EL10 value which should be considered for classification in preference to the NOELR chosen by the DS. In the available long term daphnia study (OECD 211), organisms were exposed to nominal concentrations based on Water Accommodated Fractions (WAFs). Effects were related to adult mortality and reproduction, and during the exposure no mortality in daphnia were observed. In the lowest concentration of 0.625 mg/L loading rate the reproduction decreased by 4 % compared to the control, a value which was statistically significant. Therefore, no NOELR could be derived. However, the data were sufficient to derive an EL10-value for a loading rate of 1.68 mg/L. According to the Guidance on the Application of the CLP Criteria (Jan 2024), preference should be given to the use of the EC10 over the NOEC for the classification of substances, similarly in OECD 2006: Series on Testing and Assessment Number 54, Current approaches in the statistical analysis of ecotoxicity data: a guidance to application. ENV/JM/MONO(2006)18. (Annex I: 4.1.2.7.2, " For determining chronic aquatic toxicity for classification purposes data generated according to the standardised test methods referred to in Article 8(3) shall be accepted, as well as results obtained from other validated and internationally accepted test methods. The NOECs or other equivalent ECx (e.g. EC10) shall be used"). This is consistent across all regulatory guidance and scientific understanding, where the calculated EC10 is preferred over a NOEC. The EC10 is based on actual measured values from a dose response, whereas the NOEC is an artificial value dependent purely on the test design chosen by the testing laboratory.

However, it is also noted that the decrease by 4 % in the lowest concentration (tab. xx), although statistically significant, is actually within the natural range of variation of biological systems and therefore, is of questionable relevance for hazard classification. The variation in biological systems is the reason why according to the Guidance on the Application of the CLP Criteria (Jan 2024), if available, preference should be given to EC10 over NOEC for the classification of substances. So in fact while the data are conclusive (adequate for addressing the hazard from this endpoint) they do not suggest a requirement of classification of the substance for chronic aquatic toxicity.

In addition, the registrant proposes that a new study could be conducted (as per the requirements of CORAP(see Further regulatory process information) which would remove the issues or uncertainty raised by the DS, using a method which would increase the exposure to the organisms and also allow for analytical measurement in the process.

2. Study details and methods used

All test solutions were prepared as water accommodated fractions (WAF) with the registered substance in excess of the water solubility. As defined by ECETOC, for complex substances, the concept of a single defined water solubility has no meaning, since the total amount in solution will be the equilibrium amount of all dissolved components, which may be different from the composition of the complex substance itself and will vary depending on the amount of substance added, i.e. the loading rate. Therefore, the concept of water-accommodated fractions (WAF) or loading rates (LR) was established for insoluble substances. The idea is that of a defined amount of substance, only the soluble part ends up in solution and only this part is responsible for a possible effect on organisms. The insoluble portion of the substance is removed either by filtration or centrifugation so that physical effects can be excluded. If the water solubility is exceeded, a saturated water solution of the test substance is reached. The WAF approach has been used successfully in regulatory assessment for complex substances for many years.

The purpose of ecotoxicological studies is to evaluate the inherent toxicological property of a substance or in case of multiconstituent substances or UVCBs of the critical constituent(s) rather than physical effects e.g. caused by attachment of particles to antennas of daphnia or gills of fish or cell walls of algae. ECETOC states in the monograph: Aquatic Toxicity Testing of Sparingly Soluble, Volatile and Unstable Substances: “The driving force for uptake is the dissolved concentration in the aqueous medium. The presence of undissolved substance does not influence the relationship between dissolved concentration and uptake. Hence, there is no advantage in testing above the water solubility limit in order to assess the inherent toxicity..”(ECETOC, 1996)

For complex substances containing sparingly water-soluble compounds, it is appropriate to express exposure in terms of the overall loading rate used in the preparation of the WAF or water-soluble fraction (WSF) and toxicity in terms of LL/EL/NOEL values. (ECETOC, 1996)

In addition, the CLP guidance recommends:

” For this reason, such complex substances are usually tested as a WSF or WAF, and the L(E)C50 recorded based on the loading or nominal concentrations. Analytical support data are not normally available since the dissolved fraction will itself be a complex mixture of components. The toxicity parameter is sometimes referred to as LL50, related to the lethal loading level. This loading level from the WSF or WAF may be used directly in the classification criteria**.”**

3. OECD guideline vs ECHA expectations

For poorly water-soluble substances for which adequate analytics cannot be carried out, the OECD guidelines allow reference to and use of nominal values for toxicity assessment. This approach is viewed critically under REACH, therefore, the authorities (ECHA), with reference to OECD TG 23, request sufficient accompanying analytics to ensure that the organisms are exposed to the intended substance concentrations or their soluble fractions. Studies with poorly soluble substances, which refer to nominal concentrations, are sometimes deemed to be unreliable and in some instances new studies are requested with a requirement for analytical measurements.

In the case of this chemistry, with the request for new studies (see Annex 11 – Regulatory Activities in EU), established methods had to be modified, supplemented or newly established and validated. Furthermore, extensive efforts by the registrants were made to refine the analytics to be sufficiently sensitive to monitoring even very low concentrations. However, ultimately only limited analytical measurement was possible.

Different water- and media solubilities for the substance are reported, between >1.6 mg/L for the whole UVCB substance (BASF SE, 2010) and 4 µg/L for the main constituent DTBDA (BASF SE, 2019b). Considering the whole) group of substituted diphenyl amines (SDPA) with C9 (EC 701-385-4) and C4/8 (EC 270-128-1), a water solubility of >1.6 mg/L seems too high. The water solubility of the main constituent DTBDA was determined to be in the range of 4 µg/L (BASF SE, 2019b). In this range it is difficult to maintain concentrations but using the semi-static method it was ensured that the test solutions at the beginning of each media exchange were considered to be saturated solutions. The daily exchange of the media supported the exposure of organisms at or close to the water solubility of the test substance. Thus there can be no question of a lack of exposure potential to the organisms (considering that the purpose of the analytical measurement is to confirm this exposure), as the media was refreshed regularly and the substance is known not to degrade. The exposures were maintained, and it is the Registrants opinion that a measured value would only have shown exposure to one part of the UVCB substance, not the whole part and all soluble fractions potentially present. This would have provided an artificial result ignoring the behaviour of other moieties.

Although no analytical measurement was performed due to the UVCB character of the substance and the lack of knowledge of the toxic constituent it is not clear why the DS has chosen such a conservative evaluation by considering the toxicity of the substance to be below the lowest concentration of 0.625 mg/L (nominal loading). There is sufficient information to allow the use of the EL10 value as a point of departure.

4. Study results

The following Table 33 shows the results of the study in terms of the effects on mortality and reproduction.

Table 33: Mortality and reproduction of parent after 21 days

Test groups	Test Groups Nominal [mg/L]	Reproduction		Mortality	
		Mean Living Young per surviving adult	% effect[b]	Parent animals	% effect[b]
0	0 (control)	120.3 (9.5%[a])	–	0	–
1	0.625	**115.4***	**4.07%***	0	–
2	1.25	**110.4***	**8.23%***	0	–
3	2.5	**103.6****	**13.9%****	0	–
4	5	**100.6****	**16.4%****	0	–
5	10	**83.2****	**30.8%****	0	–

*: p≤ 0.05; **: p≤ 0.01
a: Coefficient of variation for control fecundity based on surviving parent.
b: Effect relative to control. Only calculated for statistically significant effects.

Table 34: Mean values and standard deviation based on the number of offspring per parent

Conc. [mg/L]	Vessel										MV	SD
	1	2	3	4	5	6	7	8	9	10		
0.00	135	118	123	138	115	125	118	120	114	97	120.3	11
0.625	115	113	114	115	106	107	112	132	109	136	115.9	10
1.25	116	117	117	98	109	112	112	95	113	119	110.8	8
2.5	102	117	95	94	106	108	101	110	113	101	104.7	7
5	97	107	105	101	106	94	99	108	106	92	101.5	6
10	27	123	106	60	19	116	115	101	108	78	85.3	38

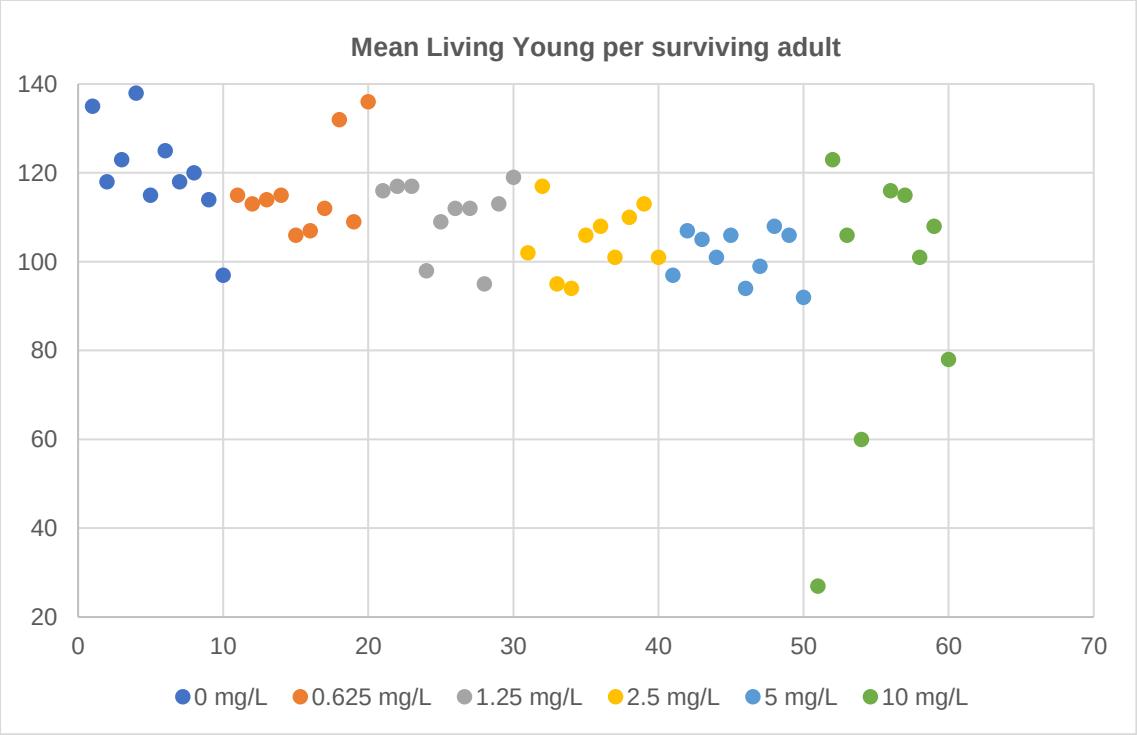
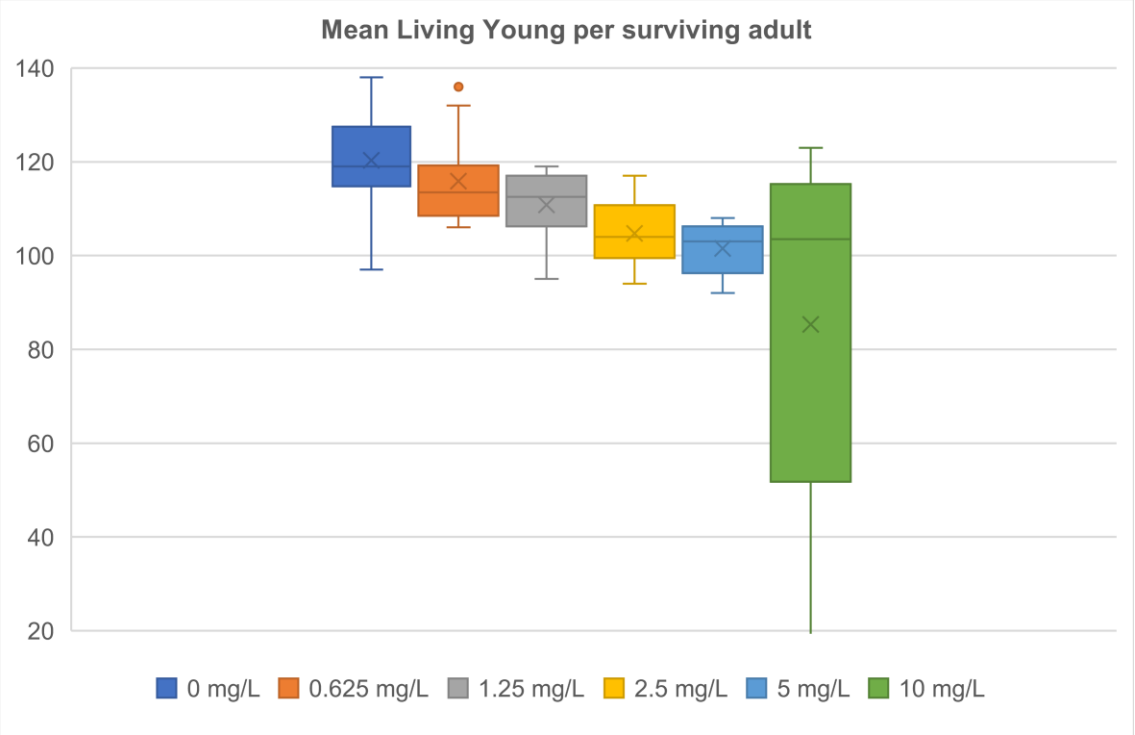


Figure 8: Number of offspring per parent: as dot plot



X in box= mean value

Figure 9: Number of offspring per parent: as box whisker plot

Biological tests aim not only to derive a level of statistical significance but also an assessment of biological relevance. This is consistent with the regulatory guidance and scientific understanding which is why EC10 is preferred over NOEC (European Chemicals Agency., 2024). While the EC10 is based on actual measured values, the NOEC is an artificial and random value in the test design chosen by the testing laboratory.

The historical controls of 2013-2017 of BASF chronic daphnia tests show a spreading in the range of 94 to 185 with a CV of 3% and 24%, respectively (Table 35). Due to the low statistical spreading (CV: 9.5%) of the controls in the chronic daphnia study, the first two following concentrations of 0.625 and 1.25 mg/l (nominal) are already significant.

Table 35: Laboratory historical control values (n = 34) of mean life offspring per parent during 2013 to 2017.

	Mean live offspring / parent	**CV**
Min	94	3%
Max	185	24%
5th Percentile	106.2	
95th Percentile	181.8	
Median	147.5	12%

Both the standard deviation and the coefficient of variance show a strong overlap of values (Table 33). This becomes particularly clear when the values are displayed in a diagram. There are practically no differences between the control points and the points of the corresponding two subsequent concentrations, with strong overlaps in the standard deviations (Table 34).

Conclusion:

From the above data it is clear that a conservative approach, such as used by the DS using values below the lowest applied concentration, is scientifically not justified. The registrants therefore fully support the testing laboratory's applied method of deriving the EC 10 value from the different effect values based on nominal exposure rates, as recommended in the OECD TG 23. Considering the EC10 of 1.68 mg/L based on nominal values is valid, there is no toxicity within the water solubility of the substance and consequently no classification and labelling is required.

Annex 11 – Regulatory Activities in EU

1. Further work proposed and other regulatory processes

Other regulatory processes are driving new work which may be highly beneficial to this CLH process. For the benefit of non-regulatory personnel (who may evaluate this information under a public consultation process) we try to put the evaluation of dossier data in a regulatory context, and to this end ECHA has published the following scheme in the Registrant's guide - How to act in substance evaluation, November 2022 (European Chemicals Agency., 2022).

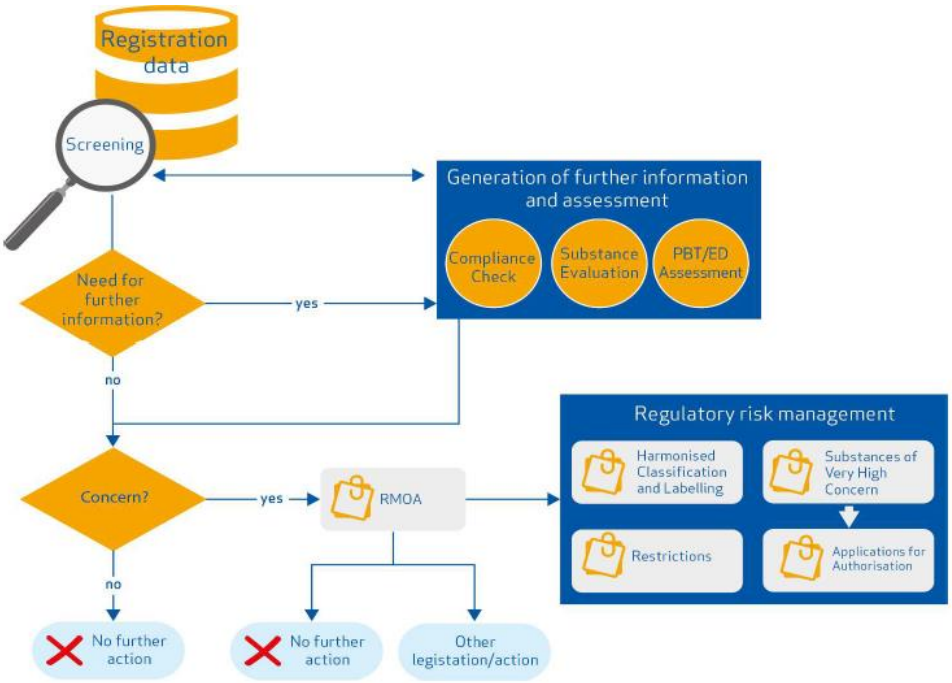


Figure 1: Substance evaluation in the regulatory context

Accordingly, data are assessed as to whether all relevant data are available and whether the data are consistent and sufficiently valid to carry out an assessment of the dossier. If ECHA and the member states in the member state committee (MSC) are of the opinion that further tests are needed, these can be requested as part of a CORAP process. Finally, the data are evaluated and in case of a concern, a harmonized classification and labeling is derived.

According to ECHA's guidelines and figure 1 "substance evaluation in the regulatory context", a CLH process carried out in parallel to a CORAP process in principle is not foreseen, if dossier data are considered as insufficiently valid or inadequate. Clearly while not foreseen, this is happening now. This substance is part of a CORAP process where ECHA has requested further studies, and these are ongoing. Even though the SeV is only based on a constituent of the UVCB registered substance (4,4'-Di-Tert-Butyldiphenylamine CAS 4627-22-9) the results of the CORAP process will affect the whole registered substance, in terms of PBT assessment. If one constituent is PBT, the entire UVCB substance will be considered to fulfill the PBT criteria.

Studies proposed under CORAP

There is currently an investigation of the persistence properties of this substance underway, and there will then follow consideration of bioaccumulation (see Annex 9 – Detailed information on bioaccumulation in fish), and finally toxicity. In ecotoxicological studies, analytical measurement has proven to be a key element in determining the toxicity of a substance, as it enables confirmation that organisms were exposed to a certain concentration and to classify the result to the toxic profile of a substance. Inadequate analytics of poorly water-soluble substances – while not reducing the value of the testing in any way – does allow room for differing interpretation of results when assessing the toxicity of a substance. This has been apparent here when the DS has adopted an unnecessarily conservative position – although this is not justified given the EC10 value which can be calculated and the valid methods used in the study to maintain exposure concentrations. However, it is fair to say that better analytical information would remove this uncertainty, and hence the Registrant intends to conduct the new study under CORAP using the passive dosing methods to maintain concentrations at a level which can be analysed.

The intention of ecotoxicological studies is to evaluate the inherent toxicological property of a substance or in case of multiconstituent substances or UVCBs of the critical constituent(s) rather than physical effects e.g. caused by attachment of particles to antennas of daphnia or gills of fish or cell walls of algae. ECETOC states in the monograph: Aquatic Toxicity Testing of Sparingly Soluble, Volatile and Unstable Substances: “The driving force for uptake is the dissolved concentration in the aqueous medium. The presence of undissolved substance does not influence the relationship between dissolved concentration and uptake. Hence, there is no advantage in testing above the water solubility limit in order to assess the inherent toxicity. This conclusion contradicts some regulatory guidelines and some reports of dose-response relationships of inherent toxicity above the solubility limit.” (ECETOC, 1996)

For complex substances, the concept of a single defined water solubility has no meaning, since the total amount in solution will be the equilibrium amount of all dissolved components, which may be different from the composition of the complex substance itself and will vary depending on the amount of substance added, i.e. the loading rate. Therefore the ‘water accommodated fraction’ (WAF) approach is of use.

For complex substances containing sparingly water soluble compounds, it is appropriate to express exposure in terms of the overall loading rate used in the preparation of the WAF or water soluble fraction (WSF) and toxicity in terms of LL/EL/NOEL values. (ECETOC, 1996)

In theory, if a UVCB or multiconstituent substance is applied in excess of the water solubility, a comparable amount of soluble fraction should be dissolved even if different loadings rates are used. Effects causing oversaturation should be avoided. However, experience shows that this principle cannot always be adhered to, therefore concentration series with poorly soluble substances do not always show a clear dose-response relationship. Without adequately sensitive analytics, it may be the case that a threshold for an effect cannot be clearly determined.

Based on recent work to establish new reliable test methods for substances like this – including analytical measurement – the Lead Registrant has proposed a sensitive and reliable method (known as “passive dosing”) to monitor the toxicity to daphnia in concentrations down to 0.001-0.01mg/l. This proposal has been endorsed by the MSC within the CORAP process. The registrants are confident they can ensure an appropriate testing procedure with reliable analytical measurement, and it should be considered that this will provide confirmation of the chronic hazard of this substance (or lack of).

It is therefore recommended that classification awaits this new information, especially given the lack of biological significance shown in the data (see above).

Annex 12 – References

BASF SE, 2021. Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene Extended One-Generation Reproduction Toxicity Study in Wistar Rats Administration via the Diet (study report No. 91R0227/13R190). BASF SE, Experimental Toxicology and Ecology, 67056 Ludwigshafen, Germany.

BASF SE, 2020a. Benzenamine, N-phenyl-, reaction products with 2,4,4-trimethylpentene Reproduction/Developmental Toxicity Screening Test in Wistar Rats Administration via the Diet (study report No. 81R0227/13C189). BASF SE, Experimental Toxicology and Ecology, 67056 Ludwigshafen, Germany.

BASF SE, 2020b. Reaction products of benzeneamine, N-phenyl with nonene (branched) Reproduction/Developmental Toxicity Screening Test in Wistar Rats Administration via the Diet and recovery period of 2 weeks (study report No. 81R0227/12C177). BASF SE, Experimental Toxicology and Ecology, 67056 Ludwigshafen, Germany.

BASF SE, 2019a. Reaction products of benzeneamine, N-phenyl with nonene (branched) Prenatal Developmental Toxicity Study in New Zealand White Rabbits Oral Administration (Gavage) (study report No. 40R0227/12R166). BASF SE Experimental Toxicology and Ecology 67056 Ludwigshafen, Germany.

BASF SE, 2019b. Water solubility of DTBDA (Analytical report No. 19L00010). BASF Competence Center Analytics, D-67056 Ludwigshafen, Germany.

BASF SE, 2010. Physico-chemical properties of test item (No. 09S01229). BASF Competence Center Analytics, D-67056 Ludwigshafen, Germany.

Beekhuijzen, M., Rijk, J.C.W., Meijer, M., de Raaf, M.A., Pelgrom, S., 2019. A critical evaluation of thyroid hormone measurements in OECD test guideline studies: Is there any added value? Reproductive Toxicology 88, 56–66. https://doi.org/10.1016/j.reprotox.2019.07.014

Beyer, B.K., Chernoff, N., Danielsson, B.R., Davis-Bruno, K., Harrouk, W., Hood, R.D., Janer, G., Liminga, U.W., Kim, J.H., Rocca, M., Rogers, J., Scialli, A.R., 2011. ILSI/HESI maternal toxicity workshop summary: maternal toxicity and its impact on study design and data interpretation. Birth Defects Research Pt B 92, 36–51. https://doi.org/10.1002/bdrb.20281

Blankenship, B., Eighmy, J.J., Hoffmann, G., Schroeder, M., Sharma, A.K., Sorden, S.D., 2016. Findings in Historical Control Harlan RCCHan™: WIST Rats from 104-week Oral Gavage Studies. Toxicol Pathol 44, 947–961. https://doi.org/10.1177/0192623316660768

Calculation of molecular dimensions related to indicators for low bioaccumulation potential, 2009. . Environment Agency, Bristol.

Cappon, G.D., Fleeman, T.L., Chapin, R.E., Hurtt, M.E., 2005. Effects of feed restriction during organogenesis on embryo-fetal development in rabbit. Birth Defect Res B 74, 424–430. https://doi.org/10.1002/bdrb.20058

Chaitra, R.S., Vani, V., Jayamma, Y., Laxmi, S.I., 2020. Estrous Cycle in Rodents: Phases, Characteristics and Neuroendocrine regulation. Karnatak University Journal of Science 51, 40–53.

Choksi, N., Jahnke, G., St Hilaire, C., Shelby, M., 2003. Role of thyroid hormones in human and laboratory animal reproductive health. Birth Defects Res. B Dev. Reprod. Toxicol. 68, 479–491.

Colicchia, M., Campagnolo, L., Baldini, E., Ulisse, S., Valensise, H., Moretti, C., 2014. Molecular basis of thyrotropin and thyroid hormone action during implantation and early development. Human Reproduction Update 20, 884–904. https://doi.org/10.1093/humupd/dmu028

Cora, M.C., Kooistra, L., Travlos, G., 2015. Vaginal Cytology of the Laboratory Rat and Mouse: Review and Criteria for the Staging of the Estrous Cycle Using Stained Vaginal Smears. Toxicol Pathol 43, 776–793. https://doi.org/10.1177/0192623315570339

Danielsson, B.R., 2013. Maternal Toxicity, in: Barrow, P.C. (Ed.), Teratogenicity Testing, Methods in Molecular Biology. Humana Press, Totowa, NJ, pp. 311–325. https://doi.org/10.1007/978-1-62703-131-8_24

Dimitrov, S.D., Dimitrova, N.C., Walker, J.D., Veith, G.D., Mekenyan, O.G., 2003. Bioconcentration potential predictions based on molecular attributes – an early warning approach for chemicals found in humans, birds, fish and wildlife. QSAR Comb. Sci. 22, 58–68. https://doi.org/10.1002/qsar.200390005

Dimitrov, S.D., Dimitrova, N.C., Walker, J.D., Veith, G.D., Mekenyan, O.G., 2002. Predicting bioconcentration factors of highly hydrophobic chemicals. Effects of molecular size. Pure and Applied Chemistry 74, 1823–1830. https://doi.org/10.1351/pac200274101823

ECETOC, 1996. ECETOC Monograph No. 26 - Aquatic Toxicity Testing of Sparingly Soluble, Volatile and Unstable Substances. European Centra for Ecotoxicology and Toxicology of Chemicals, Avenue E Van Nieuwenhuyse 4 (Bte 6), B-1160 Brussels, Belgium.

European Chemicals Agency., 2024. Guidance on the application of the CLP criteria: guidance to Regulation (EC) No 1272/2008 on classification, labelling and packaging (CLP) of substances and mixtures. Publications Office, LU.

European Chemicals Agency., 2023. Evaluating results from 55 extended one-generation reproductive toxicity studies under REACH: final report of the EOGRTS review project : March 2023. Publications Office, LU.

European Chemicals Agency., 2022. Registrant's guide: how to act in substance evaluation : November 2022. Publications Office, LU.

Ghasemi, A., Jeddi, S., Kashfi, K., 2021. The laboratory rat: age and body weight matter. EXCLI Journal; 20:Doc1431; ISSN 1611-2156. https://doi.org/10.17179/EXCLI2021-4072

Goldman, J.M., Murr, A.S., Cooper, R.L., 2007. The rodent estrous cycle: characterization of vaginal cytology and its utility in toxicological studies. Birth Defects Research Pt B 80, 84–97. https://doi.org/10.1002/bdrb.20106

Hapon, M.B., Gamarra-Luques, C., Jahn, G.A., 2010. Short term hypothyroidism affects ovarian function in the cycling rat. Reprod Biol Endocrinol 8, 14. https://doi.org/10.1186/1477-7827-8-14

Kaufmann, W., Bolon, B., Bradley, A., Butt, M., Czasch, S., Garman, R.H., George, C., Gröters, S., Krinke, G., Little, P., McKay, J., Narama, I., Rao, D., Shibutani, M., Sills, R., 2012. Proliferative and Nonproliferative Lesions of the Rat and Mouse Central and Peripheral Nervous Systems. Toxicol Pathol 40, 87S-157S. https://doi.org/10.1177/0192623312439125

Kowalczyk-Zieba, I., Staszkiewicz-Chodor, J., Boruszewska, D., Lukaszuk, K., Jaworska, J., Woclawek-Potocka, I., 2021. Hypothyroidism Affects Uterine Function via the Modulation of Prostaglandin Signaling. Animals 11, 2636. https://doi.org/10.3390/ani11092636

Lerman-Sagie, T., Ben-Sira, L., Achiron, R., Schreiber, L., Hermann, G., Lev, D., Kidron, D., Malinger, G., 2009. Thick fetal corpus callosum: an ominous sign? Ultrasound in Obstet & Gyne 34, 55–61. https://doi.org/10.1002/uog.6356

Lewandowski, T., Seeley, M., Beck, M., 2004. Interspecies differences in susceptibility to perturbation of thyroid homeostasis: a case study with perchlorate. Regulatory Toxicology and Pharmacology 39, 348–362.

Li, A.A., Makris, S.L., Marty, M.S., Strauss, V., Gilbert, M.E., Blacker, A., Zorrilla, L.M., Coder, P.S., Hannas, B., Lordi, S., Schneider, S., 2019. Practical considerations for developmental thyroid toxicity assessments: What's working, what's not, and how can we do better? Regulatory Toxicology and Pharmacology 106, 111–136. https://doi.org/10.1016/j.yrtph.2019.04.010

Li, Y., Kumazawa, T., Ishiguro, T., Kawakami, Y., Nishitani, H., Tagawa, Y., Matsumoto, Y., 2011. Hypothyroidism caused by phenobarbital affects patterns of estrous cyclicity in rats. Congenital Anomalies 51, 55–61. https://doi.org/10.1111/j.1741-4520.2011.00314.x

McClain, R.M., 1989a. The Significance of Hepatic Microsomal Enzyme Induction and Altered Thyroid Function in Rats: Implications for Thyroid Gland Neoplasia. Toxicol Pathol 17, 294–306. https://doi.org/10.1177/019262338901700206

McClain, R.M., 1989b. The Significance of Hepatic Microsomal Enzyme Induction and Altered Thyroid Function in Rats: Implications for Thyroid Gland Neoplasia. Toxicologic Pathology 17, 294–306.

Meek, M.E. (Bette), Bucher, J.R., Cohen, S.M., Dellarco, V., Hill, R.N., Lehman-McKeeman, L.D., Longfellow, D.G., Pastoor, T., Seed, J., Patton, D.E., 2003. A Framework for Human Relevance Analysis of Information on Carcinogenic Modes of Action. Critical Reviews in Toxicology 33, 591–653. https://doi.org/10.1080/713608373

Meek, M.E., et al., 2003. A Framework for Human Relevance Analysis of Information on Carcinogenic Modes of Action. Critical Reviews in Toxicology 33, 591–653.

Melching-Kollmuss, S., Fussell, K.C., Schneider, S., Buesen, R., Groeters, S., Strauss, V., van Ravenzwaay, B., 2017. Comparing effect levels of regulatory studies with endpoints derived in targeted anti-androgenic studies: example prochloraz. Arch Toxicol 91, 143–162. https://doi.org/10.1007/s00204-016-1678-y

Miyata, C., Matoba, Y., Mukumoto, M., Nakagawa, Y., Miyagawa, H., 2022. Permeability of the fish intestinal membrane to bulky chemicals. J. Pestic. Sci. 47, 86–92. https://doi.org/10.1584/jpestics.D21-055

Nichols, J.W., Bonnell, M., Dimitrov, S.D., Escher, B.I., Han, X., Kramer, N.I., 2009. Bioaccumulation Assessment Using Predictive Approaches. Integr Envir Assess & Manag 5, 577–597. https://doi.org/10.1897/IEAM_2008-088.1

OECD, 2018. OECD GUIDELINE FOR THE TESTING OF CHEMICALS No. 443 EXTENDED ONE-GENERATION REPRODUCTIVE TOXICITY STUDY.

OECD, 2000. OECD Environmental Health and Safety Publications Series on Testing and Assessment No. 19 - GUIDANCE DOCUMENT ON THE RECOGNITION, ASSESSMENT, AND USE OF CLINICAL SIGNS AS HUMANE ENDPOINTS FOR EXPERIMENTAL ANIMALS USED IN SAFETY EVALUATION. ENV/JM/MONO(2000)7.

Paccola, C.C., Resende, C.G., Stumpp, T., Miraglia, S.M., Cipriano, I., 2013. The rat estrous cycle revisited: a quantitative and qualitative analysis. Animal Reproduction 10, 677–683.

RAC-62, 2022. RAC Guidance Note: Addressing developmental neurotoxicity and neurotoxicity under the current CLP hazard classes.

Rinaldini, E., Verde Arboccó, F.C., Ezquer, M., Gamarra-Luques, C., Hapon, M.B., 2021. Effect of Thyromimetic GC-1 Selective Signaling on Reproductive and Lactational Performance in the Hypothyroid Rat. Eur Thyroid J 10, 425–433. https://doi.org/10.1159/000516432

Schoenemann, P.T., Budinger, T.F., Sarich, V.M., Wang, W.S.-Y., 2000. Brain size does not predict general cognitive ability within families. Proc. Natl. Acad. Sci. U.S.A. 97, 4932–4937. https://doi.org/10.1073/pnas.97.9.4932

Sellers, R.S., Mortan, D., Michael, B., Roome, N., Johnson, J.K., Yano, B.L., Perry, R., Schafer, K., 2007. Society of Toxicologic Pathology Position Paper: Organ Weight Recommendations for Toxicology Studies. Toxicol Pathol 35, 751–755. https://doi.org/10.1080/01926230701595300

Silva, J.F., Ocarino, N.M., Serakides, R., 2018. Thyroid hormones and female reproduction†. Biology of Reproduction. https://doi.org/10.1093/biolre/ioy115

Takayama, S., Aihara, K., Onodera, T., Akimoto, T., 1986. Antithyroid effects of propylthiouracil and sulfamonomethoxine in rats and monkeys. Toxicol. Appl. Pharmacol. 82, 191–199.

Tyl, R.W., 2012. Commentary on the Role of Maternal Toxicity on Developmental Toxicity. Birth Defects Research Pt B 95, 262–266. https://doi.org/10.1002/bdrb.21015

US-EPA, 1998. Assessment of thyroid follicular cell tumors. Downloaded from http://www.epa.gov/raf/publications/thyroid-follicular-cell-tumor.htm EPA/630/R-97/002.

Wei, Q., Fedail, J.S., Kong, L., Zheng, K., Meng, C., Fadlalla, M.B., Shi, F., 2018. Thyroid hormones alter estrous cyclicity and antioxidative status in the ovaries of rats. Animal Science Journal 89, 513–526. https://doi.org/10.1111/asj.12950

Westwood, F.R., 2008. The Female Rat Reproductive Cycle: A Practical Histological Guide to Staging. Toxicol Pathol 36, 375–384. https://doi.org/10.1177/0192623308315665

WIL Research Europe, 2014. COMBINED 28-DAY REPEATED DOSE TOXICITY STUDY WITH THE REPRODUCTION/DEVELOPMENTAL TOXICITY SCREENING TEST OF BENZENAMINE, N-PHENYL-,REACTION PRODUCTS WITH 2,4,4-TRIMETHYLPENTENE IN RATS BY ORAL GAVAGE (study report No. 85R0227/13X191). WIL Research Europe B.V., Hambakenwetering 7, 5231 DD 's-Hertogenbosch, The Netherlands.

Zoeller, R., et al., 2007. General background on the hypothalamic-pituitary-thyroid (HPT) axis. Critical Reviews in Toxicology 37, 11–53.

Zoeller, R.T., Tan, S.W., Tyl, R.W., 2007. General Background on the Hypothalamic-Pituitary-Thyroid (HPT) Axis. Critical Reviews in Toxicology 37, 11–53. https://doi.org/10.1080/10408440601123446