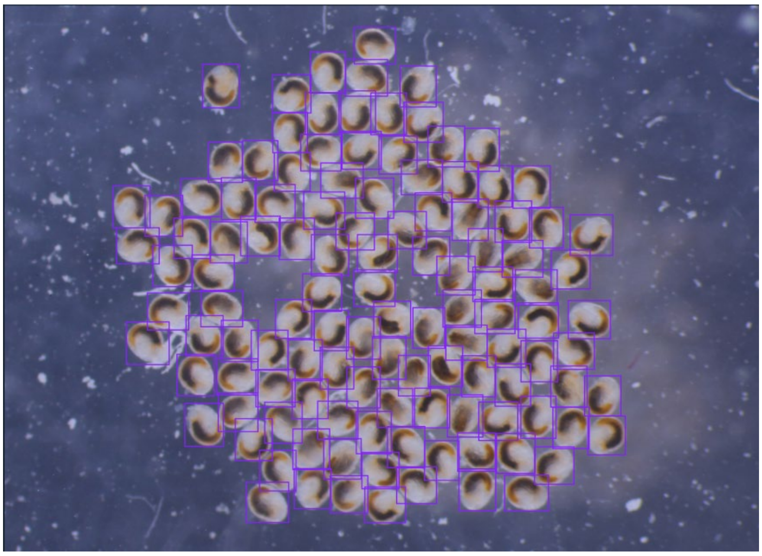
Reproductive disorders in amphipods as indicators of effects of hazardous substances in Danish waters

Scientific note from DCE – Danish Centre for Environment and Energy

Date: 27 March 2026 | **21**



AARHUS UNIVERSITY
DCE – DANISH CENTRE FOR ENVIRONMENT AND ENERGY

Data sheet

Scientific note from DCE – Danish Centre for Environment and Energy

Category:	Scientific briefing
Title:	Reproductive disorders in amphipods as indicators of effects of hazardous substances in Danish waters
Authors:	Zhanna M. Tairova, Jesper P.A. Christensen, Mihailo Azhar
Institution:	Aarhus University, DCE, Department of Ecoscience
Referees:	Iben Boutrup Kongsfelt
Quality assurance, DCE:	Iben Boutrup Kongsfelt
Linguistic QA:	Charlotte Elisabeth Kler
External comment:	The Danish Environmental Protection Agency. The comments can be found here:
Claimant:	The Danish Environmental Protection Agency, The Ministry of Environment and Gender Equality
Please cite as:	Tairova, Z.M., Christensen, J.P.A., Azhar, M. 2026. Reproductive disorders in amphipods as indicators of effects of hazardous substances in Danish waters. Aarhus University, DCE - Danish Centre for Environment and Energy, 26 s. – Scientific note no. 2026|21
	Reproduction permitted provided the source is explicitly acknowledged
Front page photo:	Zhanna M. Tairova
Number of pages:	26

Contents

Preface/Summary

This scientific briefing has been prepared in response to a request from the Danish Environmental Protection Agency, The Ministry of Environment and Gender Equality, with the assignment of assessing the utility of sub-lethal effect measurements, i.e. reproductive disorders, in amphipods *Gammarus* spp. as an indicator for toxic stress from environmentally hazardous substances in Danish coastal waters. HELCOM EG HAZ working group has recommended this indicator: "Reproductive disorders: malformed embryos of amphipods" as a supplementary indicator for toxic stress (HELCOM, 2018). Sub-lethal effects detected in amphipods were compared with sub-lethal effects detected in fish species eelpout (*Zoarces viviparus*) sampled at the same field locations as the amphipods.

This study was done by conducting field studies combined with laboratory microscopy analyses during the period of 2022-2025. Results are evaluated in relation to existing national and regional assessment criteria. It concluded that *Gammarus spp*. has the potential to be used as a bioindicator for Danish coastal waters due to the apparent sensitivity of its reproductive end-points and other qualities, such as abundance and ease of sampling.

1 Introduction

The current project aimed: 1) to evaluate the application of reproductive disorder measurements in amphipods as an indicator for toxic stress from environmentally hazardous substances in coastal waters and 2) to compare the results with corresponding indicators for toxic stress in coastal fish.

The project is related to the integrated environmental monitoring strategy employed under the National Monitoring Programme for the Aquatic and Terrestrial Environment (NOVANA) in Denmark and The Baltic Marine Environment Protection Commission – also known as the Helsinki Commission (HELCOM). In integrated environmental monitoring, chemical concentrations in marine biota and sediment are measured in parallel with biological effects/responses in biota. Biological responses are measured in selected species of organisms that can be used as bioindicators for monitored areas. These bioindicator species are selected according to several parameters, which include geographical distribution, migration patterns and sensitivity of biological response to exposure for hazardous substances.

In this project, we investigated the utility of the amphipod *Gammarus spp*. as a bioindicator of toxic effects of hazardous substances in Danish coastal waters. The occurrence of embryo aberrations in amphipods was compared with embryo aberrations in fish and eelpout (*Zoarces viviparus*) sampled at the same sites.

Previously, a Danish pilot study has suggested that the amphipod *Gammarus spp*. has the potential to be used as a bioindicator for toxic effects of hazardous substances in coastal waters (Z. M. Tairova & Strand, 2022). In the current study, this method has been explored further by collecting more empirical data at polluted and reference sites.

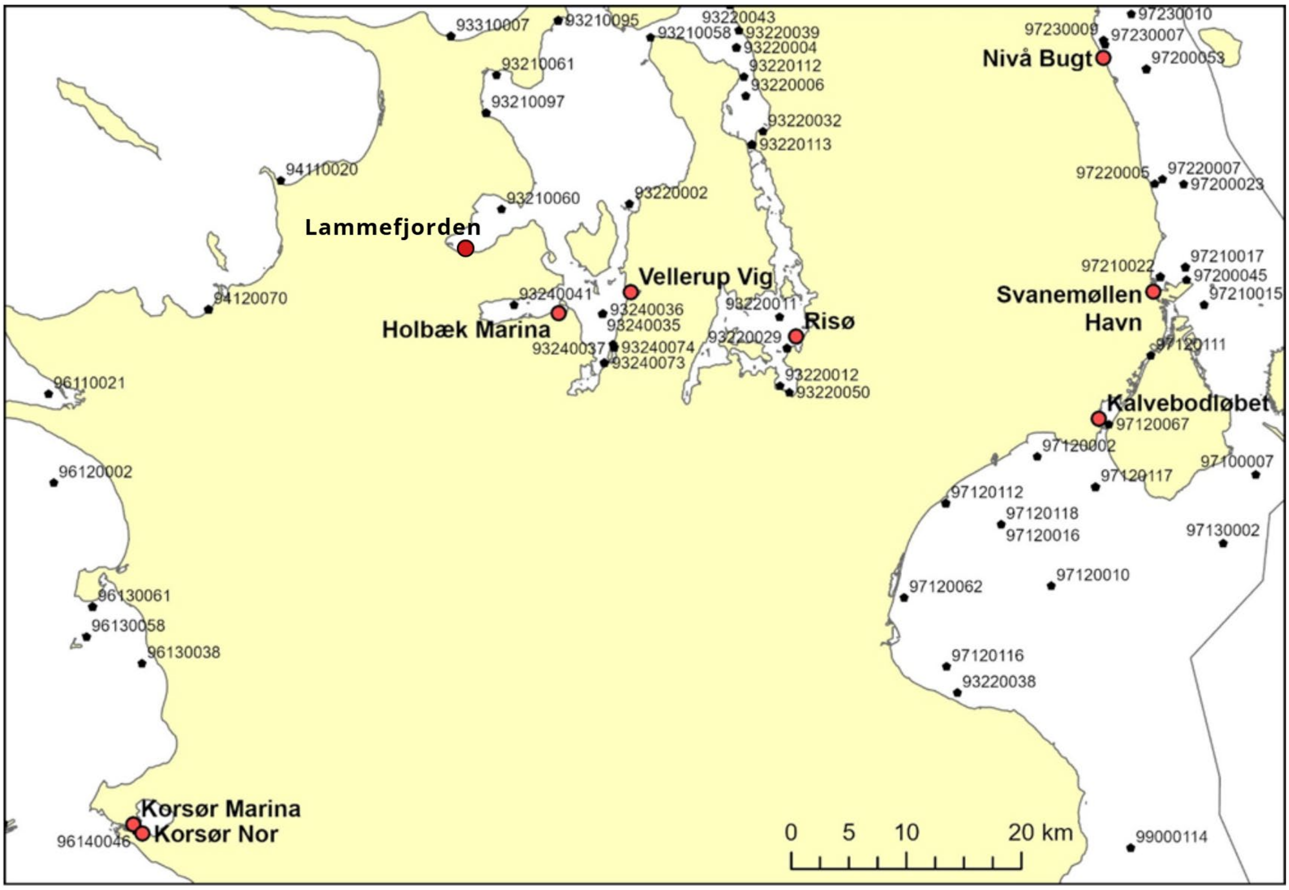
2 Materials and Methods

2.1 Amphipods

Amphipods of the genus *Gammarus spp*. were collected in May-June in the period of 2020-2025. The results from the first two years (2020–2021) were previously reported in Tairova & Strand (2022), while the results for the subsequent years (2022–2023) were presented in Tairova et al. (2024).

In 2021 and 2025, amphipod specimens were collected at Vellerup Vig, which served as a reference site due to its relatively low levels of contamination by hazardous substances. Another reference site, Nivå Bugt, was tested in 2020. The remaining stations sampled each year represented more impacted locations. The Risø site was used as an impacted site during the period 2022–2025. In 2020 and 2021, impacted stations included Holbæk Marina in both years, as well as Svanemøllen havn in 2020 and Korsør Marina in 2021. In 2024, a second impacted station, in addition to Risø, consisted of pooled samples from Korsør Nor and Korsør Marina, whereas in 2025 the additional impacted site was located at Lammefjord.

All field sites used during the 2020–2025 amphipod and fish sampling campaigns, together with the NOVANA stations (“MFS biota” — sampling sites for measurements of hazardous substances in biota), are shown in Figure 2.1.



Figur 2.1. Field sites for sampling campaigns used in the period 2020-2025 (names in bold) and NOVANA MFSbiota stations (ID number/”ObservationsstedNr”).

All the collected embryos were analyzed per hatch from individual gravid females from all stations, and the results were used to calculate three variables, i.e. the proportion of malformed embryos, the proportion of embryos with all types of aberrations and the proportion of females with more than one malformed embryo. The proportion of *malformed* embryos (i.e. embryos exhibiting structural malformations, such as fragmented internal structures, deviations in head or eye development, or damaged membranes, or undifferentiated embryos whose development ceased prior to gastrulation) was quantified separately from the proportion of embryos with *aberrations*. The latter category includes all malformed embryos as defined above, as well as underdeveloped embryos (i.e. embryos lagging by at least two developmental stages relative to the dominant stage within a brood) and dead embryos (Table 8.2).

The two variables, i.e. the proportion of malformed embryos and the proportion of females with more than one malformed embryo, are recommended for application and used as assessment criteria in supplementary indicators according to the HELCOM indicator report (HELCOM, 2018). Microscopy analyses of embryo-stages and embryo aberrations were conducted as described in Fischer 2009, with modifications (Annex: Tables 1-2).

2.2 Fish

In 2024-2025, gravid females of eelpout, *Zoarces viviparus,* were collected in October-November at NOVANA stations Kalvebodløbet in Køge Bugt and Risø in Roskilde Fjord, and the brood of each female was analysed for reproductive success. Numbers of gravid females per station/year: Kalvebodløbet in Køge Bugt in 2024 n= 24 and in 2025 n = 18; while in Risø in Roskilde Fjord the number in 2024 was n = 53 and in 2025 n = 52.

Reproductive success was measured as proportion of malformed fry (types B-G), proportion of late dead fry (type A), proportion of early dead fry (type 0) and proportion of total abnormal fry (all types) according to Technical instructions (Jakob Strand, 2013).

The eelpout (*Zoarces viviparus*) is widely used as a bioindicator for monitoring of toxic stress from hazardous substances in Danish coastal waters. This is partly because eelpout is relatively stationary in the coastal environment and because the development of embryos takes place in the female fish. Both factors make the species well-suited as a bioindicator for monitoring of environmentally hazardous substances and their effects (Jakob Strand, 2013; Z. M. Tairova et al., 2012). The hazardous substances that can affect embryo and larval development in fish include organochlorines, pesticides, polycyclic aromatic hydrocarbons (PAHs), heavy metals and organometals (Davies & Vethaak, 2012). Monitoring data for biological effect indicators can be used for both state and impact assessments and they can also be evaluated using environmentally relevant assessment criteria (ACs) developed within international scientific fora working with monitoring and assessment. The background assessment criteria (BAC) are analogous to background assessment concentrations or a natural response level. The environmental assessment criteria (EAC) represent levels of response below which unacceptable responses at higher levels, e.g. organism or population, would not be expected (Davies & Vethaak, 2012; OSPAR, 2013). These ACs have also been developed for the indicator organisms and the respective biological effect measurements deployed within the NOVANA-programme (Z. Tairova & Strand, 2021). The ACs for eelpout reported in ICES WGBEC report (2013) and OSPAR (2013) are listed in Table 2.1.

Table 2.1. Background assessment criteria (BAC) and environmental assessment criteria (EAC) for biological effect indicators in eelpout for "Reproductive success in eelpout" used in the Danish monitoring programme NOVANA (ICES WGBEC, 2013; OSPAR, 2013).

Effect indicator	BAC	EAC
Eelpout		
Mean prevalence of malformed fry (type B-G)	1%	2%
Mean prevalence of late dead fry (type A)	2%	4%
Mean prevalence of early dead fry (type 0)	2.5%	5%
Mean prevalence of total abnormal fry (all types)	5%	10%

2.3 Data analysis

Calculations of percentile, mean, median, frequency distributions and bootstrapping (100 000 runs, according to recommendation in HELCOM, 2018) were done using R© and Excel©. The map (Figure 2.1) was prepared based on stations for MFS biota extracted from OdaV2 . For data extraction R Studio 2023.06.1 and for Map production ArcGIS Pro 3.2.0 was used.

3 Results and Discussion

3.1 Reproductive disorders in amphipods *Gammarus spp.*

The embryo malformation indicator for amphipods is a multimetric indicator based on two variables measured in the sampled population: (1) the proportion of malformed embryos and (2) the proportion of females with more than one malformed embryo. In order to achieve a "good status" for an area under investigation, both variables must be below or equal to their respective threshold values (HELCOM, 2018, 2023). A third variable "proportion of embryos with all types of aberrations" was measured in this study in order to make a closer comparison with the similar variable in fish, i.e. "mean prevalence of total abnormal fry (all types)" (Table 3.4). All three variables are measured in the same pool of field-collected gravid females. As an addition to a variable "the proportion females with more than one malformation", a variable "the proportion of females with more than one aberration" was measured to explore the various variables for this indicator.

The results of reproductive disorders – mean proportions of all types of aberrations and malformations - and females with more than one type of aberration and malformation from the period 2020-2025 are shown in Figures 3.1 and 3.2.

Results from the period 2020–2023 were presented in an earlier technical report (Z. M. Tairova et al., 2024; Z. M. Tairova & Strand, 2022). In the present scientific briefing, these data are combined with data from 2024–2025 and used for the calculation of threshold values.

During 2022–2023, all amphipods underwent taxonomic analysis. The species composition of gammarids was: 64.5% *Gammarus locusta*, 14.5% *Gammarus zaddachi*, 14.5% *Gammarus tigrinus*, 4.3% *Gammarus oceanicus* and 2.2% *Gammarus salinus* (n=138). The taxonomy analysis of amphipods from two stations in 2020 (Nivå Bugt (n=15) and Holbæk Marina (n=28)) demonstrated that 88.4% were *Gammarus locusta*, while 11.6% were *Gammarus zaddachi* (Z. M. Tairova & Strand, 2022). No additional taxonomic analyses were conducted in 2024–2025.

In the previous study of the results from the period of 2020-2021 (Tairova & Strand, 2022), the mean and median values and the 90th percentile (which represents the threshold value) were calculated. For this calculation, two frequency distributions of variable of proportion of malformed embryos per brood for reference and impacted areas were used (Table 3.1). The 90th percentile represents the value below which 90% of all observations fall, meaning only the highest 10% of the measured embryo malformation rates lie above this point. In practical terms, it marks the upper boundary of what can still be considered part of the natural variation in reference conditions, and values exceeding this threshold are therefore taken to indicate an abnormal or impacted biological response.

Tabel 3.1. Threshold (90th percentiles) and mean and median values for the gammaridean amphipods *Gammarus spp.* (from Tairova & Strand, 2022).

Assessment criteria	Mean	Median	90% percentile
Proportion of malformed embryos, reference stations	0.014	0	0.041
Proportion of malformed embryos, impacted stations	0.11	0.018	0.4

Both distributions for the reference stations (Nivå Bugt and Vellerup Vig) and for all impacted stations were not normally distributed, therefore, median values (that are non-parametric) were calculated in addition to the mean values in order to compare impacted and reference stations and to compare to thresholds for gammaridean amphipods in the HELCOM supplementary indicator report (Table 3.2). Median values are lower than mean values due to not normal distribution. The median describes the typical state, but in zero-inflated data it may not capture the severity of effects present in the non-zero tail, the mean describes the average magnitude of the effect across all samples, but is sensitive to zeros and outliers.

Tabel 3.2. Secondary thresholds* for the gammaridean amphipods *Gmelinoides fasciatus*, *Pontogammarus robustoides* and *Gammarus tigrinus* (based on Gulf of Finland monitoring data, Russia) (HELCOM, 2018).

Assessment criteria	Mean	BAC	EAC	Threshold value
Proportion of malformed embryos	0.02	<0.05	>0.05	0.05
Proportion of females with >1 malformed embryo	0.15	<0.2	>0.2	0.2

* - In areas where *Monoporeia affinis*, the species used for monitoring in the Swedish National Marine Monitoring Program, does not occur naturally or is found sporadically and/or at low abundances, other amphipods with a similar life cycle and reproduction biology can be used to derive the embryo malformation indicator to establish the, so-called, "secondary thresholds" for other amphipod species belonging to gammarids (HELCOM 2018).

Resulting mean values and threshold values (Table 3.1) for the variable "proportion of malformed embryos" from baseline data (i.e. based on natural variation from reference stations, i.e. Vellerup Vig and Nivå Bugt), were below the mean values and threshold values presented in the HELCOM indicator report (see Table 3.2). Median values were lower than mean values due to not normal distribution. Mean values and threshold values from the distribution based on data from impacted stations (Table 3.1) were above the mean values and threshold values presented in the HELCOM indicator report (see Table 3.2).

For 2020–2021, the mean, median, and 90th percentile were calculated from raw data on individual broods. Since 2022, the larger dataset has allowed the mean and 90th percentile to be calculated using two different approaches:

- The values of mean and the 90th percentile were calculated for reference and impacted sites per station/per year, and then the median value for the area, i.e. "reference area" and "impacted area", were calculated based on the values per station/per year.
- The values of mean and the 90th percentile were calculated for reference and impacted areas in the same way as the abovementioned values, while using the results of distributions from bootstrapping. The application of bootstrapping to derive distributions on data for reproductive disorders is suggested in the HELCOM indicator report (HELCOM, 2018). Bootstrapping data treatment is recommended for distribution normalization with a limited number of data-points.

The results for the period of 2020-2023 are presented in Table 3.3.

Tabel 3.3. Median of mean values (means calculated per station) and thresholds (the 90th percentile) with and without the bootstrapping of the data for the period of 2020-2023.

Assessment criteria	Median of means	Median 90% percentile	Median of Means (bootstrapping)	90% percentile (bootstrapping)
Proportion of malformed embryos, reference stations	0.018	0.027	0.019	0.028
Proportion of embryos with all types of aberrations, reference stations	0.018	0.0299	0.019	0.028
Proportion of females with >1 malformed embryo, reference stations	0.2		0.13	0.23

The calculated threshold values (Table 3.3) are comparable with the threshold values calculated for gammaridean species in the HELCOM supplementary indicator report (Table 3.2). The threshold values for two variables, proportions of malformed and aberrated embryos (both at 0.028), are lower than the threshold value for "proportion of malformed embryos" presented in the HELCOM supplementary indicator report (Table 3.2). The threshold value for the variable, "proportion of females with more than one malformed embryo" is slightly higher, i.e. 0.23, than the threshold value for the same variable presented in the HELCOM supplementary indicator report (Table 3.2).

The results for the period of 2020-2025 are presented in Table 3.4.

Tabel 3.4. Median of mean values (means calculated per station) and thresholds (the 90th percentile) with and without the bootstrapping of the data for the period of 2020-2025.

Assessment criteria	Median of Means	Median 90% percentile	Median of Means (bootstrapping)	90% percentile (bootstrapping)
Proportion of malformed embryos, reference stations	0.022	0.037	0.020	0.031
Proportion of embryos with all types of aberrations, reference stations	0.023	0.042	0.022	0.031
Proportion of females with >1 malformed embryo, reference stations	0.18		0.17	0.28

After inclusion of additional data from 2024–2025 (Table 3.4), the threshold values increased slightly compared with previous estimates, but remained comparable and close to the thresholds for 2020–2023 (Table 3.3) and to those calculated for gammaridean species in the HELCOM supplementary indicator report (Table 3.2).

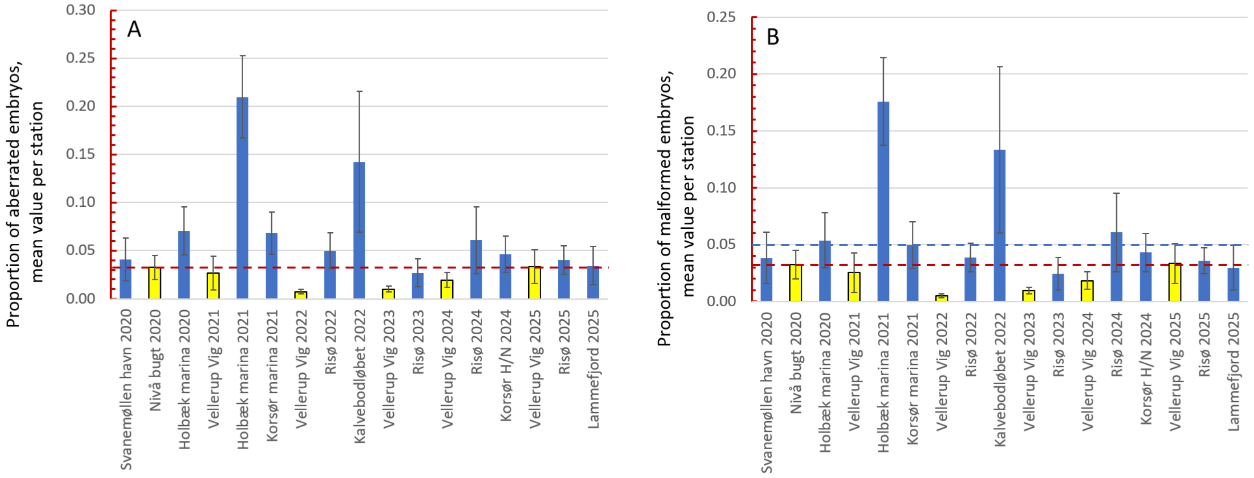


Figure 3.1. Mean proportion of all types of aberrations in embryos (A): underdeveloped, i.e. embryos at two developmental stages earlier than main stage per brood; membrane aberrations; malformations; undifferentiated embryos whose development halted before gastrulation and dead embryos in the brood. **Mean proportion of all types of malformed embryos (B)**: embryos with malformations, membrane damaged embryos and undifferentiated embryos. Vellerup Vig og Nivå Bugt (yellow columns) are the reference stations. Mean proportions were calculated by dividing the sum of all aberrations/malformations by the total number of embryos per brood). Error bars representing the standard error. **Red line**: threshold value, corresponding to 90th percentile after bootstrapping found in this study. **Blue line** threshold value from HELCOM indicator report (2018).

The results demonstrate a higher response of both variables, mean proportion of aberrations and malformations in amphipods from the stations that are considered to be impacted compared to the reference stations (Figure 3.1). The separation of the means relative to their standard errors suggests a consistent difference between impacted and reference stations. Response levels for both variables at the impacted sites (Risø and Lammefjord) in 2025 were generally lower than responses at impacted sites (except at Risø in 2023) in previous years, likely reflecting annual variability.

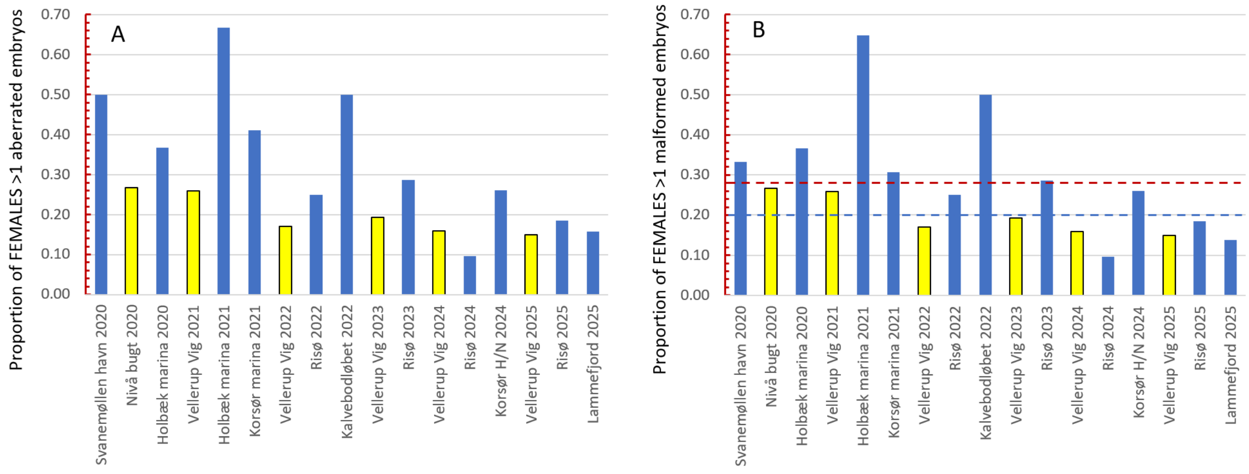


Figure 3.2. Proportion of females with more than one aberrated embryo (A) and **with more than one malformed embryo (B)**. Proportions were calculated by dividing the sum of all females with more than one aberrated/malformed embryo by the total number of embryos per brood. Vellerup Vig and Nivå Bugt (yellow columns) are the reference stations. **Red line**: threshold value, corresponding to 90th percentile after bootstrapping found in this study. **Blue line**: threshold value from HELCOM indicator report (2018).

The results show generally higher responses for both variables — the proportion of females with more than one aberration and with more than one malformation — in amphipods from impacted stations compared with reference stations. However, fewer stations exceed the threshold values following the inclusion of 2024–2025 data, which raised the threshold from 0.23 to 0.28 for the variable "Proportion of females with more than one malformed embryo". As observed for the proportions of aberrated and malformed embryos, response levels for both variables — the proportion of females with more than one aberration and with more than one malformation — at Risø and Lammefjord in the two most recent years was lower than previously recorded at Risø and generally lower than levels observed at other impacted stations in earlier years.

3.2 Reproductive success in eelpout *Zoarces viviparus*.

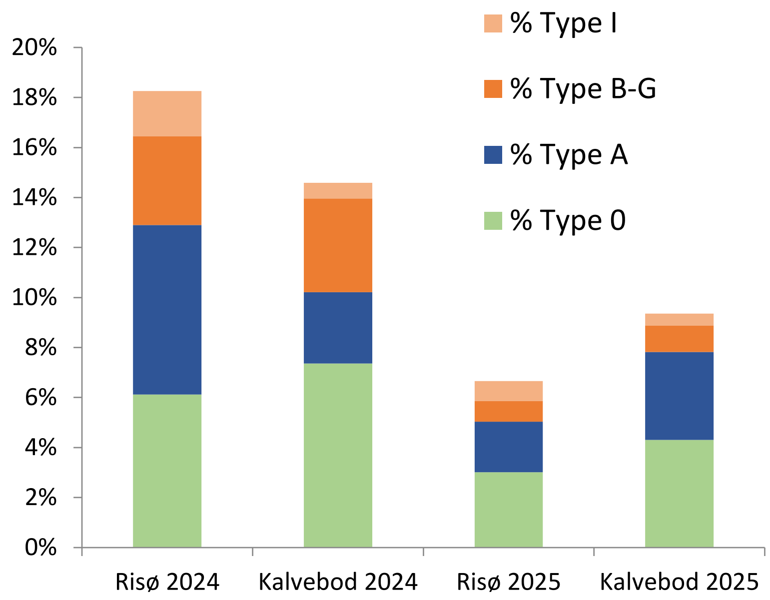
All the collected embryos were analyzed per hatch from individual gravid females from both stations. The results were used to calculate different types of abnormally developed embryos in broods of eelpout (Table 3.5).

Tabel 3.5. Proportions of abnormally developed embryos in broods of eelpout from Roskilde Fjord (Risø) and Køge Bugt (Kalvebodløbet) in 2024 and 2025.

	Risø 2024	Kalvebod 2024	Risø 2025	Kalvebod 2025
Mean prevalence of early dead fry (type 0)	6.1%	7.4%	3.0%	4.3%
Mean prevalence of late dead fry (type A)	6.8%	2.9%	2.0%	3.5%
Mean prevalence of malformed fry (type B-G)	3.6%	3.7%	0.8%	1.1%
Dwarfs (significantly smaller embryos than the mean of the brood, Type I)	1.8%	0.6%	0.8%	0.5%
Mean prevalence of total abnormal fry (all types)	17.2%	13.7%	6.5%	9.2%

Comparison of the values for different types of abnormal development in eelpout embryo with existing assessment criteria (Table 2.1) showed that several types of abnormal development were over BAC value at both stations in 2025, while all the types of abnormal development, including malformations (type B-G), were over EAC values at both stations in 2024, except the prevalence of late dead fry (type A), which was above BAC value (Table 3.5). As in the earlier study (2024), we assumed the same potential causes for the increase in abnormalities of all types observed in eelpout in 2024 as we observed in 2023 (Z. M. Tairova et al., 2024). A type 0 (early dead embryo) is also a sensitive indicator for low oxygen levels. Higher numbers in 2023 could be a result of higher temperature and occurrence of oxygen depletion in Danish coastal waters in the summer-fall 2024 (Hjorth & Josefson, 2010; J Strand et al., 2004). Prevalence of malformations (types B-G) is a sensitive indicator for pollution with hazardous substances. In 2023, the levels of hazardous substances might be increased in Danish coastal waters, including at these sites, due to the higher amount of rainfall and, consequently, increase in the inflow of the hazardous substances from land this year (DMI, 2024).

Figure 3.3. Proportions of abnormally developed embryos in broods of eelpout from Roskilde Fjord (Risø) and Køge Bugt (Kalvebodløbet) in 2024 and 2025, presented as mean values for different categories of aberrated development.



A comparison of reproductive disorder results in amphipods and fish revealed elevated responses in eelpout at two NOVANA stations, Risø in Roskilde Fjord and Kalvebodløbet in Køge Bugt (Figure 3.3). Similar elevations in responses relative to the reference site Vellerup Vig were observed in amphipods from Kalvebodløbet (KL) in 2022 and from Roskilde Fjord (Risø) in 2024 (Figures 3.1 and 3.2).

3.3 Chemical status assessment.

The assessment of chemical pressures at the study sites was based on the presence of local point sources, such as intense marine traffic in marinas, proximity to wastewater treatment plant discharges (e.g. Lynetteholm at Svanemøllen havn) and runoff from channels connected to agricultural catchments, as observed at Lammefjord. In addition, the assessment of the study sites was informed by previous results on biological responses from indicators used within the NOVANA monitoring programme, e.g. the Risø site in Roskilde Fjord (Hansen, 2012).

All study sites are located in coastal areas and, therefore, no true reference (uncontaminated) sites are available. Table 3.5 provides an overview of the chemical status assessment, including the substances responsible for failure to achieve chemical status compliance in each water district, as well as the total number of priority substances assessed against the environmental quality standards (the Danish national MKK). No data are available for Vellerup Vig. However, the low level of marine traffic and the absence of marinas, industry and urban areas suggest that this site is likely less affected by potential anthropogenic contamination sources than the other sites.

Earlier studies have shown strong correlations between occurrence of embryo aberrations in marine amphipods *Monoporeia affinis* to exposure to elevated levels of metals, PAHs and PCBs (Kolesova et al., 2024; Löf et al., 2016). All three contaminant groups are represented in Table 3.6 by substances occurring at concentrations above the environmental quality standards (EQS).

Tabel 3.6. The table is based on data from the Danish Environmental Protection Agency's status assessment conducted in connection with the revision of the 2021–2027 Chemical Status Management Plans (MiljøGis, 2026).

EU Water District (Vandområde) ID	Water District Name/ Study sites	Name of the priority substance causing failure to achieve "Chemical Status" compliance	Number of EQS (DK) checked
DKCOAST6	”Nordlige Øresund”/ Nivå bugt, Kalvebodløbet	Anthracene, BDE (sum), Benzo(a)pyrene, Lead, Cadmium, Mercury, Nickel	17
DKCOAST2	”Roskilde Fjord, indre”/ Risø	Cadmium, Mercury	17
DKCOAST206	”Smålandsfarvandet, åbne del”/ Korsør harbour	Benz(a)pyren, Lead, Cadmium, Mercury, Nickel	17
DKCOAST165	”Isefjord, indre”/ Vellerup Vig, Lammefjord, Holbæk marina	BDE (sum), Benz(a)pyren, Lead, Cadmium, Mercury, Nickel, Tributyltin	16
DKCOAST16	Korsør Nor	Lead, Cadmium, Mercury, Tributyltin	16

3.4 Integrating AI into workflows.

An initial pilot study was conducted to evaluate the use of AI to accelerate microscopy workflows for detecting reproductive disorders in amphipods. Detecting abnormalities can be separated into four steps (Figure 3.4): locating the embryos in the image, identifying the life stage they are at, identifying whether an abnormality is present and, lastly, classifying the abnormality into one of the five classes: Dead, Membrane aberration, Malformation, Undifferentiated (see Figure 3.5).

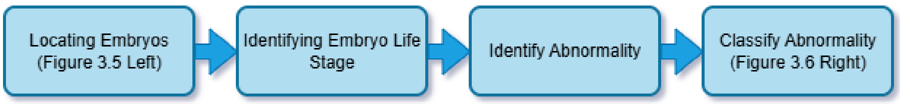


Figure 3.4. Steps for detecting and classifying embryo normality and abnormality.

The dataset was limited to 65 sites from 2025; a total of 5604 annotations were used, including variations with image saturation and brightness augmentations. Evaluation was performed using an 80:20 training-testing split; 80% of annotations were used to train the model, and the remaining 20% were used as a test set unseen during training. Localization was performed using the RT-DETR model producing bounding boxes around embryos in the image, which were then fed to a separate classifier to determine whether embryos were abnormal or normal. For the test dataset, a detection precision (mAP) of 0.98 and a classification performance with a weighted F1 score of 0.91 were obtained (closer to 1.0 is better). Performance on "Malformation" (n=18) and "Undifferentiated" (n=27) aberrations was poor at an F1 score of 0.20 and 0.1, respectively, due to their low number of samples, while "Membrane" aberrations(n=87) scored well at 0.82.

Figure 3.5. Steps 1, 3 and 4 of the embryo detection workflow. In Step 1, embryos are first detected (left), then classified separately (right) based on their aberrations in Steps 3 and 4. For high resolution embryo microscope images, the final image is stitched together before inference by the model.

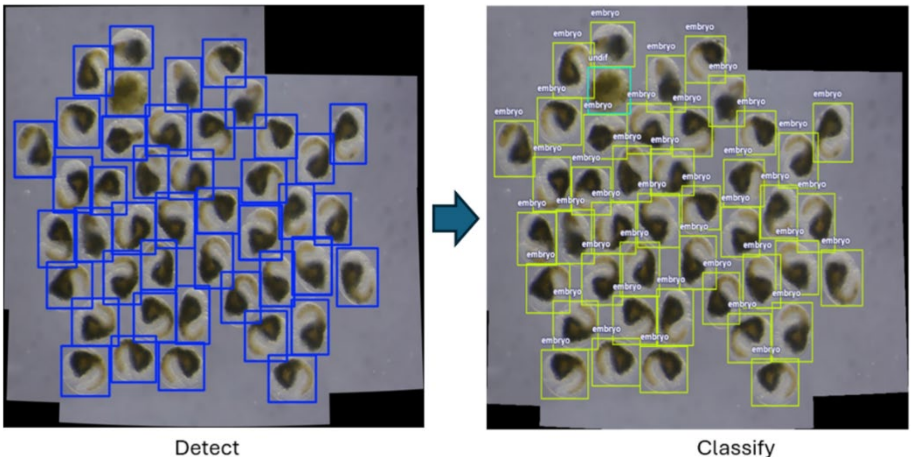
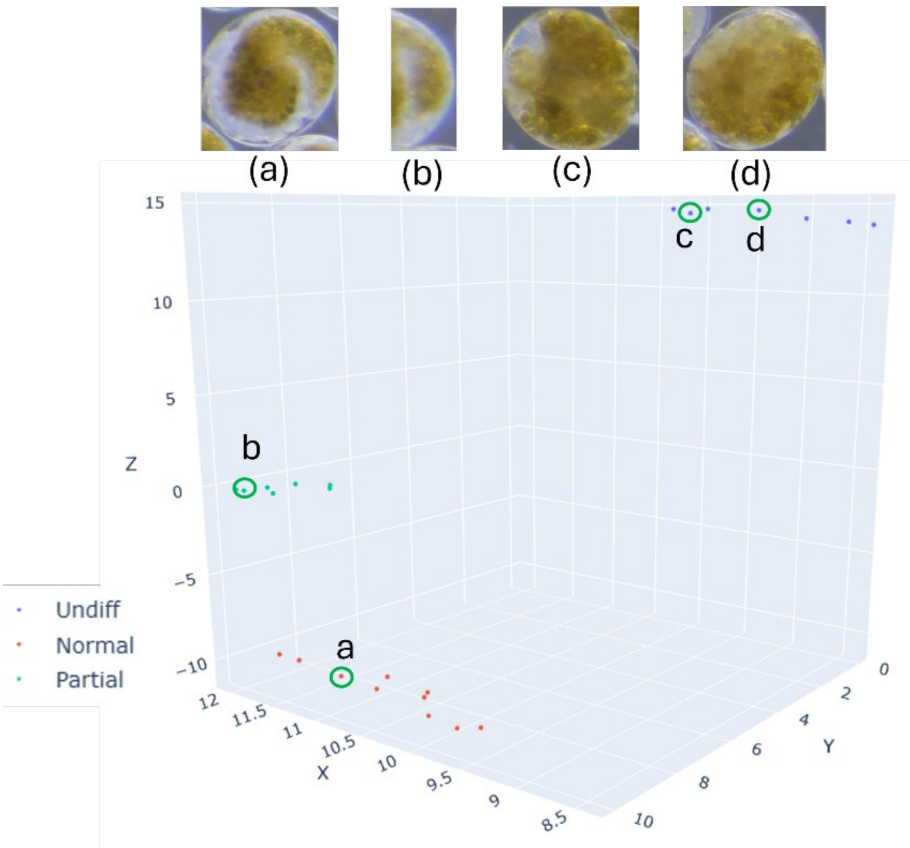


Figure 3.6. Clustering of amphipod embryos using the neural network embeddings after localization. Three classes are present in the sample: (a) regular embryos, (b) partially visible embryos, and (c) & (d) undifferentiated embryos.



As part of this pilot, we examined the possibility of leveraging large neural networks pretrained on different data for clustering of the detected embryos without further network training (unsupervised clustering). Detected embryos were passed through a large neural network that extracted a 1024-dimensional vector encoding the visual appearance of the detected embryos in the image and then clustered them based on similarity. Figure 3.6 shows a 3D projected view of the clustering; the points represent the embryos, and the color denotes their class. Aberrations, such as undifferentiated embryos, are clustered separately from the regular embryos and partially visible embryos (those at the edges of microscope images). In both cases, detection and unsupervised clustering, the potential for automating the evaluation of amphipod embryos is strong, but more data is needed to draw concrete conclusions on the robustness of such a tool in these workflows.

4 Conclusions

Reproductive disorders are a HELCOM recommended indicator for biological effect monitoring of environmental pollution with hazardous substances. The results of this study indicate that reproductive disorders in the amphipods *Gammarus spp.* are a sensitive endpoint for detecting sub-lethal toxicity of hazardous substances in coastal waters. Therefore, this indicator can be viewed as a potential indicator for Danish waters for environmental assessment of areas impacted with chemical pollution. Reproductive disorders as an indicator for environmental pollution have the advantage of linking the sub-lethal response to a higher organizational level in the ecosystem – observed effects in reproduction in an indicator organism can be linked to potential effects in populations, which increases the ecological relevance of this indicator.

It was possible to derive a threshold value consistent with threshold values for gammarids presented in HELCOM (2018) based on the data collected in the period 2020-2025. Results from a pilot study and the present study allowed calculating the threshold values (Table 3.3 and 3.4), which are comparable with the threshold values calculated for gammaridean species in the HELCOM supplementary indicator report (Table 3.2). The threshold values for the two indicators, proportions of malformed and aberrated embryos, both at 0.031, are lower than the threshold value for "proportion of malformed embryos", at 0.05, presented in the HELCOM supplementary indicator report (Table 3.2). The threshold value for the indicator, "proportion of females with more than one malformed embryo" is higher, i.e. 0.28, than the threshold value for the same indicator presented in the HELCOM supplementary indicator report, i.e. 0.2. Although the interspecies differences between the threshold values for three gammaridean species from the Gulf of Finland (Table 3.2) and *Gammarus spp*. from Danish waters could be expected, the threshold values observed in this study are comparable to those presented in the HELCOM supplementary indicator report. These threshold values can be applied for assessment of environmental conditions, which can affect the reproduction in coastal amphipods.

Elevated levels of reproductive responses in amphipods from Kalvebodløbet in Køge Bugt and Risø in Roskilde Fjord are consistent with elevated levels in reproductive responses observed in fish studied in the same two NOVANA stations.

Gammarus spp. in Danish coastal waters fulfills the criteria as a bioindicator organism due to such qualities as:

- Ecological relevance for Danish coastal areas;
- Relative abundance and ease of sampling in most of the areas that were studied in this project;
- Sensitivity of this indicator was comparable to the HELCOM supplementary indicator thresholds;
- Sensitivity of this indicator was comparable to routinely used reproductive responses in fish in Danish monitoring programme.

Due to the abovementioned reasons, it is possible to conclude that *Gammarus spp*. has the potential to be used as a bioindicator for environmental pollution in Danish coastal waters.

5 Suggestions for future studies

Based on the experience with establishing indicators of biological responses to pollution with hazardous substances in Denmark and other countries, we recommend conducting further baseline studies comprised of repetitive and seasonal field sampling of amphipods at sites with various degrees of pollution to determine spatial and temporal variability of the suggested indicator – reproductive disorders in *Gammarus spp*.

For future studies, chemical analyses should be conducted for the same period as measurements of this biological effect indicator in relevant environmental matrices (e.g. sediment, water, bivalves), as corresponding chemical data are not always available in monitoring databases for all study sites. In addition, laboratory studies incorporating complementary biomarkers — such as enzymatic activity — are recommended, as these can provide biological evidence of chemical pollution effects in amphipods and help evaluate pollutant specificity. Furthermore, the pilot study demonstrates that AI-based image analysis can substantially accelerate microscopy workflows for detecting embryo abnormalities and holds promise for automating the assessment of reproductive disorders; however, expanding baseline datasets through repeated and seasonal sampling across sites will be essential to support robust model training and validation, particularly for underrepresented abnormality classes.

6 Acknowledgments

The authors gratefully acknowledge Pia Hyrland (Landsforeningen Levende Hav) and Gunni Ærtebjerg (NGO Red Isefjord) for their valuable assistance in collecting amphipod specimens during the 2025 fieldwork.

7 References

Davies, I. M., & Vethaak, D. (2012). Integrated marine environmental monitoring of chemicals and their effects. ICES cooperative research report Nr. 315.

DMI. (2024). Sammendrag af året 2024.

Hansen, J. W. (2012). MARINE OMRÅDER 2011. NOVANA. Videnskabelig rapport fra DCE – Nationalt Center for Miljø og Energi nr. 34. (Issue 34).

HELCOM. (2018). Reproductive disorders: malformed embryos of amphipods. HELCOM supplementary indicator report. (Issue July).

HELCOM. (2023). Reproductive disorders: malformed embryos of amphipods. HELCOM supplementary indicator report.

Hjorth, M., & Josefson, A. B. (2010). Marine områder 2008. NOVANA. Tilstand og udvikling i miljø- og naturkvaliteten. Danmarks Miljøundersøgelser, Aarhus Universitet. Faglig rapport fra DMU nr. 760 (Issue 760). http://www.dmu.dk/Pub/FR760.pdf

ICES WGBEC. (2012). Report of the Working Group on Biological Effects of Contaminants (WGBEC). In ICES WGBEC Report (Issues 12-16 March, Porto, Portugal, ICES CM 2012/SSGHIE:04.).

Kolesova, N., Sildever, S., Strode, E., Berezina, N., Sundelin, B., Lips, I., Kuprijanov, I., Buschmann, F., & Gorokhova, E. (2024). Linking contaminant exposure to embryo aberrations in sediment-dwelling amphipods: a multi-basin field study in the Baltic Sea. Ecological Indicators, 160(February), 111837. https://doi.org/10.1016/j.ecolind.2024.111837

Löf, M., Sundelin, B., Bandh, C., & Gorokhova, E. (2016). Embryo aberrations in the amphipod Monoporeia affinis as indicators of toxic pollutants in sediments: A field evaluation. Ecological Indicators, 60, 18–30. https://doi.org/10.1016/j.ecolind.2015.05.058

OSPAR. (2013). Background document and technical annexes for biological effects monitoring, Update 2013. Monitoring and Assessment Series. Publication number: 589/2013.

Strand, J, Andersen, L., Dahllöf, I., & Korsgaard, B. (2004). Impaired larval development in broods of eelpout (Zoarces viviparus) in Danish coastal waters. Fish Physiology and Biochemistry, 30(2004), 37–46.

Strand, Jakob. (2013). Biologisk effektmonitering i fisk. DCE Teknisk anvisning. In DCE - Danish Centre for Environment and Energy (Issue TA. nr. M26). https://bios.au.dk/fileadmin/bioscience/Fagdatacentre/MarintFagdatacenter/TekniskeAnvisninger2011_2015/TA_M26_Biologisk_effektmonitering_i_fisk__ver_1_.pdf

Tairova, Z. M., Christensen, J. P. A., & Strand, J. (2024). Reproductive disorders in amphipods as indicators of effects of hazardous substances in Danish waters. Aarhus University, DCE - Danish Centre for Environment and Energy, Scientific note no. 2024|10 (Issue March).

Tairova, Z. M., & Strand, J. (2022). Biological effect measurements in Gammarus spp. and Corophium volutator as indicators of toxic effects of hazardous substances in Danish coastal waters. Aarhus University, DCE – Danish Centre for Environment and Energy, 24 pp. Technical Report No. 237. In Aarhus University, DCE – Danish Centre for Environment and Energy © (Issue 237). https://doi.org/10.1057/978-1-349-96042-2_1651

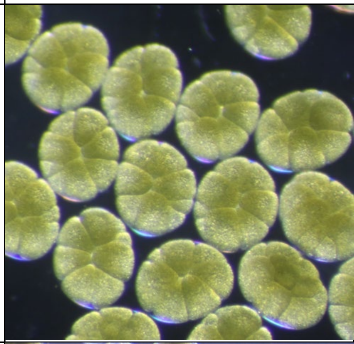
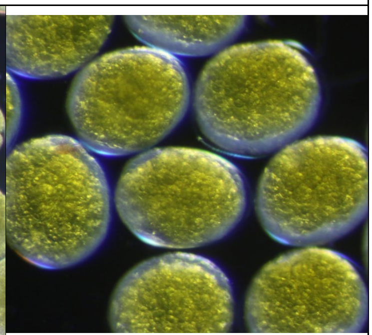
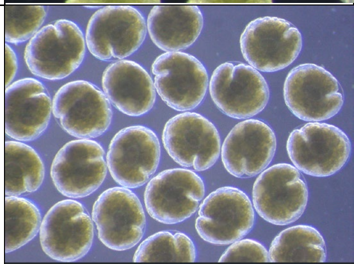
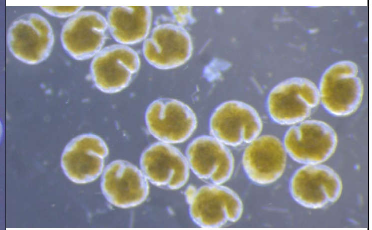
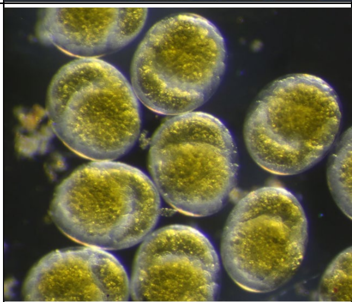
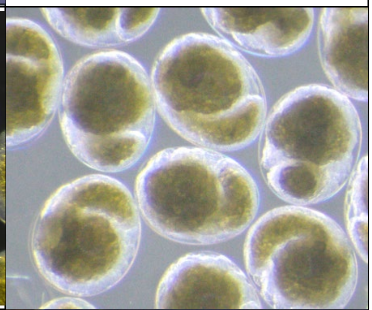
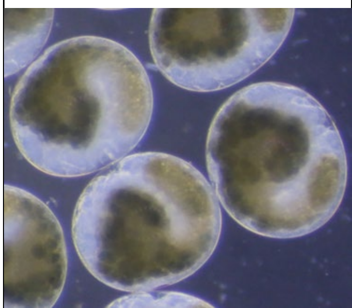
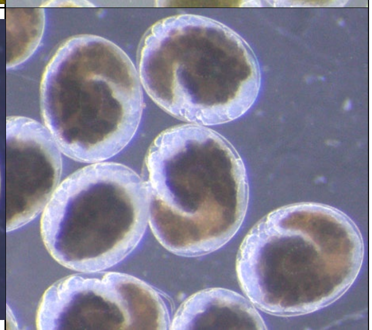
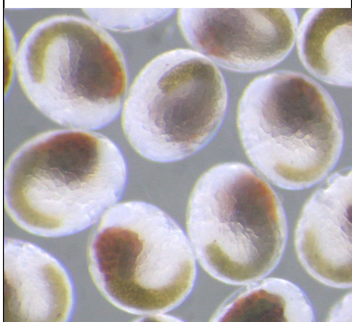
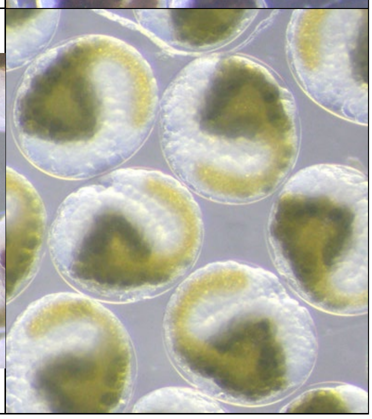
Tairova, Z. M., Strand, J., Chevalier, J., & Andersen, O. (2012). PAH biomarkers in common eelpout (Zoarces viviparus) from Danish waters. Marine Environmental Research, 75, 45–53. https://doi.org/10.1016/j.marenvres.2011.09.005

Tairova, Z., & Strand, J. (2021). Status for biological effect indicators monitored in Danish marine ecosystems. November, 34. https://dce.au.dk/fileadmin/dce.au.dk/Udgivelser/Notater_2021/N2021_59.pdf

MiljøGIS, 2026. MiljøGIS for genbesøg af vandområdeplaner 2021-2027", https://miljoegis.mim.dk/spatialmap?profile=vandrammedirektiv3genbe-soeg2024", Retrieved 04 february, 2026.

8 Annex

Table 8.1. Embryo-stages for *Gammarus* spp. (Amphipoda)

Stage	Distinguishing marks	Description	Illustrations
1	Cluster of cells, single cells more visible	Cluster of cells with no clear structure, single cells more visible	 
2	Small fracture/fractures in cell structure	Small fracture/fractures in cell structure.	 
2B	Fracture widens, shape is like a "keyhole".	Fracture is wider; "head" and "tail" are shaped.	 
3	"Comma"-shaped, The "comma" takes up most of the space in the embryo. The "tail" part is made of different type of cells	"Comma"-shaped. Legs are not clearly visible yet, but more like white 'foam'. The "comma" takes up most of the space in the embryo.	 
4	"Comma"-shaped. The comma takes up less space and the legs more in the embryo	Still "Comma"-shaped, legs are now clearly visible and look more like legs. The "comma" takes up less space and the legs more in the embryo.	 

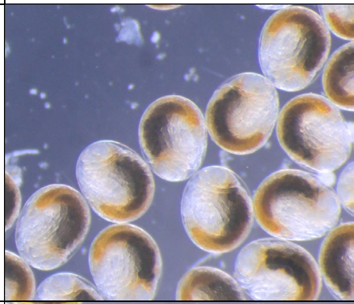
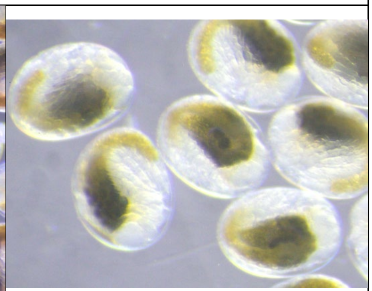
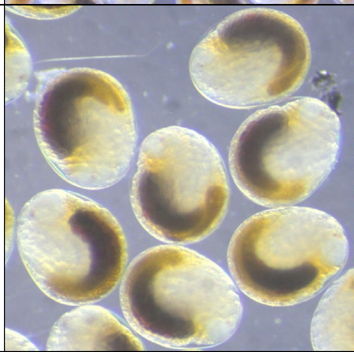
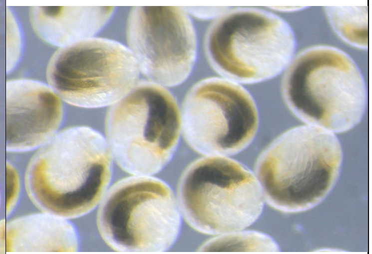
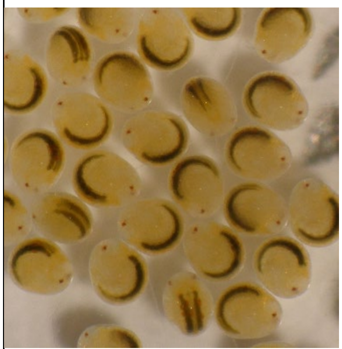
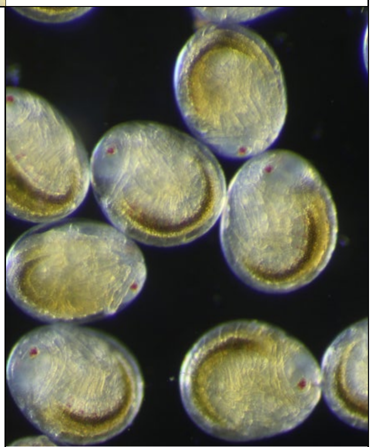
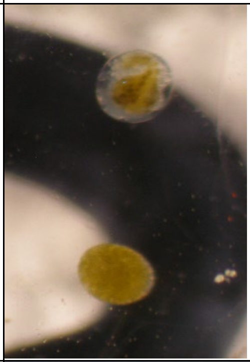
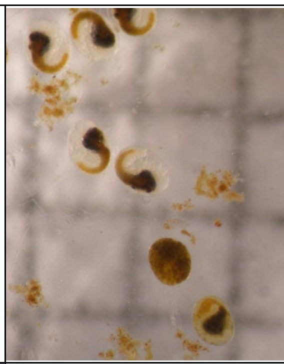
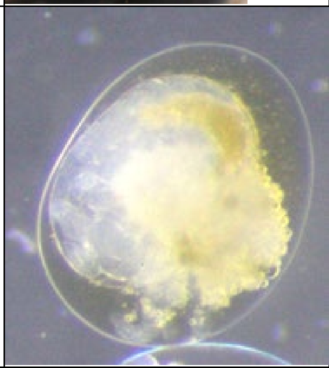
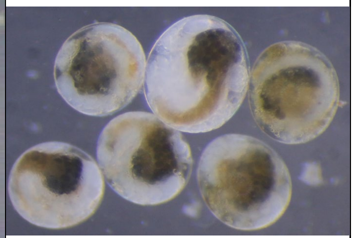
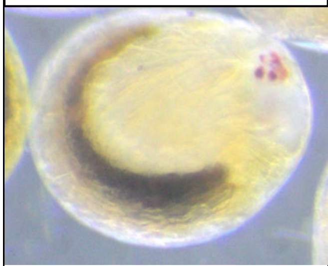
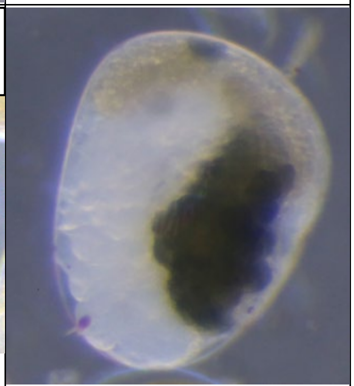
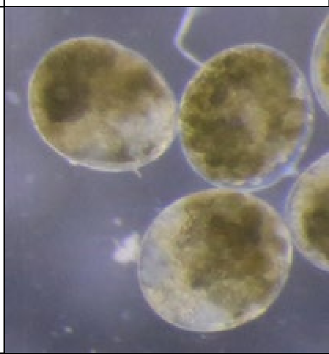
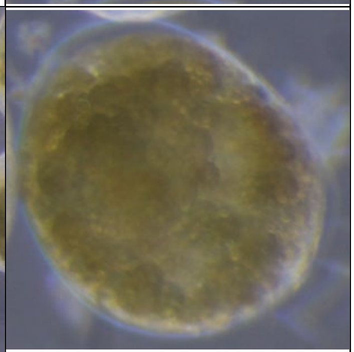
Stage	Distinguishing marks	Description	Illustrations	
5	Less "comma"-shaped, with distinguishable legs. No eyes yet.	Less "comma"-shaped, with clear legs and a white see-through "head". No eyes yet.		
5B	Similar to Stage 6 – two stripes, eyes however are not existing or only a suggestion for an eye (not red but a light spot).			
6	"Banana"-shaped, two stripes, with clear red eyes.	"Banana"-shaped, with clear red eyes. Shape of juvenile is clear in embryo.		
7	Small copy of adults.	Juveniles, still curved and/or hatching from the egg and/or potentially free swimming.		

Table 8.2. Embryo aberration forms *Gammarus* spp. (Amphipoda)

	Name	Description	Illustrations	
1	Underdeveloped	Embryos in other developmental stage than rest of the brood		
2	Membrane abberations	Enlarged membranes and/or leakage between inner and outer membrane.		
3	Malformations	Malformed embryos	Example: Head and eye aberration 	
4	Undifferentiated embryos	Embryos where the features of the cell structure is no longer clear		
5	Dead	Embryo is non-transparent in color, e.g. milky-white or dark.