RAPPORT L.NR. 7767-2022

Analytical procedures and interpretation basis for biomarkers used in environmental effect monitoring for the water column at the Norwegian continental shelf

Foto: NIVA

Norwegian Institute for Water Research

REPORT

Main Office
Økernveien 94
NO-0579 Oslo, Norway
Phone (47) 22 18 51 00

NIVA Region South
Jon Lilletuns vei 3
NO-4879 Grimstad, Norway
Phone (47) 22 18 51 00

NIVA Region East
Sandvikaveien 59
NO-2312 Ottestad, Norway
Phone (47) 22 18 51 00

NIVA Region West
Thormøhlensgate 53 D
NO-5006 Bergen Norway
Phone (47) 22 18 51 00

NIVA Denmark
Njalsgade 76, 4th floor
DK 2300 Copenhagen S, Denmark
Phone (45) 39 17 97 33

Internet: www.niva.no

Title	Serial number	Date
Analytical procedures and interpretation basis for biomarkers used in environmental effect monitoring for the water column at the Norwegian continental shelf	7767-2022	19.09.2022
Author(s) Beyer, Jonny Brooks, Steven J.	Topic group Environmental monitoring	Distribution Open
	Geographical area Norway	Pages 61

Client(s)	Client's contact person
Offshore Norge (formerly NOROG, Norwegian Oil and gas)	Rolf C. Sundt
Client's publication:	Printed NIVA Project number 200262/8

Summary

Herein, method description documents for the biomarker parameters that are required or recommended by the current edition of the guidelines for offshore water column monitoring at the Norwegian shelf (M-300, revised version 2020) are systemised, examined and discussed. The objective of this study is to assess the quality status of open and available method and analysis descriptions, identify any significant shortages on biomarker method descriptions and method validation issues and to discuss/suggest possible correcting and improving measures.

Four keywords		Fire emneord	
1.	Offshore water column monitoring program	1.	Offshore vannsøyleovervåking
2.	Biomarker procedures	2.	Biomarkør prosedyrer
3.	Blue mussel and fish	3.	Blåskjell og fisk
4.	Method improvements	4.	Metodeforbedringer

This report is quality assured in accordance with NIVA's quality system and approved by:

Jonny Beyer	*Steven J. Brooks*	*Adam Lillicrap*
Project Manager/Main Author	Quality Assurance	Research Manager

ISBN 978-82-577-7503-2

NIVA-report ISSN 1894-7948

Analytical procedures and interpretation basis for biomarkers used in environmental effect monitoring for the water column at the Norwegian continental shelf

WCM2020 R&D

Preface

This study was performed by NIVA as part of the research activities in the Norwegian offshore monitoring program, section water column monitoring (WCM). The funding client of the WCM program is the association of oil and gas producing companies at the Norwegian continental shelf (Offshore Norge, formerly NOROG, Norwegian Oil and gas). The work with this report was performed mainly by Jonny Beyer with contributions and QA assistance from Steven J. Brooks, the latter who also is PL for WCM project activities at NIVA.

Oslo, Sept. 13th 2022

Jonny Beyer

Table of contents

Summary

Offshore water column monitoring (WCM) activities on the Norwegian continental shelf (NCS) includes the use of selected contaminant responsive biomarkers in fish and mussels among the monitoring parameters. For these biomarker parameters, there is a need to validate their analytical determination and their suitability and interpretability for the type of contamination issues and monitoring needs that are targeted by the WCM program. Methodological validations concern the availability of technical analysis procedures that meet analytical quality requirements and that, preferably, are harmonized internationally. Validation of suitability will basically concern the existence of fit for purpose documentation, including appropriate sensitivity, dose-responsivity, adequate noise-signal ratios and insight about any relationships between the biomarker and organism function and fitness. Validation of biomarker interpretability concerns measures to optimise QA and QC needs for biomarker detections involving the systematic development/provision of reference materials for facilitating data comparisons between studies and across different users. In this report, technical description documents for biomarker parameters required/recommended by the current edition of the guidelines for offshore water column monitoring (WCM) at the Norwegian shelf are examined and discussed, both to identify any pressing shortages on method validations and for discussing/suggesting possible correcting measures. When technical demands for analysis quality of the biomarker assays are sufficiently met, the relevance of the given biomarker response(s) in relation to other more higher hierarchy effects can be addressed more adequately. Regarding the selection of biomarkers that are employed in WCM, it remains uncertain whether the present parameters and procedures described by the official guidelines represent the optimal choices, as there are still many shortages on data and knowledges that validate and confirm their suitability for the kind of environmental conditions and contamination situations that occur in the mixing zones for discharges from offshore oil and gas production installations. The present study can serve to provide some insight on the state of art on how different biomarkers are being measured in the WCM program. Needs are identified for knowledge regarding normal range of variation of biomarkers and for documentation of dose- and time-response relationships, especially for field realistic exposure scenarios. Some key improvements for the interpretation basis of WCM relevant biomarkers are discussed, including the possibilities for establishing objective judgement criteria of observed biomarker signals.

Sammendrag

Tittel: Analytiske prosedyrer og tolkningsgrunnlag for biomarkører brukt i miljøeffektovervåking for vannsøylen på norsk kontinentalsokkel

År: 2022

Forfattere: Jonny Beyer and Steven J. Brooks

Utgiver: Norsk institutt for vannforskning, ISBN 978-82-577-7503-2

Offshore vannsøyleovervåking aktiviteter på norsk kontinentalsokkel inkluderer bruk av utvalgte forurensnings-responsive biomarkører i fisk og blåskjell blant overvåkingsparameterne. For disse parameterne er det behov for å validere deres analytiske bestemmelse og deres egnethet og tolkbarhet for den typen forurensningsproblemer og overvåkingsbehov som gjelder på norsk sokkel. Metodiske valideringer gjelder tilgjengeligheten av tekniske analyseprosedyrer som oppfyller analytiske kvalitetskrav og som fortrinnsvis er harmonisert internasjonalt. Validering av egnethet vil si eksistens av dokumentasjon som viser egnet sensitivitet, dose-responsivitet, tilstrekkelige støy-signalforhold samt innsikt i eventuelle sammenhenger mellom den aktuelle biomarkør-respons og organismens biologiske funksjon og tilstand. Validering av biomarkørtolkbarhet dreier seg om tiltak for å optimalisere QA- og QC-behov for biomarkørmålingene i overvåkingsprogrammet, for eksempel ved systematisk utvikling/bruk av referansemateriale som kan forenkle datasammenligninger på tvers av studier og på tvers av ulike brukere. I denne rapporten er tekniske metodebeskrivelser (SOPer) for de ulike biomarkørparametere som inngår i gjeldende utgave av retningslinjene for offshore vannsøyleovervåking på norsk sokkel blitt samlet og vurdert, både for å påvise mulige mangler og for å diskutere/foreslå mulige korrigerende tiltak. Når tekniske krav til analysekvaliteten til biomarkøranalysene er tilstrekkelig oppfylt, vil det også bli lettere å vurdere/dokumentere relevansen av de(n) gitte biomarkørresponsen(e) i sammenheng med bruken offshore og i forhold til om de kan linkes til andre høyere-ordens biologiske effekter. Når det gjelder det nåværende utvalget av biomarkører som inngår i veilederen, er det fortsatt usikkert om parameterne og prosedyrene samlet sett utgjør et optimalt valg i forhold til behovene som denne miljøovervåkingen har. Fortsatt er det mangler på data og kunnskap som kan validere og bekrefter deres egnethet for de spesielle miljøforhold og forurensningssituasjoner som fins i blandesonene for utslipp fra offshore olje- og gassproduksjonsinstallasjoner. Denne rapporten gir en oversikt over hva som er metodemessig og kunnskapsmessig status for vannsøyledelen av den biologiske effektovervåkingen ved norske olje og gass plattformer på norsk sokkel. Det identifiseres et generelt behov for bedre kunnskap om normal variasjon av biomarkører og for dokumentasjon av dose- og tid-respons-forhold som kan validere egnetheten for disse parameterne i overvåkingsaktivitetene, og da spesielt for feltrealistiske eksponeringsscenarier. Noen sentrale forbedringer for tolkningsgrunnlaget for de aktuelle biomarkørene diskuteres, inkludert nødvendigheten av å etablere objektive vurderingskriterier for observerte biomarkørsignaler.

1 Introduction

1.1 Background

Oil and gas producing companies operating on the Norwegian Continental Shelf (NCS) have for more than 25 years performed offshore monitoring surveys in sediments and water column in areas close to oil and gas producing installations in order to investigate possible environmental impacts of the operations and to provide the environmental authorities with a better basis for regulation. Recommendations to the strategies and practical contents of marine monitoring activities have for several decades been communicated from transnational expert organisations such ICES-WGBEC, OSPARCOM, MEDPOL and HELCOM, which also has formed a basis for the development and revisions of offshore monitoring guidelines such as from OSPARCOM (2004) and Norwegian Environment Agency (Miljødirektoratet, 2020; Miljødirektoratet, 2015; Miljødirektoratet, 2011). The two key parts of the current effect monitoring on the NCS include ecological community-based effect surveys (sediment monitoring program) and the biomarker-effect-based water column monitoring program (WCM) (Figure 1), the latter which is the focus of this report.

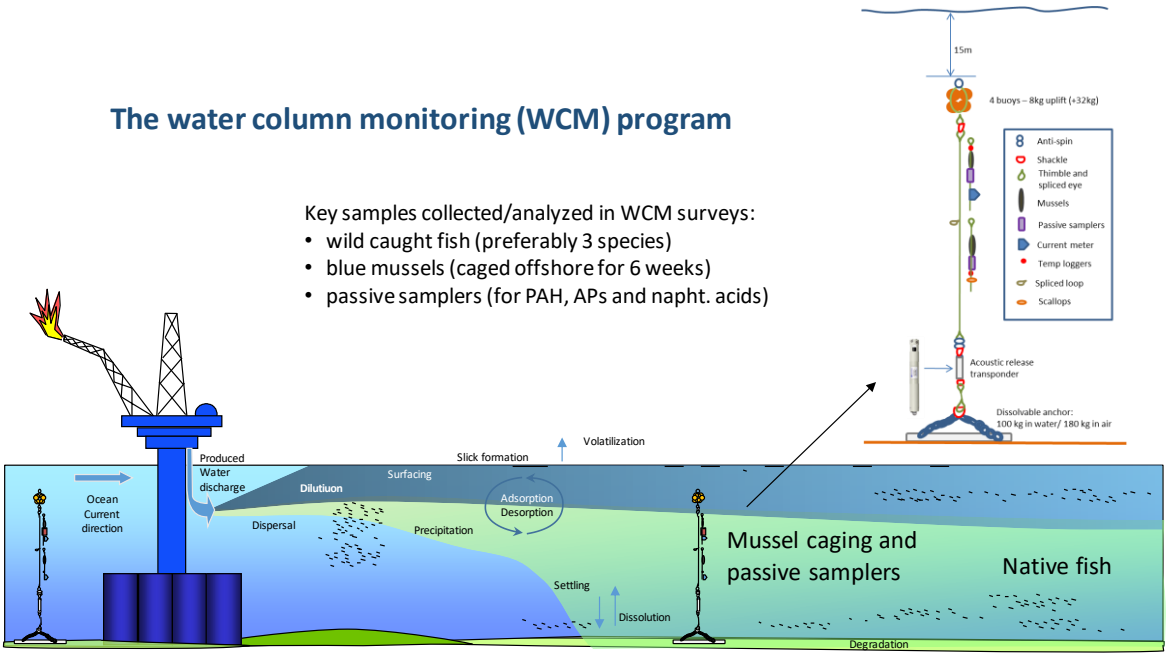

Figure 1: The Norwegian Water Column Monitoring (WCM) program in its present shape is designed to investigate exposure and effect markers in caged blue mussels and native fish specimens from the area close to offshore oil and gas production installations.

The WCM surveys, earlier used a strategy of caging sentinel mussels and fish in the water column with various distances from selected offshore installations. When the WCM program recently was merged with the offshore Condition Monitoring program, the focus in the fish part of WCM was changed to the use of native fish specimens obtained from platform-close areas. Currently, WCM surveys with transplant caged blue mussel and native fish are required to be performed each third year at selected offshore fields, whereas lab-based studies are done during the two intermediate years with the aim to further refine and quality assure the program content (Miljødirektoratet, 2020). The results from the WCM surveys performed up to now have been summarised in multiple

survey reports, and overviews of these are found in several review articles (Hylland et al., 2008; Brooks et al., 2011b; Bakke et al., 2013; Beyer et al., 2020).

Produced Water (PW) and rock cuttings produced by the drilling operations are the two main sources of oceanic chemical contamination from the offshore oil and gas operations on the NCS. The PW is treated to lower the content of unwanted substances and then reinjected to the formations deep under the seabed or discharged to the water column. After treatment, the PW discharges still contain dispersed crude oil, polycyclic aromatic hydrocarbons (PAHs), alkylphenols (APs), metals, and many other chemical constituents of various environmental relevance. Risk-based regulation, greener offshore chemicals and improved cleaning systems have reduced the environmental risks of such PW discharges, but discharge of PW is still the largest operational source of oil pollution to the sea from the offshore petroleum industry on the NCS. A treated PW discharge typically contains dispersed oil (normally 10–100 mg/L range), dissolved hydrocarbon (HC) gases, suspended particles (e.g. clay), inorganic salts, organic acids, aromatic hydrocarbons, ketones, phenol/alkylphenols, heavy metals, and naturally occurring radioactive materials (NORM) and often also chemicals added to the production system to aid the extraction processes or to protect against biofouling and corrosion. In addition, many PW compounds remain unidentified and in chromatographic analyses they elute in what is called the Unresolved Complex Mixture (UCM).

In this report, method description documents for chemical and biomarker parameters in the WCM part of the offshore monitoring guidelines are collected and assessed. The quality and suitability of the obtained documents are assessed and possible needs for further refinements are discussed. Method descriptions that are publicly accessible (via Internet), and free of charge, are given priority, provided they meet adequate technical and academic quality demands. Web-links to each procedural document are shown whenever possible. The second main objective of the present report is to address the interpretation basis for the addressed selection of markers. The biomarker interpretation issue concerns both the aspects of analytical quality (precision and accuracy) of each biomarker parameter assessed in environmental monitoring work as well as the issue of interpreting the biomarker effect signal into an ecotoxicological adverse outcome context, that means whether the biomarker signals an early warning for possible adverse impacts/outcomes later in the effect chain (i.e., effects on higher-hierarchical endpoints). In the present report, we discuss interpretation relevant aspects for some of the biomarkers that most frequently have been used in the WCM program, including:

- typical normal range of biomarker in controls (background signal)
- dose and time response relationships, intensity of responses
- overview of uncertainty elements (confounding factors)
- reference materials for interpretation of biomarker signals

1.2 SOP, QA and interpretation aspects of WCM biomarkers

To pursue proper quality of analysis data are key when ecotoxicity biomarkers are used as tools for environmental pollution effect monitoring. The availability of biomarker analysis procedures that have been validated, quality assured and internationally harmonized is of key importance for facilitating good technical quality of biomarker measurements in pollution effect monitoring.

According to the regulation, dispersal and risk modelling must be done for each PW case to make predictions of the concentrations of different key constituents from PW which possibly can be encountered within the downstream water column bodies and with different distances from discharge points. Thus, based on controlled exposure systems with a set of PW mixtures a system for suitability-verification of monitoring parameters can and should be made, both regarding the

parameters that are already in use in WCM and for parameters that are candidates for possible future inclusion. Likewise, such suitability verification studies should also enable testing of how specified sets of effect-based parameters perform together within an integrated exposure-effect assessment approach, as such approaches are likely to become more important for marine effect monitoring programs in the years ahead given the requirements and recommendations of the EU Marine Strategy Framework Directive (MSFD). However, to address the methodological aspects and tools of integrated pollution effect monitoring strategies is a rather complex task. These issues will therefore be covered rather superficially in this report.

1.3 BAC and EAC assessment criteria of WCM data

Marine pollution monitoring programs such as WCM are in need for a harmonized and standardized way of interpreting ecotoxicological monitoring data from analyses of chemical contaminant endpoints as well as biological effect endpoints in marine sentinel study species. The development of such an internationally harmonized framework for interpretation of pollution monitoring data has over years been addressed by transnational expert working groups under the auspices of ICES and OSPAR (Davies and Vethaak, 2012; Vethaak et al., 2017).

To establish an international harmonized interpretation basis for biological effect endpoints, significant efforts conducted under the auspices of the ICES WGBEC group and within the Oslo and Paris Commission framework have concerned development of assessment criteria for biological and ecotoxicological effects parameters (biomarkers) in selected marine sentinel species.

The two assessment criteria which are made for assessment of biological effects include:

- Background Assessment Concentration (BAC)
- Environmental Assessment Criteria (EAC)

The present status and outcome of these important efforts are available at this website:

https://dome.ices.dk/ohat/trDocuments/2019/help_ac_biota_biological_effects.html

The ICES work aimed to establish a transnational and improved interpretation basis and assessment criteria for marine pollution monitoring data. In the initial phase, the work was concentrated on chemical monitoring endpoints, i.e., to define typical background concentrations (BCs), background assessment concentrations or criteria (BAC), and environmental assessment concentration or criteria (EAC) of concentration data of hazardous chemical contaminants in specific categories of samples. Later, an analogous assessment criteria strategy was adopted for integrative interpretation of biological effects data. For biological effect endpoints, BAC should represent the upper threshold value for a typical background situation/range and EAC should represent the border between acceptable and unacceptable effect levels for the given biomarker and the given bioindicator species.

The first set of recommended BAC and EAC assessment criteria for biological effects parameters in various marine bioindicator species was provided in Davies and Vethaak (2012) and in updated background documents and technical annexes (OSPAR, 2013). These assessment criteria recommendations were subsequently agreed to by OSPAR (OSPAR, 2012).

The ICES Study Group for the Integrated Monitoring of Contaminants and Biological Effects (ICES SGIMC) proposed BAC/EAC threshold level data for a collection of biological effect parameters/biomarkers in marine fish species commonly used in pollution effect monitoring and suggested also the use of a simplified three-colored “traffic light”-like guidance system for easy interpretation/classification of biomarker level intensities in relation to BAC and EAC. The actual

choice of the three colors has apparently varied somewhat with time, but in the most recently updated reports on this issue, e.g., Lyons et al., (2017); Vethaak et al. (2017), BLUE represents response levels below BAC, GREEN represents response levels above BAC but below EAC, and RED represents response levels above EAC. The ICES SGIMC working group has recommend that GES be expressed as some high percentage compliance with the green-red boundary (ICES, 2011).

2 Biomarker tools used in WCM

2.1 Biomarker parameter selections

A special feature of the Norwegian WCM surveys (and the associated laboratory effect studies) is the integrated monitoring strategy used and particularly the involvement of ecotoxicological effect parameters (biomarkers) to detect possible adverse effects in sentinel organisms due to the operational offshore discharges that are being monitored in offshore surveys. Essentially, a biomarker in ecotoxicological monitoring is a parameter that can be determined in a sentinel organism and that have been shown to be responsive to the type of contamination issue (involving both the type of contaminants and their realistically expected concentrations) that is addressed in the monitoring study. The method description resources that have been examined in this study are listed below (Table 1). The addressed selection of procedures was chosen based primarily on the biomarker parameters that are required or recommended parameters in the WCM section of the most recent edition of the Norwegian offshore monitoring guidelines (Miljødirektoratet, 2020) (see parameters listed in Table 1). As produced water (PW) is the main category of operational discharges that is addressed by the effect-based studies in WCM, the different biomarkers that are required/recommended by the guideline would logically be expected to be responsive to contaminants typically found in such offshore PW discharges.

The Norwegian offshore monitoring guidelines inform about the biomarker parameters to measure in the sentinel organisms, but they don't provide any technical descriptions for *how* to carry out these analyses. Detailed analytical method descriptions or standard operating procedures (SOP) for the different monitoring parameters are not provided. Moreover, for most of the effect parameters that are required in the monitoring, it is a challenge that the degrees by which the parameters are fit for the purpose in the monitoring is typically incompletely and insufficiently documented. It is also a question whether we can understand and describe the combined effects of several chemicals acting together within the exposed sentinel organisms (and in its population and community). The importance of these issues has been recognised by different transnational expert advisory organizations such as ICES-WGBEC (Working Group on Biological Effect of Contaminants), OSPARCOM and JPI Oceans. A key report that includes method advisory and assessment criteria information of a broad range of biomarkers in various marine bioindicator species and for integrated marine monitoring studies is the updated background document from OSPAR on biological effects monitoring (OSPAR, 2013).

Table 1: Overview of technical procedure resources for monitoring parameters for the water column that are required or recommended by the most recently edition of the offshore monitoring guidelines. For web-based hyperlinks to the different SOP resources see the electronic version of the report.

Activity	**SOP name**	**Link to procedure or SOP resources**
Method descriptions or SOPs for blue mussel specimens	Sampling and dissection of *Mytilus* mussels	(OSPAR, 2013)
	Size and condition (CI) data of whole mussel	Link
	Speciation of mussels at the individual level	(NIVA, 2017; Brooks and Farmen, 2013; Inoue et al., 1995)
	Reproductive maturity & spawning status of blue mussels	(Duinker et al., 2008; Sunila et al., 2004)
	General health status of blue mussels assessed by «stress on stress»	Link
	PAH concentration in mussel soft tissue measured by GC-MS	Link
	Metal (Hg, Pb, Cd, Ba) analysis in mussel soft tissue with ICP	NS EN ISO 17294-2 and NS 4768 (Hg)
	Micronuclei (MN) in mussel non-granular haemocytes	(Bolognesi and Fenech, 2012), (NIVA, 2020)
	Lysosomal membrane stability in blue mussel digestive gland tissue	Link, Link
Method descriptions or SOPs of fish specimens	Determination of species, size, gender and age of fish	Link
	Determination of fish liver and gonad somatic index (LSI and GSI)	Link
	Determination of fish condition index (CI)	Link
	PAH/NPD concentration in fish liver measured by GC-MS	Link
	PAH metabolite concentration in fish bile measured by GC-MS/LC-FD	Link
	PFAS in fish blood	(NIVA, 2022)
	Histological, including lysosomal, changes in fish liver	Link (Au, 2004)
	CYP1A1-induction in fish liver assessed by EROD-activity determination	Link
	CYP1A1-induction in fish liver assessed by CYP1A ELISA	(Goksøyr, 1991)
	CYP1A1-induction in fish liver assessed by qPCR of AHR and CYP1A1 mRNA	Method source TBD, Meier et al 2020, with supplement?
	DNA damage assessment in fish by (a) DNA adduct measurement in liver	Link
	DNA damage assessment in fish by (b) comet assay of strand breaks in lymphocytes	Link
	DNA damage assessment in fish by (c) micronuclei formation in RBC	(Barsiene et al., 2004; D'Agostini and La Maestra, 2021)
	Acetylcholine esterase (AChE) inhibition in fish muscle	Link
Other WCM relevant SOPs and method resources	Determining polymer-water and polymer-polymer partition coefficients of organic compounds	Link
	Passive sampling of hydrophobic contaminants in water using silicon rubber samplers	Link
	Guidelines on quality assurance	Link, Link, Link

TBD: to be decided

2.2 Water column monitoring parameters in mussels

2.2.1 Size and condition index (CI) of whole mussel

Blue mussels (*Mytilus edulis* spp.) (Figure 2) have been much used as sentinels in WCM and other marine pollution monitoring surveys (Beyer et al., 2017). Mussels must be treated carefully during sampling and dissection so as not to cause tissue damage that may hamper the quality of the tissue sections for histology analyses and for the other endpoints, e.g. clean digestive gland sample without gonad tissue or stylus removed. A detailed procedure for sampling and dissection of *Mytilus* sp. has been provided by OSPAR in the updated background document and technical annexes for biological effects monitoring, (OSPAR, 2013).

According to the offshore guidelines, mussel length (L) and soft tissue dry weight (W) must be estimated in representative sub-samples of mussels per caging group for estimation of the condition index (CI). Hansson (2017) published a report through ICES TIMES that explains the relevance of common biological variables like shell volume, flesh weight, shell weight, and CI in relation to the general health status of blue mussels, and how to quantify these parameters. The CI expresses the relationship between the body mass dry weight and the total length of the shell and can be calculated in different ways, perhaps most commonly by means of the formula: $CI = W/L^3$ (Riisgård et al., 2014) whereas NIVA most often uses the formula: (soft tissue dry weight/shell dry weight) *100 (Moschino and Marin, 2006; Orban et al., 2002). The CI of blue mussels vary considerably from population to population and during the year cycle within each population, especially since the CI is affected by the main spawning events.

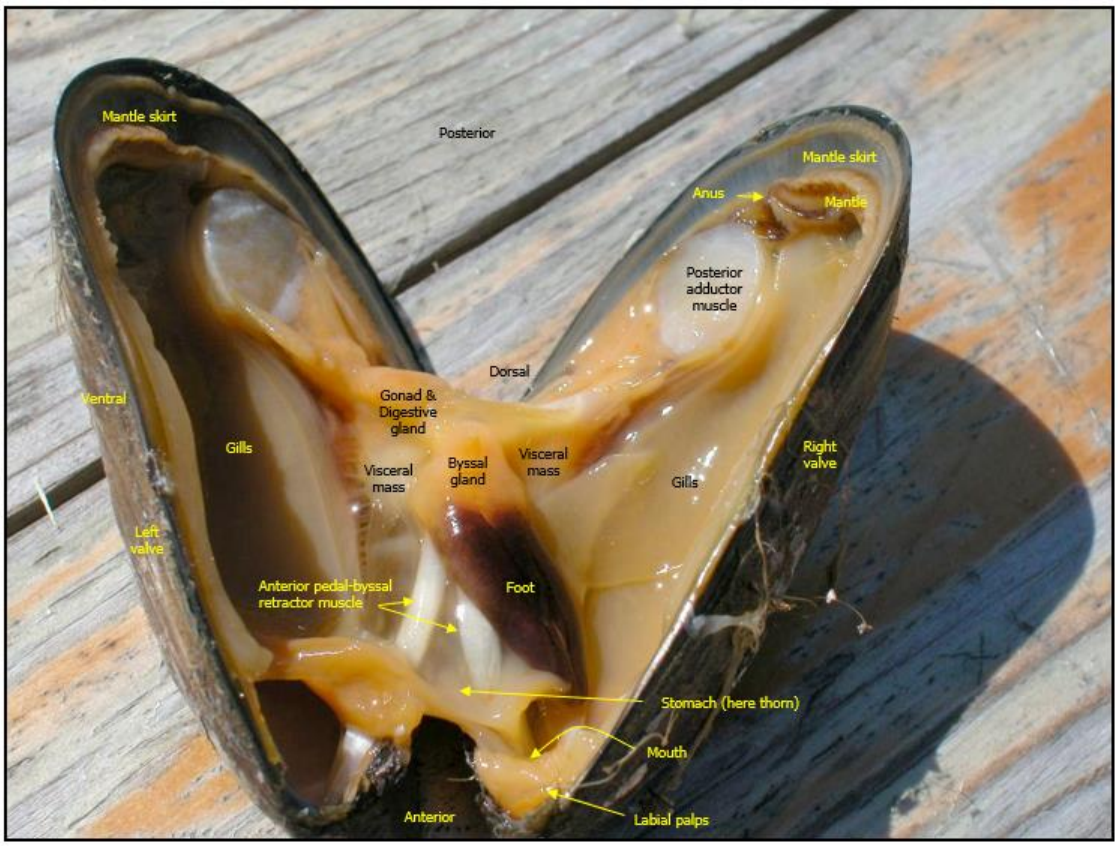

Figure 2: Overview of the gross body architecture and anatomy features of a blue mussel.

2.2.2 Speciation of mussels at the individual level

Differences in contaminant bioaccumulation and biomarker responses may occur among the blue mussel sub-species *Mytilus edulis*, *M. trossulus*, and *M. galloprovincialis*. Determination of species / subspecies / hybrid composition is thus to be carried out on a representative subsample of each mussel group in WCM surveys and the grouping and distribution of mussels to different caging stations shall be done so that the species composition is the same at all stations. As the external morphology of the different subspecies is influenced by environmental factors, species identification by visual traits is never certain, but subspecies identification should rather be achieved based on measuring the variability of the polyphenolic adhesive protein gene after polymerase chain reaction amplification of gill tissue DNA. In NIVA, a confidential SOP is used to analyse this parameter, but method descriptions are also accessible from several publications, e.g., Inoue et al., (1995); Bignell et al. (2008); Brooks and Farmen (2013), clarifying that the basepair-size (bp) of amplified DNA segments from the *Glu* gene, which is determined by 4% agarose gel electrophoresis, show specific bands for *M. edulis* (180 bp), *M. trossulus* (168 bp), and *M. galloprovincialis* (126 bp), and with the presence of double bands suggesting *Mytilus* hybrids (Brooks and Farmen, 2013).

2.2.3 Reproductive maturity & spawning status of blue mussels

Blue mussels have seemingly a relatively simple gross anatomy (Figure 2, Figure 3), but it contains organs that are highly specialised and efficient. During main spawning events, blue mussels will lose a large part (40-59%) of their total bodyweight (Cossa, 1989; Duinker et al., 2008) and such a sudden and dramatic change of body-mass represents a significant physiological challenge for the individual and will also affect the bioburdens of chemical contaminants that are sequestered in the mussels tissues. To assess reproductive maturity and spawning status of mussels is therefore relevant when mussels are used in pollutant monitoring as well as in pollutant effect monitoring. According to the offshore guidelines, a histological sample should therefore be taken of all mussel individuals that are analysed in order to analyse the gender and the spawning status. Studies that covers the relevance of reproductive maturity and spawning status of blue mussels have been reported by Sunila et al. (2004) and by Duinker et al. (2008).

According to the offshore monitoring guidelines, a mussel histology sample must be obtained for documenting the maturity and spawning status, preferably as a transverse incision that include the main internal organs and gonadal tissue and gonad products (e.g., as shown in Figure 3). The dissection process is an extremely important stage in the histological process, and standardized dissection procedures facilitates between specimen and sample comparability and will also simplify the interpretation of any downstream histological analysis. Recently, Eggermont et al. (2020) reported a detailed description and visualization of the vascular-related anatomy of *Mytilus edulis* by using high resolution X-ray computed tomography and histology combined with 3D-reconstruction using AMIRA-software. The study presents a detailed description and visualization of the vascular-related anatomy of and provide image resources that can aid the practical work with mussel dissection as well as haemolymph withdrawn from the posterior adductor muscle or from other positions and tissues within the mussel body from where haemolymph samples can be obtained.

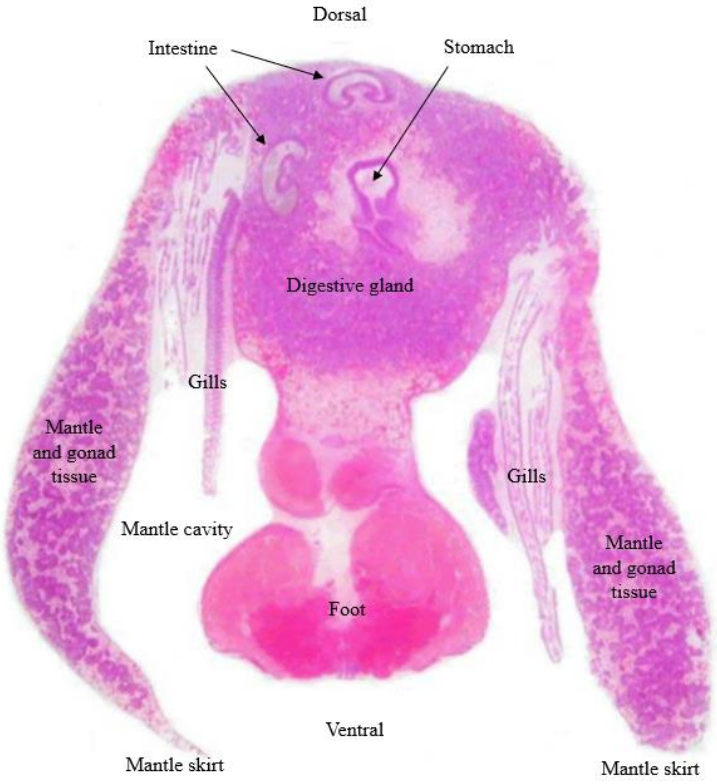

Figure 3: Transverse histological section showing an overview of the body architecture and key organs and tissues of the blue mussel, shown with standard haematoxylin and eosin histological coloration. The tissue tags are put on top of a picture obtained from Bignell et al. (2012) with permission.

2.2.4 Blue mussel histopathology

Mussel histology/histopathology biomarkers have previously been addressed in WCM surveys, but these parameters are currently omitted due to uncertainties about method and fit-for purpose issues. Firstly, there is a need for high-quality SOPs for the histology/histopathology parameters in different tissues and body compartments of blue mussels, e.g. digestive gland, gills and gonads. Secondly, an appropriate amount of documentation of responsivity/suitability must be provided for these parameter candidates, especially with respect to contamination-categories and contamination-dosages that realistically can be encountered close to PW discharges offshore. A detailed procedure document of blue mussel sampling procedures and histology/histopathology biomarkers for publication in the ICES TIMES series has been under preparation for some time (Bignell et al., in prep). It is possible that this SOP, when it is completed, will address a similar collection of parameters as described in the mussel histology/histopathology background documents which earlier have been communicated by ICES (Bignell et al., 2012) and OSPAR (2013), i.e., cell type composition in digestive gland epithelium, digestive tubule epithelial atrophy and thinning, lysosomal alterations in digestive gland cells, and inflammation in various tissues (ibid.). Assessing the applicability for these parameters for use in WCM would be needed, as noted above. However, any such suitability assessments may also choose to target other histology/histopathology parameters in mussel sentinels if interesting parameter candidates can be identified from the available and relevant research literature, e.g., (Au, 2004; Bignell et al., 2008; Garmendia et al., 2011c; Garmendia et al., 2011b; Garmendia et al., 2011a), and many others. Such a broader review will be needed but is outside the scope of this report.

2.2.5 Blue mussel stress on stress (SOS) parameter

According to the offshore guidelines, the physiological condition of caged blue mussels should be assessed by the «stress on stress, SOS» biomarker, which is a measure of general health status at the whole-organism response level as it is related to the available energy reserves (ATP) required to maintain shell closure (and thereby survival of mussels when kept in air). Recently ICES TIMES published a SOP of the SOS parameter in blue mussels (Thain et al., 2019).

Although being time consuming to measure, the SOS biomarker is considered a simple and low-cost biomarker of the general health condition of *Mytilus* mussels. The assay measures the time (number of days) it takes before the test mussels no longer can keep their valves closed. Background documents on the SOS parameter in mussels have earlier been provided by ICES (Chapter 16 in (Davies and Vethaak, 2012)) and by OSPAR (Chapter 19 in (OSPAR, 2013)). Background SOS responses up to 18 and 16 days have been observed for *M. galloprovincialis* and *M. edulis*, respectively. OSPAR previously suggested that blue mussels may be considered healthy, stressed but compensating, or severely stressed if the SOS time is more than 10 d, between 5 and 10 d, or less than 5 d, respectively (OSPAR, 2013). Those two response thresholds (10 and 5 days) was suggested to be used as the BAC and EAC for stress on stress measurements in blue mussels (Martinez-Gomez et al., 2017; Vethaak et al., 2017; Thain et al., 2019).

The BAC and EAC assessment criteria for the SOS parameter in blue mussels from ICES working groups are defined to 10 and 5 days, respectively, with high values indicate healthy mussels (ICES, 2011).

Although the SOS response can be correlated quantitatively to tissue concentrations of single contaminants, it also indicates the overall impact of multiple stressors. Several consider it to be a relevant marker for integrated effect measurements. Studies have found SOS in mytilids to be affected by various kinds of contaminant stress, such as heavy metals, organic pollutants, organometals and contaminated field sediment, but also by several natural factors such as the water temperature, size of mussels and time of year (deZwaan and Eertman, 1996; Hellou and Law, 2003; Hamer et al., 2008). A SOP of the SOS parameter is in blue mussels was recently published in ICES TIMES (Thain et al., 2019), see hyperlink provided in Table 1. Although most studies have demonstrated shorter SOS closure time in mussels when exposed to xenobiotics, some experiments with low concentrations of PAH from contaminated sediments, have observed a slight increase in the survival, possibly suggesting an hormesis effect (Eertman et al., 1995).

2.2.6 PAH and NPD concentration in mussels measured by GC-MS

Experience from many previous surveys in the WCM program suggests that the concentrations of petrogenic polycyclic aromatic hydrocarbons (PAH) and related organic compounds in mussel soft tissue have been suitable for discriminating mussel groups caged with different distances downstream from offshore installations (Hylland et al., 2005; Sundt et al., 2006; Hylland et al., 2008; Brooks et al., 2011b; Brooks et al., 2011a; Harman et al., 2011; Sundt et al., 2012; Brooks et al., 2013; Pampanin et al., 2013; Harman et al., 2014; Brooks et al., 2015; Pampanin et al., 2019). According to the monitoring guideline for the NCS, the concentration of NPD (naphthalene, phenanthrene and dibenzothiophene, including their C1–C3 alkyl homologues) and total PAH (quantified by the EPA 16 priority PAHs) shall be measured in mussel soft tissue, although that collection of PAHs is not fully representative for petrogenic PAH mixtures. The EPA 16 PAHs include: naphthalene, acenaphthene, acenaphthylene, fluorene, phenanthrene, anthracene, fluoranthene, pyrene, benz [a]anthracene, chrysene, benzo[b]fluoranthene, benzo [k]fluoranthene, benzo[a]pyrene, benzo [ghi]perylene,

indeno [1,2,3,cd]pyrene, and dibenz [a,h]anthracene. A reference sample must be used when measuring PAH in mussels in order to facilitate quality control and comparison of analysis data between caging locations and survey years.

2.2.7 Metal (Hg, Pb, Cd, Ba) analysis in mussel soft tissue with ICP

According to the offshore guidelines, concentrations of heavy metals (Hg, Pb, Cd, Ba) are to be measured in mussel soft tissue by use of an ICP procedure and reference samples must be used in order to facilitate quality control and comparison of analysis data between caging locations and survey years. The most common metals in PW discharges include arsenic, cadmium, copper, chromium, lead, mercury, nickel and zinc. An open source SOP that covers these analyses is not available through ICES TIMES, but relevant methodological information on this parameter is accessible from many literature sources, e.g., (Vaidya and Rantala, 1996; Giusti and Zhang, 2002; Strizak et al., 2014). According to discharge summary reports from the industry and the Norwegian Environment Agency, the highest levels of lead, mercury and zinc in PW discharge streams are more than a factor 1000, and arsenic and cadmium more than a factor 100, above the Norwegian coastal water background levels (Beyer et al., 2020). The highest concentration reported for arsenic, cadmium, copper, and lead was from one low volume PW source from a gas and condensate field. If these values are excluded the levels of all metals except mercury were a factor <100 above naturally levels in seawater. Barium and iron constituents in PW are also exceeding background concentrations in seawater (by a factor more than 1000).

2.2.8 Micronuclei (MN) in mussel non-granular haemocytes

The micronuclei (MN) assay is one of the most widely used genotoxicity biomarkers in toxicology and ecotoxicology with several thousand published studies in the Web of Science literature database, of these several hundred studies have addressed blue mussel as the bioindicator organism. The MN assay provides a measure of chromosomal DNA damage in different cell categories, occurring as a result of either chromosome breakage or chromosome mis-segregation during mitosis because of cellular spindle apparatus malfunctions, the lack or damage of centromere or chromosomal aberrations (Fenech, 2000).

In *Mytilus* sp., the MN assay is commonly applied in circulatory non-granular haemocytes or in cells extracted from enzymatically treated gill tissue. An open source SOP that covers these analyses is going to be available through ICES TIMES in a few months' time (S. Brooks, pers. comm.). However, MN protocol relevant information is accessible from many literature sources, e.g., (Fenech, 2000; Barsiene et al., 2004; Bolognesi and Hayashi, 2011; Bolognesi and Fenech, 2012). The classic method for assessing MN is by using microscopy scoring of different cell types for the relative occurrence of micronuclei (MNi) and nuclear buds (Bolognesi and Hayashi, 2011; Bolognesi and Fenech, 2012). Recently, steps have been taken to improve this manual assay, such as by using Cytospin™ cytocentrifuge with cytofunnel™ sample chambers in order to provide perfect density of cells for simplifying cell scoring, but further optimisations of these techniques are still required (S. Brooks, pers. Comm.).

Considerable attention has been devoted to possible refinements of the MN detection methods. Especially the involvement of automatic scoring and quantification of micronuclei traits by use of image analysis (IA) protocols appear to be interesting and was suggested already by Fenech et al. (1988). Method development and testing activities for such IA based MN analyses is ongoing at NIVA, and preliminary proof shows it is feasible to automate the scoring of MN in both fish and mussel samples (S. Brooks, pers. comm.). Other automatization strategies have employed flow cytometry-

based scoring methods or Coulter Counter assays (Roman et al., 1998; Avlasevich et al., 2011). The rationale for developing automatized MN scoring assays is because manual MN-scoring is very time-consuming procedure, measurements must be performed blindly by anonymizing and randomizing of samples, and because microscopy work in the field (on a vessel at sea) is both difficult and challenging.

As MN are relatively rare events, an accurate automatized determination of their frequency clearly benefits from interrogating more cells per replicate. Based on practical considerations, microscopy and image analysis tend to limit analyses to 1000–2000 cells per replicate, whereas flow cytometry provides the opportunity to readily acquire MN frequency information based on ≥5000 cells, usually in timeframes of 2–3 min (Avlasevich et al., 2011). This potential to increase the number of cells evaluated and thereby enhance statistical power may be especially important for work with weak genotoxicants or even potent MN inducers when one is interested in carefully defining the low end of the dose–response curve. However, the reliability of a flow-cytometry method for mussel samples with different cell types must be thoroughly validated before it can be considered as ready for use. Regarding QA of MN quantifications in blue mussels, recent activity (in 2018) under the biomarker division of BEQUALM included micronuclei assessment in mussel haemocytes based on the protocol by Bolognesi and Fenech (2012). Nine laboratories participated in the intercalibration exercise and there was reasonable agreement between laboratories in the frequency of micronuclei, nuclear buds and binucleated cells when normalised to 1000 cells (ICES, 2021).

The ICES WGBEC group has defined background assessment criteria (BAC) for the micronucleus assay in blue mussel haemocyte cells to 2.5 ‰, with low values indicating healthy specimens. See also BAC guidance data for the micronucleus assay with other marine species later (in Table 13).

2.2.9 Lysosomal membrane stability in blue mussel

In mussels, macrophage like cells (haemocytes) are important for the general defence system. These cells contain many lysosomes, i.e., membrane-surrounded pockets of digestive enzymes assisting in the combat of virus and bacteria, for removal of harmful foreign substances as well as in the breakdown of feed and recycling of the cell's contents. A decrease of the lysosomal membrane stability (LMS) in haemocyte cells is often observed in mussels exposed to chemical pollutant stress. LMS in blue mussel is one of the most frequently used biomarker parameters used in conjunction with marine pollution effect studies. Many studies have found lowered LMS in mussels from polluted, urban- and industrial-related areas when compared to individuals from control, clean seawater sites.

Two main methods for assessing LMS in mussels: and an *in vitro* cytochemical Neutral Red Retention Time (NRRT) method using haemolymph cells (Lowe et al., 1992) (Figure 4) and an enzyme cytochemical method (lysosomal labilization period assay) which involves the use of cryostat sections of digestive gland tissue. With the classical manual microscopy based NRRT assay, the LMS value stability of the lysosomal membrane is estimated based on the ability of the vesicles to retain the toxic neutral red (NR) dye before they break. Cells are extracted from the test animals and incubated with NR which accumulates in the lysosomes of the cells. The assay value to be recorded is normally the time response when dye loss is numerically assessed in 50% of the cells. Analysis protocols for both alternative procedures are described in an ICES TIMES report developed by Martínez-Gómez et al. (2015), with which the NRRT assay representing an upgraded procedure as it also considers the size of the lysosomes. In WCM surveys, LMS in macrophage-like mussel cells have typically been

done by means of the Neutral Red Retention Time (NRRT) assay, but more recently the analyses are preferably done with the cryostat section method of digestive gland samples partly due to significant challenges with analysing NRRT with microscopy on board a survey vessel in rough seas (S. Brooks, pers. com.).

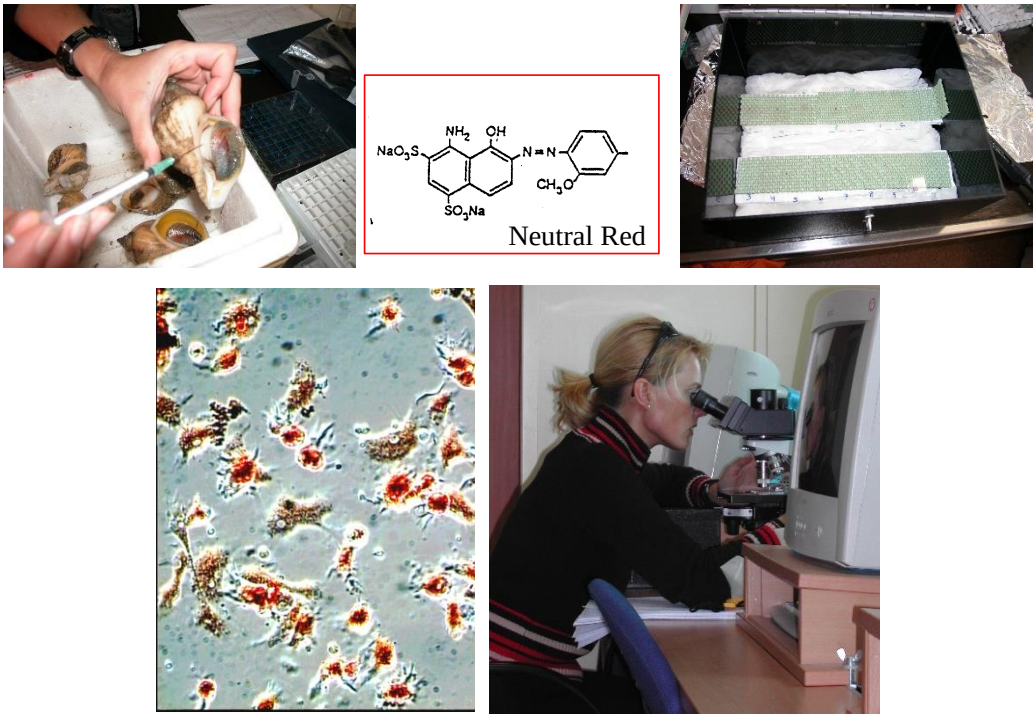

Figure 4: Analyses of lysosomal membrane stability (LMS) in macrophage like cells in invertebrate haemolymph with NRRT assay using light microscopy scoring procedure. Haemocyte cells are extracted from the mollusc test animal (upper left) and incubated with Neutral Red dye in an incubation chamber (upper right). The dye is taken up by the cells and accumulated in the lysosomes (lower left) and LMS is assessed as the EC50 time, i.e., the number of minutes before 50% of the cells have changed colour and shape.

The NRRT method has been validated through several international biomarker monitoring programmes (e.g. BEQUALM, BEEP, etc). ICES WGBEC has established BAC and EAC values of LMS in blue mussels of 120 and 50 min when measured with the NRRT method (Vethaak et al., 2017), i.e., suggesting non-stressed specimens having NRRT of 120 minutes or more, moderately stressed mussels having 120-50 min, whereas specimens with NRRT of 50 minutes or less being severely stressed. Spawning animals should not be included in any tests, since NRRT will normally decrease dramatically at this time. During spawning, the lysosomes are enlarged and will generally start leaking more rapidly compared with non-spawning mussels.

The ICES WGBEC group has defined the Background Assessment Criteria (BAC) and Environmental Assessment Criteria (EAC) for the NRRT assay in blue mussel haemocyte cells to 120 and 50 minutes, respectively. Likewise, the BAC and EAC for lysosomal labilization period in blue mussel have been defined to 20 and 10 minutes, respectively. For both lysosomal parameters, high values are indicating healthy specimens.

2.3 Water column monitoring parameters in fish

2.3.1 Biological data and somatic-organ indices of fish

As supportive parameters to biomarkers, several gross health parameters are to be assessed in fish samples. The size of whole fish is to be assessed by detection of weight and length data. Body weight is usually measured as total weight including stomach and gut contents. If, however, the stomach is full of undigested food items that can be easily removed, these need to be excluded from the total weight. The gender of fish specimens is to be determined by use of macroscopic assessment, or possibly by involving the use of magnifier examination of gonads during fish necropsy, and the age of fish specimens is to be determined by use of microscopical examination of fish otoliths which are obtained during fish necropsy.

The liver somatic index (LSI) and the gonad somatic index (GSI) of fish specimens are to be expressed as the percent of the organ weight to the total somatic weight.

The condition index (CI) of fish specimens is most commonly determined by the Fulton's condition factor (*K*), but also often termed as *CF*, which is calculated as the gram-weight of the fish divided by the cm-length cubed and then multiplying with a scaling factor of 100 (to get it close to 1) by using the formula shown below. For key gadoid fish species used in offshore monitoring studies, such as Atlantic cod (*Gadus morhua*) and haddock (*Melanogrammus aeglefinus*), a value of *K* close to 1 is typically indicating a fish in healthy and "normal" condition (Nash et al., 2006; Hansson et al., 2017).

K = (gram weight / cm length^3) x 100

Weight is given in grams and length is in centimetres

2.3.2 PAH/NPD concentration in fish liver measured by GC-MS

Similar as for mussels, determination of PAH exposure markers is relevant for fish, but because fish have a much better capacity for metabolizing PAH, the PAH exposure signals in fish are typically different compared to mussels. According to the offshore guidelines, the concentration of PAH shall be measured in fish liver by use of a GC-MS procedure, e.g., Klungsøyr et al. (1988); Ehrhardt et al. (1991). Reference samples are to be employed to facilitate quality control and comparison of analysis data between study locations and survey years. Previously, concentrations of PAH/NPD were analysed in fish muscle samples, this is mostly due to food safety considerations. However, the PAH/NPD concentrations detected in fish muscle in the WCM and condition monitoring surveys were typically very low (most often below detection limits). In 2020, the requirement was therefore changed to apply to liver which in gadoids is very fat-rich and therefore could be more likely to accumulate the hydrophobic PAHs above detectable signals, and it is also seen as a supportive parameter for the detection of biliary PAH metabolites (see below).

2.3.3 PAH metabolite concentration in fish bile measured by GC-MS/LC-FD

The concentration of fluorometric PAH metabolites in fish bile (Figure 5) have in many studies been demonstrated to be a sensitive exposure biomarker for determination of recent PAH exposure of fish in field and laboratory based exposure situations (Beyer et al., 2010). In offshore surveys on the NCS, increased concentrations of bile metabolites of petrogenic PAHs (and alkylphenols) have been detected in fish caged as far as 10 km downstream from oil production platforms in the assumed direction of the PW plume (Aas et al., 2002b; Hylland et al., 2008). As fish have an efficient metabolic capacity for oxidising PAH parent compounds, and to transform them into more water soluble

metabolites which temporarily are accumulated in the bile fluid, the analysis of PAH metabolites in bile can have an advantage over PAH parent compound detection as a measure of recent and/or ongoing PAH exposure.

As required by the offshore monitoring guidelines, the concentration of PAH metabolites shall be measured in fish bile by use of GC-MS/LC-FD procedures and reference samples are to be employed in order to facilitate quality control and comparison of analysis data between study locations and survey years. Since the bile fluid is varying in concentration with the feeding cycle, the stomach content (full/empty) of fish specimens shall be registered during necropsy to support the interpretation of the maturity of the bile mixture.

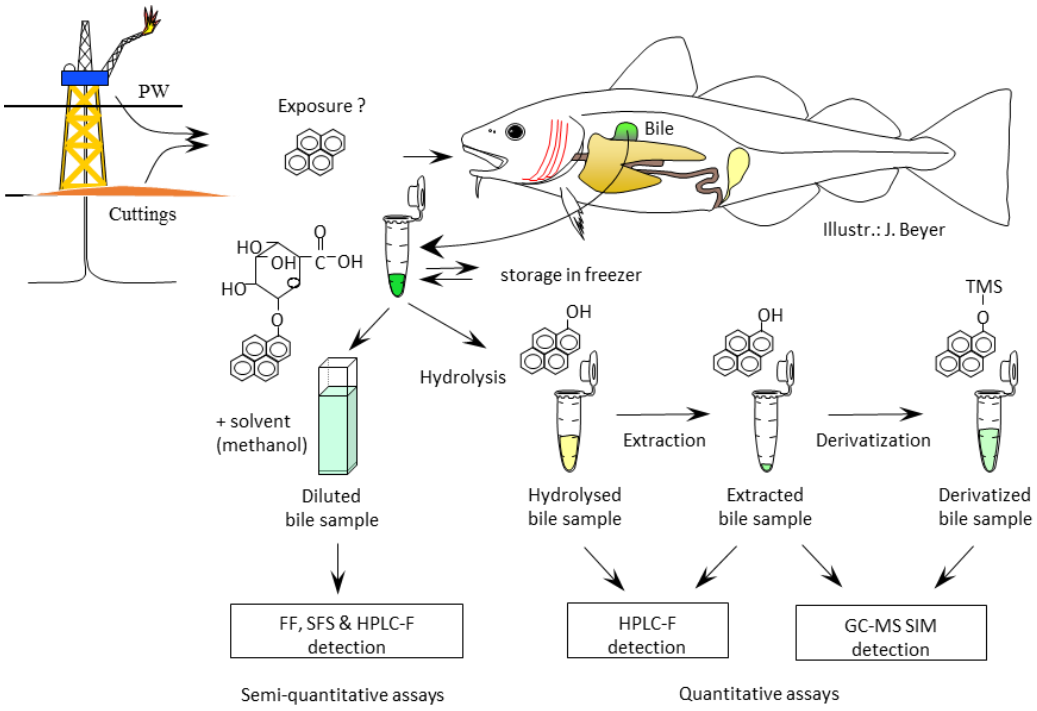

Figure 5: Overview of method alternatives for detection of PACs in fish bile, here with using the four-ring pyrene as an example PAH contaminant.

In vertebrates like fish, the bile is a major route for excretion of biotransformation products of many substances of endogenous and exogenous origin and certain constituents of the bile (steroid based bile acids and bile salts) emulsify lipids and thereby facilitate digestion of dietary fats and oils. Fresh bile is produced by the liver cells and is collected and temporarily stored in the gall bladder before it is evacuated into the intestinal tract after feeding. The bile is a key route for excretion of hydrophobic substances, including toxic organic contaminant compounds, after these have been metabolically transformed by enzymatic detoxification processes, for which the liver plays a major role. Typically, parent compounds of lipophilic pollutants are made more water soluble in the hepatocyte before the metabolites are secreted into the bile and temporarily stored and up-concentrated there. Quantitative detection of such biliary pollutant metabolites can serve as a sensitive proxy for how much parent contaminant have recently been taken up by the fish, either from the surrounding water or from contaminated food. Most commonly this biliary pollutant metabolite approach is used for estimating recent (or ongoing) exposure of fish to polyaromatic hydrocarbons (PAHs), and both semi-quantitative and quantitative methods can be utilised for this

detection (Ariese et al., 2005; Beyer et al., 2010). Biliary PAH metabolites are strong fluorophores so the analysis of bile samples can be performed with rapid fluorescence screening assays (fixed fluorescence assay or synchronous scanning fluorometry) or with HPLC-F analysis. These assays are only semi-quantitative, but experience has shown that they can still provide a fairly good indication of the overall PAH stress and the presence of major PAH metabolite classes in the bile sample (Aas et al., 2000; Jonsson et al., 2004). Interestingly, these rapid fluorometric assays even provide a good indication of the origin of the biliary PAH mixture, e.g. whether it is predominantly a petrogenic (oil-related) or pyrogenic (combustion-related) PAH mixture (Figure 6).

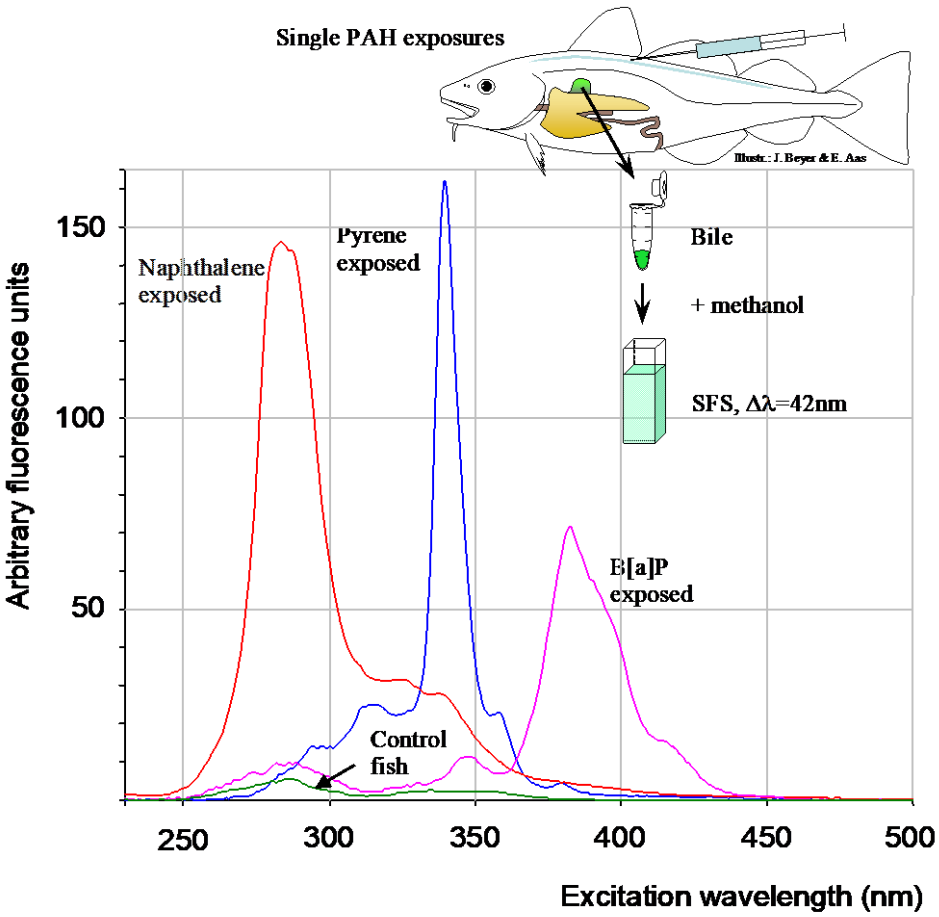

Figure 6: PAC metabolite specificity in SFS screening assays. Fluorescent metabolites of three separate parent PACs are discriminated to a reasonable degree by SFS analysis of diluted bile. The wavelength interval used in all scans shown above is 42 nm. SFS scan data obtained from (Aas et al., 2000) with permission.

2.3.3.1 Interpretation basis of petrogenic PAH metabolites in fish bile

If fish are exposed to some petrogenic PAH mixtures, i.e., PAHs derived from petroleum related sources, the biliary PAH metabolite pattern will normally be dominated by species with 2-4 aromatic ring structures (and alkylated forms of these) although certain special cases, like exposure to heavy weathered oil or heavy bunkers oil, can result in also larger PAH species (e.g. five ring structures) being present in significant amounts. For “small” PAH metabolites (e.g. naphthalene related metabolites) experience has shown that the GC-MS provides better metabolite separation and quantification power in comparison to HPLC based assays. For medium sized PAH metabolites (four

ring PAHs) both approaches are equally good. Whereas for larger PAH structures (five rings and more) HPLC is the best analytical approach (Jonsson et al., 2003). Some background data and response intensities, based on earlier project studies of PAH metabolites analysed by means of GC-MS determinations, are provided below (Table 2).

Efforts under the auspices of the ICES WGBEC group have addressed the development of assessment criteria (BAC and EAC) for biliary PAH metabolites in a set of marine sentinel fish species (Atlantic cod, dab, flounder, haddock) (Table 3). Semiquantitative and quantitative measurements of fluorescent bile metabolites in different marine fish species from presumed reference locations and/or presumed PAH exposed locations offshore at the NCS (e.g. close to offshore oil and gas production installations at the NCS) are available from a huge number of study reports and peer-reviewed articles, see and overview in (Beyer et al., 2020).

Table 2: Provisional response interpretation guide for petrogenic PAH metabolites in gadoid fish bile based on GC-MS SIM analyses of specimens collected at offshore oil fields.

Parameter	Value	Background signal	Trace signal	Strong signal
1-OH-naphthalene	ng / g bile	5	< 10	>20
2-OH-naphthalene	ng / g bile	5	< 10	>20
C1-OH-naphthalene	ng / g bile	250 – 300	< 1000	> 2000
C2-OH-naphthalene	ng / g bile	200 – 250	< 1000	> 2000
C3-OH-naphthalene	ng / g bile	400 – 600	< 2000	> 4000
1-OH-phenanthrene	ng / g bile	0	< 10	> 20
C1-OH-phenanthrene	ng / g bile	100-150	< 500	> 1000
C2-OH-phenanthrene	ng / g bile	70 – 150	< 500	> 1000
1-OH-pyrene	ng / g bile	0	< 10	>20

Table 3 Background assessment concentration (BAC) and environmental assessment criteria (EAC) for selected PAH metabolites in different fish species of relevance to the WCM program. Data reproduced from ICES website.

bile metabolite	species	common name	method	BAC	EAC
1-OH pyrene	Gadus morhua	cod	HPLC-F	21.00	
			GC-MS		483
	Limanda limanda	dab	HPLC-F	16.00	
	Platichthys flesus	flounder	HPLC-F	16.00	
	Melanogrammus aeglefinus	haddock	HPLC-F	13.00	
1-OH pyrene equivalents	Gadus morhua	cod	SSF	1.10	35
	Limanda limanda	dab	SSF	0.15	22
	Platichthys flesus	flounder	SSF	1.30	29
	Melanogrammus aeglefinus	haddock	SSF	1.90	35
1-OH phenanthrene	Gadus morhua	cod	HPLC-F	2.70	
			GC-MS		528
	Limanda limanda	dab	HPLC-F	3.70	
	Platichthys flesus	flounder	HPLC-F	3.70	
	Melanogrammus aeglefinus	haddock	HPLC-F	0.80	

- HPLC-F is high performance liquid chromatography - fluorescence, GC-MS is gas chromatography - mass spectrometry, and SSF is synchronous scan fluorescence 341/383 nm
- units are ng ml−1 for HPLC-F, ng g−1 for GC-MS, and pyrene-type µg ml−1 for SSF
- low values indicate healthy organisms

2.3.3.2 Uncertainty elements

The most significant confounding factor related to biomarker use of biliary metabolite detection in fish is the variation in the bile density during a feeding cycle and in feeding fish as compared to non-feeding fish (Table 4). The gall bladder is partly or totally emptied into the intestine subsequent to significant intake of food, and this food intake also triggers the hepatocyte to produce new bile (choleretica), which then is accumulated in the gall bladder, pending for the next food intake. Collier and Varanasi (1991) showed in studies with pollutant exposed English sole that the bile concentration of both natural bile constituents such as biliverdin and total protein as well as xenobiotic metabolites measured as fluorescent compounds (PAH-FACs) was markedly decreased in feeding fish as compared to non-feeding fish, and that the concentration of biliverdin, total protein and bile FACs increased during a starvation period of up to one month. Several later studies have confirmed these observations that during starvation, the total volume of bile, total amounts of PAH metabolites in the gall bladder, and biliary biliverdin concentrations all significantly increase, e.g. (Richardson et al., 2004).

The concentration of total protein and biliverdin tends to increase in bile of starving fish, especially during the first few days of starvation. Hence, it should be possible to normalise the concentration of xenobiotic-metabolites to the protein or biliverdin, i.e., to reduce the variance caused by the bile density factor. This kind of normalisation was suggested already by Collier and Varanasi (1991). But unfortunately, this kind of normalisation removes at best only parts of the variability, and in some situations, it apparently adds more variability to the statistical sample (Aas et al., 2000). Thus, biliverdin normalisation of metabolite data is not recommended as a standard procedure. Still, the biliverdin (or protein) concentration should always be measured in the bile sample and this information should always be considered as one of the explaining factors in the result interpretation of bile metabolite detections.

Table 4: Confounding factors of biliary PAH metabolite signals in fish.

Factor	Influence	Suggested actions
Increasing bile density during starvation	Increased metabolite concentration with increasing density.	Analyse bile biliverdin to estimate general density and consider these values during data interpretation. Avoid very different feeding status in study groups. Avoid comparing bile metabolite levels in bile samples that are significantly different in bile density.
Empty gall bladder	No sample	In laboratory studies, starve the study fish for a certain number of days (e.g. 4d) prior to sampling. In field studies, avoid taking samples of fish with empty gall bladders.
Very high bile density	Signal inhibition in screening assays at high bile concentrations (inner filter effect)	In fixed fluorescence and synchronous F scans, ensure appropriate dilution of bile prior to analysis

Since the gall bladder is partly or totally emptied in connection with feeding, it is recommended that fish that is used for bile metabolite detection have not been eating at least a couple of days (optimal 4 days) prior to necropsy. In field studies, this can be achieved by only using fish with full gall bladders for sampling. In general, a recommended procedure to minimise the influence of the bile density factor is, whenever this is possible, to make steps to ensure that all study animals have more or less the same starvation status. In fish that have not been feeding for a relatively long time (e.g. more than two weeks) the bile density is often very high, and the bile has a dark green colour. If fluorescence screening assays (FF or SFS) are used with such very dense bile samples the analyst should be aware of the potential influence of inner filter effects that can depress the fluorescence signal (Aas et al., 2000). The procedure to avoid this problem is to utilise a sufficient dilution of the bile sample prior to the fluorescence detection. From our experience a dilution factor of 1600x in solvent (50% methanol) is most often sufficient, but in extreme cases a higher dilution-factor may be needed, see (Aas et al., 2000) for details.

2.3.3.3 Quality assurance

When applying PAH bile metabolite detection in monitoring studies, reproducibility and accuracy of the analytical data are crucial, so quality control measures need to be implemented. Laboratory reference materials (LRMs; often shortened to “reference materials” - RMs; also known as “control materials”) play an important role in every quality control scheme. They can be used for method optimization, for statistical purposes, e.g. long-term reproducibility checks, or for interlaboratory comparisons, such as the exercises performed under BEQUALM or QUASIMEME and under WGBEC activities. RMs can be pure compounds, calibrant solutions, extracts, spiked samples or real samples, each with different purposes. Various options for assuring the quality of PAH metabolite measurements in fish bile are provided in the ICES TIMES report by Ariese et al. (2005). Among the options to check for systematic errors of an analytical procedure, one of the best is to analyse a certified reference material (CRM) and check the results against the certified value. Previously, two CRMs of bile from PAH exposed fish were produced and made commercially available, one representing exposure to a typical pyrogenic type of PAH mix and one representing exposure to a petrogenic mix of PAHs (Ariese et al., 2004). Laboratories that aim to improve the quality of fish bile PAH metabolite detections as biomarker may adopt to the recommendations in Ariese et al. (2005).

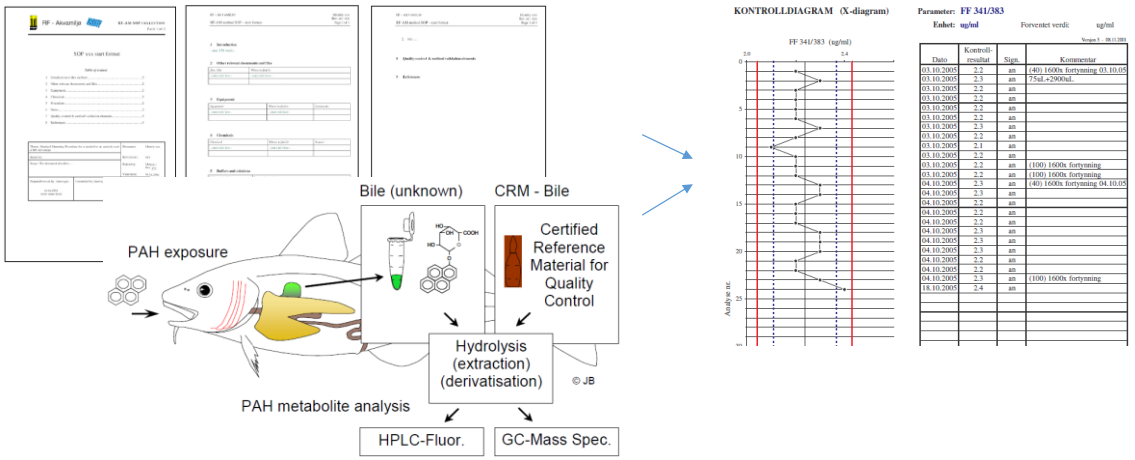

Figure 7: Biliary PAH metabolites in fish are among the easiest ecotoxicology biomarkers for which a combined access to biomarker SOPs and adequate reference materials will facilitates systematic method validation and analysis quality assurance and control measures.

2.3.3.4 Examples of biliary PAH metabolite detection in offshore field studies

Several studies on biliary PAH metabolites in fish at offshore oil and gas fields have indicated that fish in the water column around platforms are exposed to low concentrations of PAH pollutants; in particular when they are caged and thereby fixed in the plume downcurrent from the platform. The field work of the BECPELAG study (2001) was conducted at the Tampen field (Statfjord B) and in the German Bight. The study involved both caging of Atlantic cod and collection of native fish specimens. In bile from caged fish a weak but still very clear increase of the PAH metabolite level was detected with proximity to the Statfjord B platform (Aas et al., 2002b).

PAH metabolite measurement in fish bile has also been used in the two offshore Water Column Monitoring (WCM) studies that have been conducted so far. In the WCM studies, sub-groups of transplant study organisms (cod and blue mussels) have been located at specified distance to the study platform for a period of one month. After retrieval of the transplant specimens, a set of chemical markers and biomarkers in the organisms are examined for possible pollutant effects. The WCM 2003 survey was conducted at the Troll B platform and the results are reported by (Børseth and Tollefsen, 2004). The observed PAH metabolite levels in bile from the caged fish indicated a significant exposure at the stations downstream the platform (Figure 8).

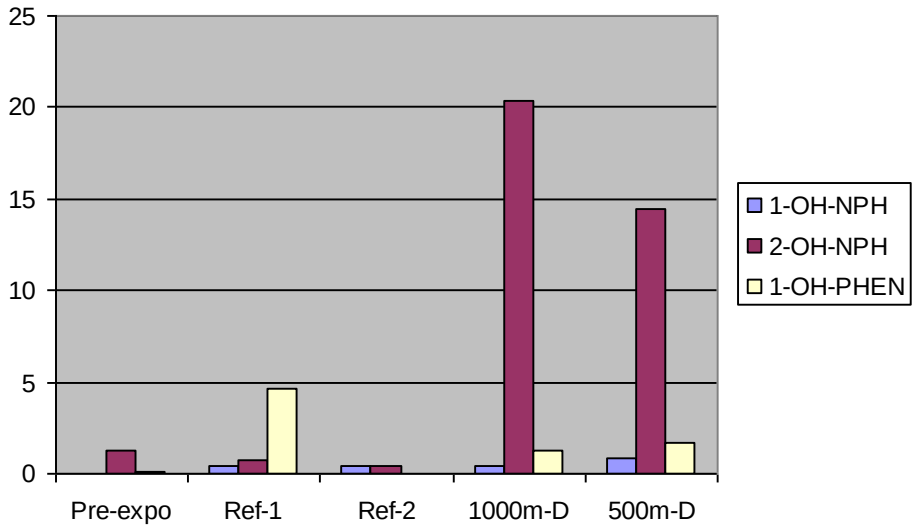

Figure 8: Quantitative PAH metabolites in bile from Atlantic cod caged for one month at the different stations in the WCM 2003 study. PAH metabolite results obtained from (Børseth and Tollefsen, 2004).

The species factor is essential for many biomarkers in fish, also for biliary PAH metabolites as indicated by a recent field campaign at the Frøy field in the Frigg area of the North Sea (Beyer, 2004). In this study a range of fish species were collected at an abandoned oil and gas production site. The results of bile analyses indicated that PAH metabolite level was generally higher in fish species which are more in direct contact with the sediment in behaviour and during feeding, e.g. haddock and flatfish species (such as lemon sole *Microstomus kitt* and witch *Glyptocephalus cynoglossus*), as compared to fish species which has a more pelagic behaviour, i.e., such as cod and whiting (Figure 9). With respect to the haddock, it is known that this fish species during the search for various burying invertebrate prey organisms ingests large quantities of sediment (Figure 10). This feeding habit obviously brings the haddock into very close contact with sediment particles and consequently also any lipophilic organic chemicals that might occur in the sediment. Organic pollutants are sometimes

significantly enriched in the sediment surface at sites around oil platforms. As judged from the species variation found in the Frøy study (Figure 9), judgements of biliary PAH metabolite concentrations could be feasible across species borders, i.e., between demersal species that in a sufficient degree share a common habitat (e.g. sediment-linked behaviour), but not between demersal and pelagic species.

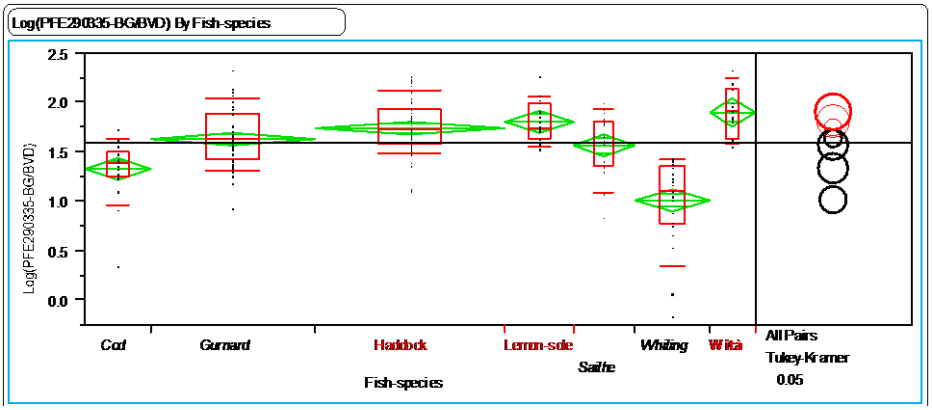

Figure 9: Overall trend in two-ring PAH fluorescence in bile samples from various fish species collected in the Frøy region in 2003.

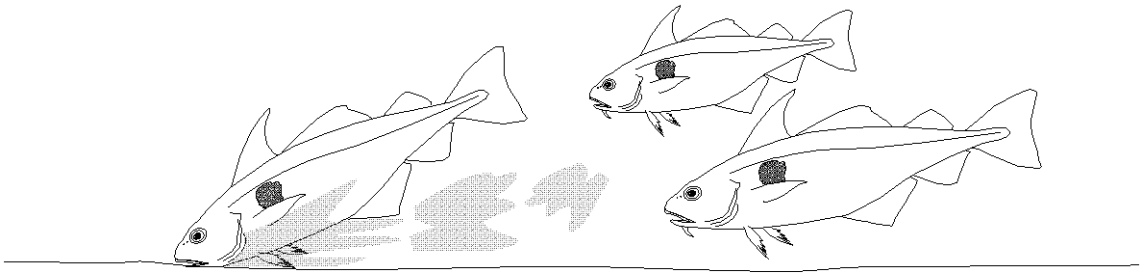

Figure 10: During feeding search for invertebrate prey on soft sediment bottoms the haddock will often rework the sediment surface and swallow considerable quantities of sediment material. In contaminated areas, this feeding habit brings the haddock in direct contact with pollutant chemicals in the sediment.

2.3.4 Histological changes, including lysosomal changes, in fish liver

The offshore monitoring guidelines state that liver of fish should be analysed for the parameter "tissue changes, including lysosomal changes" with use of "histology methods". It may be warranted that the guidelines can provide some more detailed advice regarding the possible choices of parameters, for example by referring to method advisory documents from relevant expert groups (such as ICES WGBEC) that may offer specific parameter selections and method descriptions on which parameters and methodologies to address. For example, ICES TIMES no. 36 would possibly be relevant in this regard as it include background descriptions and analytical protocols for lysosomal effect methodology in fish (and in mussels) (Moore et al., 2004). The most recent report from the WCM program (Pampanin et al., 2019), refers to (Bernet et al., 1999) and (Feist et al., 2004), as method resources for assessment of liver lesion parameters in fish, however, none of those include protocols for lysosomal biomarkers in fish livers.

The study by Bernet et al. (1999) describes a standardised tool for assessment of histological findings which can be applied to different organs, including gills, liver, kidney and skin. For liver samples, the study includes criteria for examination and diagnosis under categories on “early non-neoplastic toxicopathic lesions”, “foci of cellular alteration”, “benign neoplasms”, and “malignant neoplasms”. Descriptions of the normal appearance of liver tissue are also included.

The report by Feist et al. is an ICES TIMES method description document. It addresses the two marine flatfish species dab (*Limanda limanda* L.) and flounder (*Platichthys flesus* L.) and includes criteria for examination and diagnosis of the following liver parameters: architectural and structural alterations; plasma alterations; deposits; nuclear alterations; atrophy; necrosis; and vacuolar degeneration; two progressive changes (hypertrophy and hyperplasia); inflammation and tumour development (benign and malignant).

Many consider histological techniques and biomarkers to have a great (and largely unused) potential for determining health status and chemical stressor responses in ecotoxicological monitoring. Multiple changes to subcellular organelles and integrities of cells and tissues occur as sub-lethal responses to chemical and biotic (e.g. parasitic) stressors. Histology also provides possibilities of anchoring different sub-cellular and cellular perturbations in a chemically stressed bioindicator organism to organ and system functions at higher levels of biological organisation in the affected individual, ultimately providing links to fitness parameters (such as growth and reproduction), and also enabling insights into causes for population effects. In fish, the liver is often the target organ for histopathology assessments and several histopathological alterations in fish livers that can be induced by chemical stress are quite easy recognizable and scoreable. However, there are also several key challenges associated with the use of histopathological biomarkers in any bioindicator organism in conjunction with ecotoxicological studies. As discussed in Bernet et al. (1999), standardized methods for the assessment of normal histological features as well as different toxicant induced changes to different tissues are often lacking for the study species or the life phase addressed in the investigation. Generally, in all histopathology, it is key important that the normal range of variation as well as the possible confounding influence of different natural factors such as species, age, gender, sexual maturity stage, sampling season, feeding, and others, are well understood and can be controlled for.

Regarding the WCM program, the different gadoid fish species that are addressed in the field surveys must also be the species targets that are used for the building up of required knowledge and measurement capability on different histopathology traits provided that these are proven (by controlled exposure effect studies) to have relevance (being responsive and sensitive enough) for the main contamination scenarios that are addressed and evaluated by the monitoring. A substantial volume of work must be carried out in order to achieve this goal.

2.3.5 CYP1A1-induction and CYP1A activity (EROD) in fish liver

The cytochrome P450 1A family (CYP1A) of enzymes is responsible for the initial metabolism of hydrophobic and planar organic pollutants such as multi-ring polyaromatic hydrocarbons (PAHs) like benzo[a]pyrene. When fish are exposed to such pollutants, CYP1A enzyme activity is induced in metabolically active cells and tissues, such as in liver hepatocytes and the endothelial cells that form the inner lining of blood vessels. Liver in fish is the most common organ for determination of CYP1A induction and subcellular fractions such as S9 supernatant or microsomes are used for CYP1A mediated enzyme activities such as the EROD-activity or for determination or CYP1A proteins by ELISA or by quantitative PCR (qPCR) analyses of AH-receptor gene and other CYP1A1-gene expression products.

Induction of hepatic EROD (7-ethoxyresorufin O-deethylation) activity is indeed one of the most frequently used biomarkers in fish. Numerous studies have concluded that EROD induction in fish liver is a feasible response parameter for scaling the toxic potency of planar organic pollutants such as polychlorinated dioxins, furans, biphenyls and carcinogenic polyaromatic hydrocarbons. Nevertheless, as a biomarker parameter EROD induction is clearly hampered by the potential influence several confounding factors, including variation related to species, season, gender, sexual-maturation, temperature adaptation, inhibition and possibility of disappearance of signal in cases of improper treatment of samples. These factors should as far as possible be controlled for when EROD is used as a biomarker. This can be difficult to fully achieve in field investigations. For example, if native fish are collected at several locations in a recipient in which the water column is stratified, such as in an estuary, the confounding influence on the EROD levels caused by variation of the water temperature cannot be ruled out. In this case, the use of caging exposure of the fish would sometimes be a better approach. It is also clear that the EROD parameter is generally better suited for laboratory studies where the control of the confounding factors normally is better than in field studies. Since the hepatic EROD level in fish can be influenced by confounding factors, natural variation as well as by certain assay-factors (e.g. use of plate reader instead of manually method), it is difficult to provide a simple list of typical background values. Nevertheless, an overview of typical EROD background data and response intensities based mostly on data from our own studies is provided below (Table 5).

The ICES WGBEC group have defined of background assessment criteria (BAC) for hepatic EROD activity levels in gadoids and other marine fish species that have been much used as bioindicators in pollution effect monitoring studies (Table 6).

Table 5: Provisional response interpretation guide for hepatic EROD (7-ethoxyresorufin O-deethylase) activity in juveniles of selected classes of marine fish.

Fish sample	**Value**	**Lower range of normal**	**Higher range of normal**	**Weak induction response**	**Strong induction response**
Liver PMS fraction, (post mitochondrial supernatant, e.g. S9 or S12)	pmol ethoxyresorufin /min/mg protein	Gadoids: 1 - 5 Eel: 1- 5 Flounder: 5 – 10 Dab: 15 - 20	Gadoids: 15 - 20 Eel: 15 – 20 Flounder: 40 – 50 Dab: 50 - 60	200% of control group average	> 400% of control group average
Liver microsomes	pmol ethoxyresorufin /min/mg protein	Gadoids: 5 - 10 Flounder: 10 – 20 Dab: 30 - 40	Gadoids: 30 - 40 Flounder: 80 – 120 Dab: 100 - 150	200% fold of control group average	> 500% of control group average

Table 6 Background assessment criteria (BAC) for hepatic EROD activity levels in various marine fish species. Data reproduced from ICES website.

species	common name	sex	matrix	BAC
Gadus morhua	cod	both	liver microsome	145.0
Limanda limanda	dab	female	liver S9	178.0
		male	liver S9	147.0
		both	liver microsome	680.0
Callionymus lyra	dragonet	both	liver microsome	202.0
Platichthys flesus	flounder	male	liver S9	24.0
Lepidorhombus boscii	four spotted megrim	both	liver microsome	13.0
Pleuronectes platessa	plaice	male	liver S9	9.5
		both	liver microsome	255.0
Mullus barbatus	red mullet	male	liver S9	208.0

- units are pmol min−1 mg S9 protein−1 or pmol min−1 mg microsomal protein−1 for the liver S9 and liver microsome matrices respectively
- there are no EACs
- low values indicate healthy organisms

2.3.5.1 Uncertainty elements

EROD induction in fish is vulnerable to the confounding influence of internal, external and temporal factors (Table 7), which all carefully should be controlled for in laboratory and monitoring studies. A detailed review of all relevant factors falls beyond the scope of this report, but a good overview is provided in (Bucheli and Fent, 1995).

Table 7: Major confounding factors of hepatic EROD induction in fish.

Factor	Influence	Suggested actions
Suboptimal sample handling	Decrease of CYP1A catalytic activity due to inactivation of CYP1A enzymes. Increase of P420/P450 ratio.	Implement quality assurance (QA) at all steps in the sample handling process, including preparation, storage and analysis. Evaluate P420/P450 ratio of cytochrome P450 in sample extracts.
Water temperature	Cold water adaptation in fish leads to increased hepatic EROD activity	In field surveys, always record the water temperature and consider this factor when interpreting CYP1A catalytic data.
Gender and sexual development status	Influence EROD in several ways.	Minimise as far as possible the difference between study groups with respect to these factors.

Biological factors that can influence the EROD activity include species, fish size and age, gender and reproductive development and status. External factors which may exert influence on EROD levels include handling treatment, contaminant exposure route, and the use of anaesthesia during capture. It is known that water temperature can drastically affect induction of EROD in fish, and this environmental variable is often a major confounding factor in field studies. A variety of chemicals and chemical mixtures are known to inhibit the induction of EROD in fish; including heavy metals, organometallic compounds, some polychlorinated biphenyl (PCB) congeners and organotin compounds, such as TBT. Due to the possible action of confounding factors, the presence or absence

of EROD induction in fish from a field site should always be interpreted with prudence. This also underlines the importance of addressing more than one study marker (chemical detection and biomarkers) in animals obtained from contaminated field locations. However, since EROD induction is well characterized in a number of fish species, the influence of the most important confounding factors can normally be controlled for by means of a proper study design.

2.3.6 Hepatic CYP1A1 induction assessed by qPCR of AHR and CYP1A1 mRNA

According to the offshore guidelines, quantitative PCR assay should be used to evaluate the expression in fish liver of selected biomarker genes such as cyp1a (Cytochrome P450 1, alpha) and ahrr (aryl hydrocarbon receptor repressor). However, there appear to be a shortage of open source method descriptions for these assays. Preparation of these procedural descriptions can possibly be achieved based on the study reported by Meier et al. (2020) (including the supplement method descriptions).

2.3.7 PAH associated DNA adducts in fish liver

According to the offshore guideline, PAH induced DNA damage in fish shall be determined by measurements of DNA adducts in liver tissue. Hepatic DNA adduct formation in fish measured by the 32P-postlabelling method has been widely demonstrated as a highly sensitive method for assessing carcinogenic DNA adduct formations and genotoxic stress of PAH contaminants in both field and laboratory exposure situations. However, due to the technical demands and challenges in performing the assay, only a relatively small number of laboratories outside the field of human health studies have ever established (and maintained) the capability needed for performing the 32P-postlabelling assay on a routine basis. During the recent 20 years, a clear decline can be seen in the number of ecotoxicological studies using DNA adducts as a PAH responsive biomarker in non-human samples. This trend is probably due to a combined effect of the high costs for conducting the 32P-postlabelling analyses, the relatively high amounts of the radioactive 32P isotope being required and a shrinking number of analytical laboratories that are willing and capable of performing the analyses on a routine basis. For considerable time, much effort has been made to try developing new methods for quantifying PAH associated DNA adducts, such as analyses based on different mass spectrometry (MS) alternatives. The aim has been to develop analyses that are less challenging and less costly than, but equally or more sensitive as, the 32P-postlabelling method. These MS based method developments have so far been most promising in relation to studies in human toxicology.

DNA adduct formations rests on the ability of some pollutant chemicals to be metabolically activated and subsequently form stable, covalent binding to the nitrogen bases of DNA; inducing abnormal bulky structures of the DNA double helix. Already in the study of (Brookes and Lawley, 1964) a direct relationship between the carcinogenic potency of different PAHs and their level of covalently binding to DNA was demonstrated, and the toxicological significance of bulky PAH DNA adduct formation in cancer development is well characterised in numerous studies, e.g. (Baird and Mahadevan, 2004). The development of the classical endonuclease P1 32P-postlabelling assay, was done in the lab of Kurt Randerath in the early 1980s (Gupta et al., 1982). In brief, the procedure involves first an extraction of the DNA from a tissue sample, followed by an enzymatic hydrolysis of the purified DNA extract to form mononucleotides. The non-adducted mononucleotides are then removed from the extract by the action of an endonuclease P1 enzyme. The remaining (adducted) mononucleotides in the extract are subsequently enriched and then radio-labelled with 32P followed by separation of the mixture of radiolabelled molecules in the extract by means of thin-layer chromatography (TLC), and

finally the amount of adducts are semi-quantitatively determined by detection of radioactive emissions from the TLC sheet by means of autoradiography (Figure 11).

In marine ecotoxicology studies and pollution monitoring, various species of fish have been the most commonly used sentinels for DNA adduct assessments, including several studies conducted within the WCM program, as recently summarised by Pampanin et al. (2017). The experience from these studies has produced insight for some key sentinel fish species regarding what are typical background levels as well as response intensities observed in different PAH associated exposure scenarios in the field or in the lab. An overview of DNA adducts in liver of relevant gadoid fish species is listed below (Table 8).

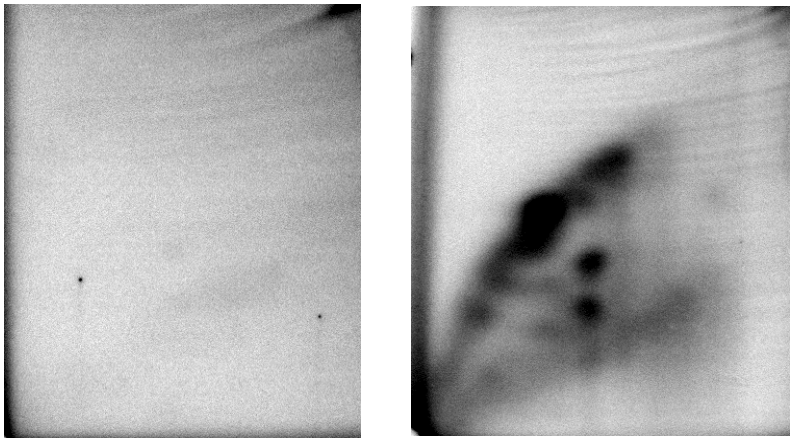

Figure 11: The autoradiogram is the critical endpoint of the 32P-postlabelling assay. The degree of blackening reflects the amount of DNA adducts in the analytical sample. In the case above, the results of one reference and one exposed sample of Atlantic cod liver are shown.

Table 8: Provisional response interpretation guide for interpretation of bulky DNA adducts in selected gadoid species as measured by means of 32P-postlabelling assay.

Fish species	Value	Background range	Trace signal range	Strong signal
Atlantic cod	nmol adducts / mol normal nucleotides	0 - 2	4 - 8	> 20
Polar cod	nmol adducts / mol normal nucleotides	0 - 2	3 - 5	> 10
Haddock	nmol adducts / mol normal nucleotides	2 – 10 *	10 – 15 *	> 30 *

* Haddock spot not subtracted (see discussion below)

Table 9: Assessment criteria for interpretation of DNA adduct levels in four species of marine fish as proposed by ICES working groups. Values are given for both background assessment levels (BAC) and environmental assessment criteria (EAC).

Fish species	Unit	BAC	EAC
Atlantic cod	nmol adducts / mol normal nucleotides	1.6	6
Haddock		3	6
Flounder		1	6
Dab		1	6

2.3.7.1 Uncertainty elements

DNA adduct formation is generally less influenced by confounding factors and uncertainty elements than many other biomarkers. Still, a certain influence may occur (see main factors in Table 10). The 32P-postlabelling assay is a very sensitive method for quantification of covalently bound pollutant chemicals in the DNA material. But, one disadvantage is the technical complexity of the method, and also the fact that it involves the use 32P, a potent radioactive isotope. The interpretation of the autoradiograms produced in the assay requires a substantial level of experience (and skills) with the method. Different analytical laboratories may utilise slightly different analytical conditions during DNA treatment, clean-up procedures, TLC-separation of nucleotides, interpretation of autoradiogram and adduct quantification. And this may in some situations lead to different results. In fish exposed to a mixture of putative adduct inducing compounds (e.g. a PAH mixture), it is also likely that the method underestimates the amount of adducts in the sample. This is due to the fact that the assay is optimised according to a DNA adduct of benzo[a]pyrene (7R,8S,9S-trihydroxy-10R-(N2-deoxyguanosyl-3´-phosphate)-7,8,9,10-tetrahydro-benzo[a]pyrene) (Figure 12) which is the only PAH-DNA adduct that is available for analytical standardisation.

Table 10: Confounding factors of bulky DNA adduct response signals in fish (see also Table 16).

Factor	Influence	Suggested actions
Different methods	Lowered comparability of analytical result data between studies or laboratories	Establish communication between labs. Implement the use of similar reference material and/or reference samples.
Fish age and size	Stronger signal with age and size	Avoid different age & size range between groups, test statistically effect of age/size within groups.
Fish mobility	Increased variability within study sites in field studies	Impossible to fully control for. Avoid the use of highly migratory fish species. Don't conduct field studies during spawning season. Ensure sufficient number of samples and consider excluding clear outliers.

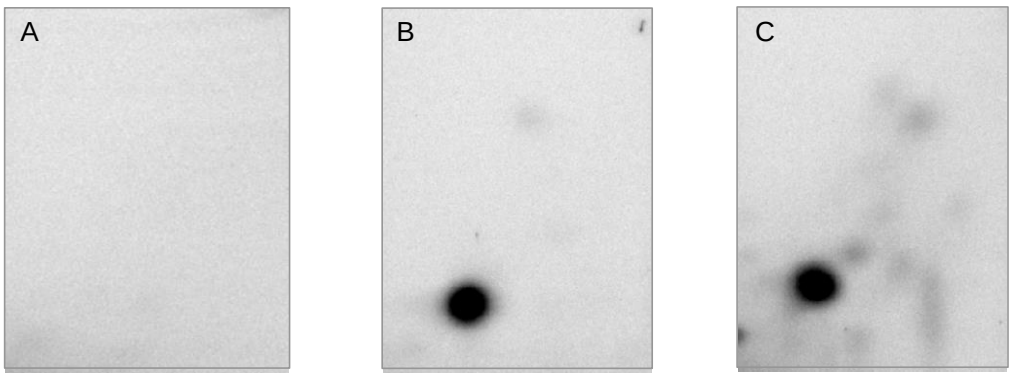

Figure 12: Standard and reference samples used for analyses of PAH-DNA adduct concentrations in fish liver by means of the 32P-postlabelling method. 32P-postlabelling autoradiograms analysed with negative reference sample (A) and positive standard samples (B), and a positive ring test reference sample (C). The Standard DNA-BaPDE adduct is 7R,8S,9S-trihydroxy-10R-(N2-deoxyguanosyl-3´-phosphate)-7,8,9,10-tetrahydro-benzo[a]pyrene.

The relation between hepatic DNA adduct level and age factors, e.g. illustrated by fish size or body length, is relevant to take into consideration at the interpretation of DNA adduct data. In prolonged exposures situations the amount of DNA adducts in liver tissue might be expected to accumulate over time. In field studies, it is often necessary to accept a certain variation of fish size, and this may add variation to a biomarker such as DNA adduct formation. In haddock collected at the Frøy region in 2003, there was apparently no clear differences in adduct levels between the three study locations but there was a weak correlation (r= xxx) between fish size and hepatic DNA adduct levels (Figure 13).

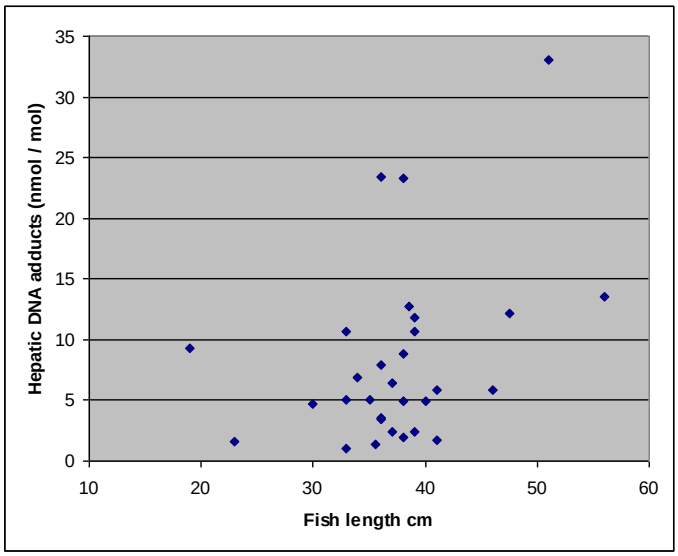

Figure 13: The relation between hepatic DNA adducts and fish length in haddock collected in the Frøy region in 2003. Results from (Beyer, 2004).

2.3.7.2 DNA adducts in haddock (*Melanogrammus aeglefinus*)

One relevant example to illustrate the importance of establishing similar interpretation of autoradiograms between different laboratories and studies is referring to the use DNA adducts in haddock as biomarker in field studies at offshore fields. From several recent studies in our group it is clear that haddock obtained from some non-polluted and pristine areas (Icelandic waters and Barents Sea) very often display a distinct adduct characteristic in liver tissue, appearing in 32P postlabelling autoradiograms as a strong and distinct spot outside the normal radioactive sector at the autoradiogram (Figure 14, see for example subfigure C). Interestingly, the haddock is also well known by its large black spot on each side of the fish body. The peculiar “haddock spot” in autoradiograms is not always present (e.g. Figure 14B). In some cases, the “haddock spot” may contribute with more than 10 nmol/mol of the total adduct level of the autoradiogram. This kind of autoradiogram signal in reference-fish is not observed in any other gadoid species or in any other fish species analysed (Lennart Balk, personal communication).

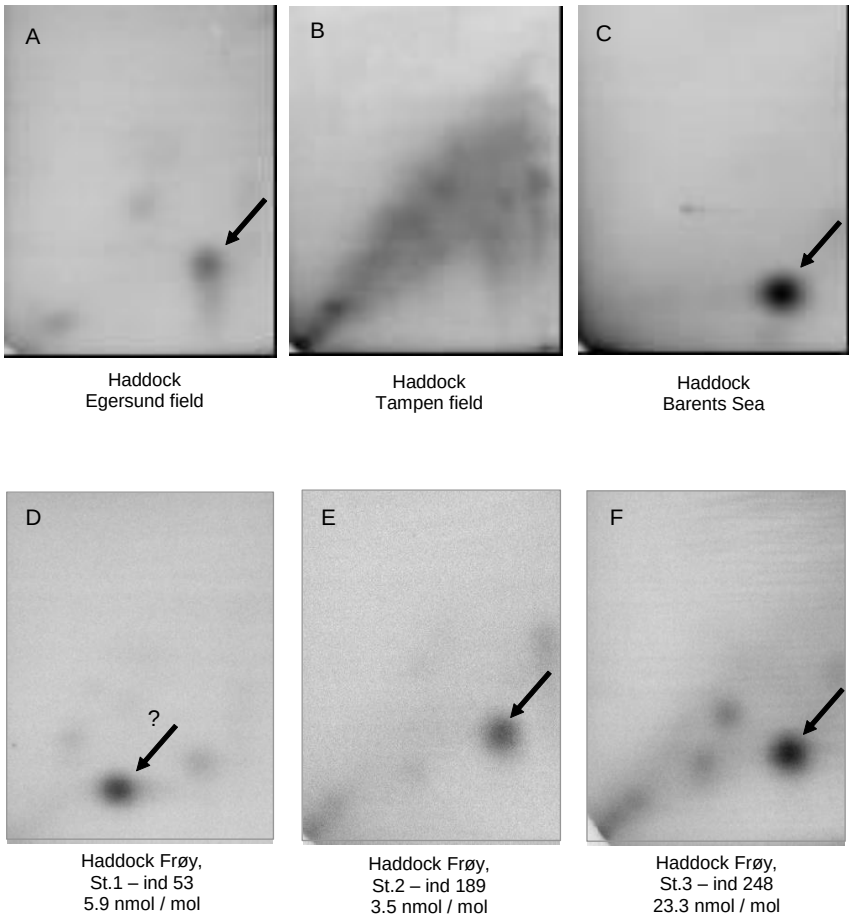

Figure 14 (A-F): A selection of 32P-autoradiograms haddock liver samples from different offshore locations. The presence of the "haddock spot" is indicated with an arrow.

It is no doubt that the "haddock spot" is representing adducted nucleotides, but the question is: what is the causal agent? Is it an endogenous substance or a foreign pollutant compound? One possible explanation of the "haddock spot" is related to the feeding behaviour of haddock on soft sediment bottoms, and the natural formation of certain adverse chemical species during the degradation of organic materials in soft sediments. As noted earlier in the bile metabolite section, the haddock is known to plough soft bottom sediment during prey search (Figure 10). In this process it swallows large quantities of sediment, so the intestine is entirely filled up. The organic matter which is being degraded in soft sediment would produce a range of organic chemical species as degradation products. Some of these species would contain aromatic structures that after being taken up by the haddock are being bioactivated and would then form stable adducts in DNA. In addition, due to its feeding behaviour, the haddock will come in much closer contact with these than any other fish species does. The contribution from the "haddock spot" results in a higher background signal in haddock than in other fish species (see Table 8). Anyway, a relevant technical question is whether the distinct area on the autoradiogram which is covered by the "haddock spot" should be excluded at the estimation of DNA adduct level in each haddock specimen? Presently, this issue is a matter of an ongoing discussion.

2.3.8 DNA strand breaks in lymphocytes (comet assay)

The alkaline single cell gel electrophoresis or comet assay is a technique for measuring DNA damage in terms of DNA strand breaks (Figure 15). The comet assay was originally described by (Östling and

Johanson, 1984) and later developed further by (Singh et al., 1988) for use on human cells, but the assay has also been applied on many test species including fish and aquatic invertebrates. A detailed protocol is available online from (Tice et al., 2000). The assay involves electrophoresis of cell DNA in an agarose gel under alkaline conditions. Since smaller DNA fragments migrate more rapid in the electrophoresis field the presence of many small DNA fragments will form a “comet with tail”. The level of DNA strand breaks is usually presented as the % DNA in comet tail, but tail moments and tail lengths are also used. The assay is considered to be a sensitive technique for the detection of non-specific DNA strand breaks. In human DNA, DNA breaks are detected in the level of 0.1 per 10^9 Daltons (Gedik et al., 1992).

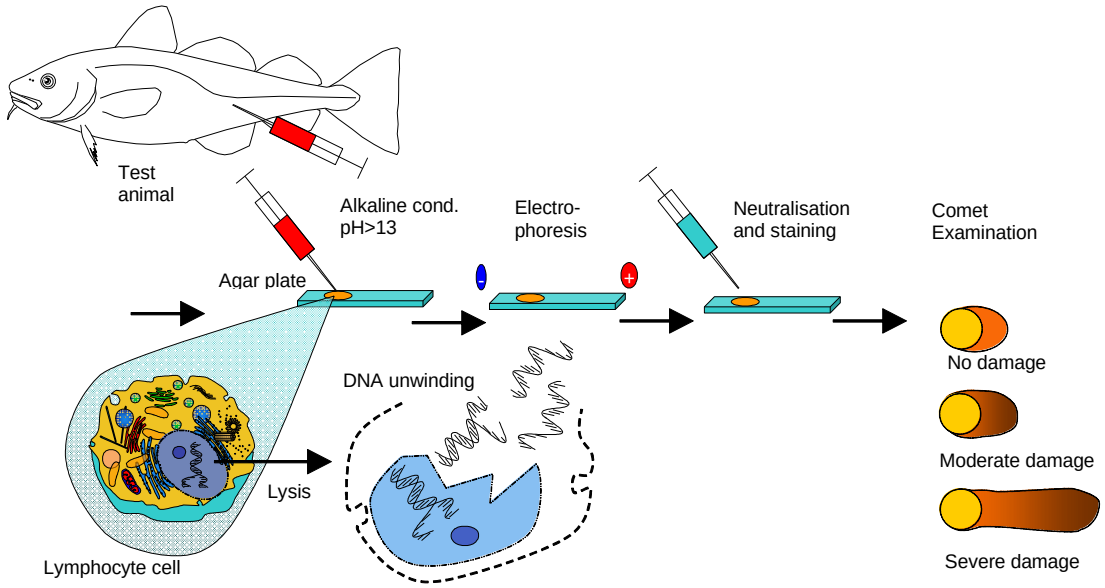

Figure 15: The principles of the single cell comet assay used for determination of DNA strand breaks. After ethidium bromide staining, an increased presence of small DNA fragments is visualised as larger comets with more DNA in the tail region.

The comet assay is used in a large number of studies with aquatic species (fish and invertebrates) and several reviews about this issue are available, e.g. (Lee and Steinert, 2003). Rather few studies, apart from those conducted in our research group, have addressed the effect of petroleum hydrocarbon type of exposures in marine organisms. Results from WCM relevant studies are reported by (Bechmann and Taban, 2001; Aas et al., 2002a; Bjørnstad et al., 2002; Taban et al., 2003; Taban et al., 2004; Baussant, 2004; Mamaca et al., 2005) and summarised by Hylland et al. (2006) and Brooks et al. (2011b). The findings in these studies are used for suggesting the guideline values shown in Table 11.

The ICES WGBEC group has defined background assessment criteria (BAC) for the comet assay in blue mussels to 10 %, and for measurement of marine fish (Atlantic cod and dab) to 5 %, with low values indicating healthy mussels.

Table 11: Provisional response interpretation guide for interpretation of comet assay results (% DNA in comet tail) in a selection of marine invertebrate species.

Species	**Cell type**	**Background range** % DNA in comet tail	**Trace signal** Relative to control	**Strong signal** Relative to control
Blue mussels (*Mytilus edulis)*	Haemocytes	7 - 12	110% (of mean value in control sample)	>130% (of mean value in control sample)
Scallop *(Pecten maximus)*	Haemocytes	9-12		
Icelandic scallop (*Chlamys islandica*)	Haemocytes	9-15		
Sea urchin (*Strongylocentrotus droebachiensis*)	Coelomocytes	10-13		
Mytilus edulis	Sperm	11-17		
Sea urchin (*S. droebachiensis*)	Sperm	9-10		
Corkwing wrasse (*Symphodus melops*)	Blood cells	7-13		

2.3.8.1 Uncertainty elements

In analyses with the comet assay, the background level which is detected in controls is relatively high in comparison to the signals observed in affected specimens. In our studies the baseline level is typically in the area of 10% when the % DNA in tail measure is used. For offshore monitoring surveys, such high background levels can be due to the challenges of sampling offshore. Electrophoresis cannot be performed on the vessel and the blood samples loaded on agar, are stored in lysis buffer for transport back to the lab. This transport step and the delay in electrophoresis may lead to increases in the background. In addition, different individual comets from the same study specimen always show a great deal of variation, a situation which normally leads to overlapping signal ranges in control groups and exposed groups. Consequently, it is important that the procedure include a large enough number of comet scores to ensure an appropriate discrimination power between study groups.

Several studies indicate that in situations of very long exposure, the responses measured by the comet assay may be reduced (Shaw et al., 2002; Large et al., 2002; Baussant, 2004). These observations could indicate that mussels may be able to adapt to the exposure possibly due to induction of repair mechanisms. Hence if no response is detected in wild organisms this does not necessarily mean that they have not been exposed. This also can be an argument for using caging instead of collection of native specimens in field studies.

2.3.8.2 Quality assurance

The scoring of DNA in comet tails is carried out by means of a computer-based image analyser software. As for MN scoring, automated scoring by computer assisted image analysis will soon be available for comet, which will dramatically improve the number of samples that can be run and improve data quality.

2.3.9 Micronuclei formation

The micronuclei (MN) test depends upon detection of membrane surrounded structures of enclosed chromatin occurring in the cytoplasm outside the nucleus after cell division (Figure 16). This kind of abnormal structure results from chromosomal breaks and dysfunction of the spindle apparatus and represents an irreversible genetic lesion. MN can be induced by chemical exposures. In the cytoplasm

of cells in interphase, micronuclei can be detected by normal DNA staining and microscopy examination of cell samples. An open source SOP that covers measurement of micronuclei formation in mussel and fish is in preparation and will be published through ICES TIMES (Brooks, pers. Com.), however, appropriate methodological information is already accessible from several literature sources, e.g., (Barsiene et al., 2004; Bolognesi and Fenech, 2012; D'Agostini and La Maestra, 2021; Ek-Huchim et al., 2022).

MN formation can be assessed by a simple microscopy-based technique involving adding blood or cell smears to a glass slide, drying, DNA staining and microscopy scoring. In blood cells, MN formation can also be assessed by means of a flow cytometric analysis. MN formation is used as a biomarker in regulatory drug testing for genotoxicity. In marine environmental studies, the MN method is used with cells from fish and invertebrates (preferably mussels). Various cells can be examined. In fish studies immature erythrocytes are normally the study cells (Figure 16). MN detection has been shown to be a sensitive tool for discerning exposed from control specimens of fish and mussels exposed to PAH, alkylphenols, crude oil dispersions and produced water mixtures (Aas et al., 2002a), (Beyer et al., 2005). The possibility of using an automatized image-analysis assay for measuring MN formation in fish red blood cells (RBC) smears is presently being explored by NIVA.

The ICES WGBEC group have defined of background assessment criteria (BAC) for micronuclei determination in a selection of marine species and cell types (Table 13).

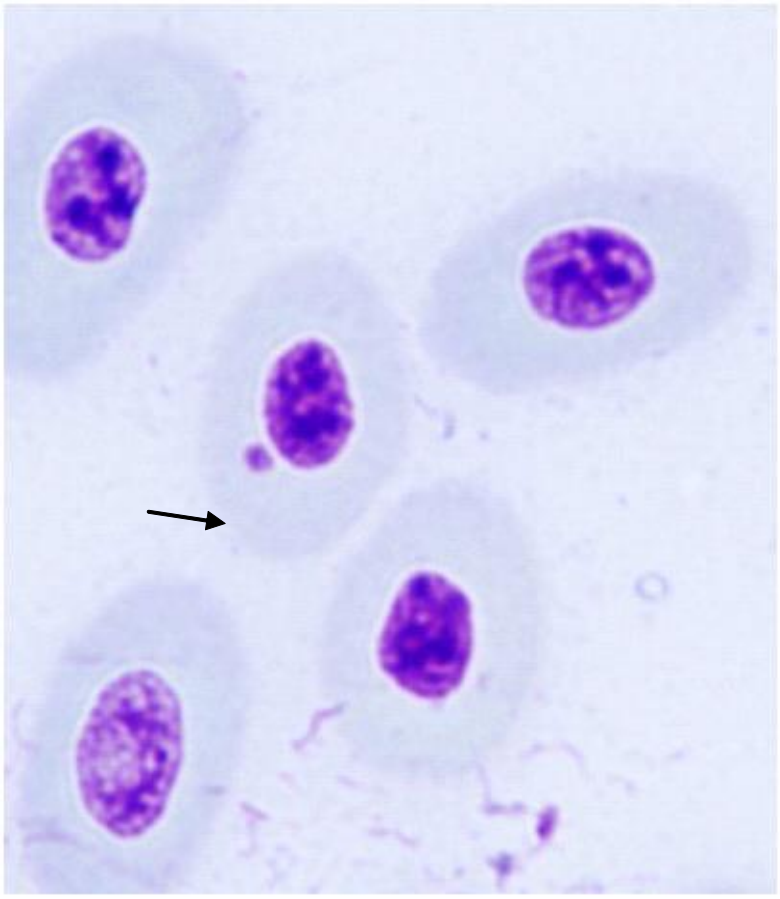

Figure 16. Micronuclei detected in immature erythrocytes in cephalic kidney of Atlantic cod. Results from Beyer et al. (2005).

Table 12: Provisional response interpretation guide for of micronuclei determination in a selection of marine species and cell types.

Species	Cell type	Background range MN / 1000 cells	Trace signal MN / 1000 cells	Strong signal MN / 1000 cells
Blue mussel	Gill cells	1-3	3-4	>6
Blue mussel	Haemocytes	1-3	3-4	>6
Atlantic cod	Immature erythrocytes from liver	0.1-0.3	200% (of mean value in control sample)	>400% (of mean value in control sample)
Atlantic cod	Immature erythrocytes from cephalic kidney	0.1-0.3		

Table 13 Background assessment criteria (BAC) for micronuclei determination in a selection of marine species and cell types. Data reproduced from ICES website.

species	common name	matrix	BAC
Mytilus edulis	blue mussel	gill	2.5
		blood	2.5
Mytilus galloprovincialis	Mediteranean mussel	blood	3.9
Mytilus trossulus	bay mussel	blood	4.5
Platichthys flesus	flounder	blood	0.3
Limanda limanda	dab	blood	0.5
Zoarces viviparus	eelpout	blood	0.4
Gadus morhua	cod	blood	0.4
Mullus barbatus	red mullet	blood	0.3

2.3.10 Acetylcholine esterase (AChE) inhibition

In all higher organisms, the enzyme Acetylcholine Esterase (AChE) is an essential enzyme involved in the transmission of nerve signals in the central nervous system, autonomic ganglia and neuromuscular synapses (Figure 17). The main function of AChE is to hydrolyse the neurotransmitter acetylcholine into choline and acetate in the cholinergic junctions. By hydrolysing the signal transmitter molecule acetylcholine AChE switches off the nerve signal and brings the synapse back to be ready for the next firing (normal speed for a cholinergic synapse is 1000 firings per second). In addition, new research has indicated AChE to have several important functions beyond neurotransmissions, e.g. (Zhang et al., 2002).

Many toxic chemicals are known to affect AChE based functions and can act as efficient inhibitors of the AChE activity. Best known are different chemical warfare agents and the organophosphate and carbamate pesticides which inhibit the catalytic function of AChE; an effect which in mild cases impairs a range of biological functions and in extreme cases leads to paralysis and death of the exposed organism. Exposure to organophosphate insecticides leads normally to an irreversible

inhibition whereas carbamate insecticides may lead to a slowly reversible effect. The concentration leading to a 50% inhibition of AChE is denoted IC 50.

Also common environmental pollutants such as petroleum related polycyclic aromatic hydrocarbons (PAHs) have been found to influence on AChE activities (Kang and Fang, 1997; Oropesa et al., 2007; Hauser-Davis et al., 2019; Fu et al., 2018; Olivares-Rubio and Espinosa-Aguirre, 2021), although these latter effect relationships are more unclear.

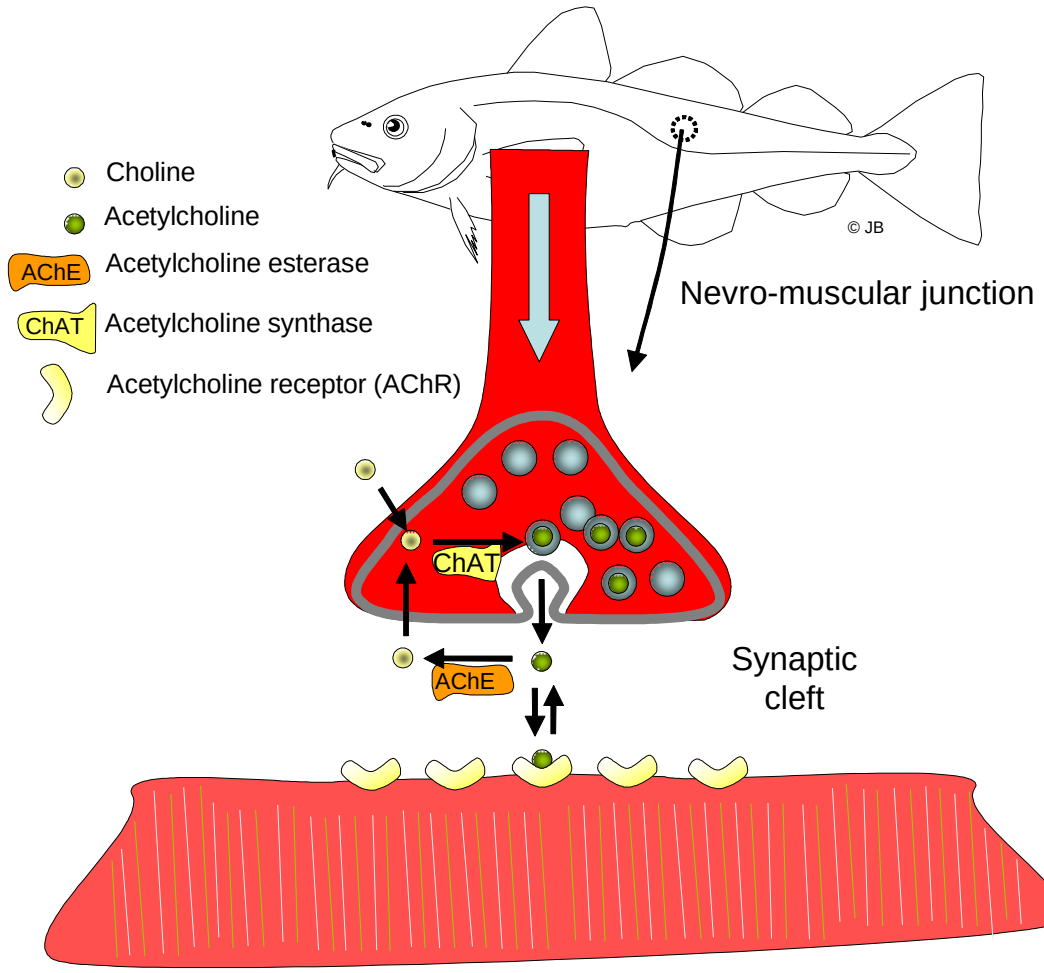

Figure 17: The basic principles related to the function of acetylcholine esterase in cholinergic synapses, such as motor end plates.

Pollutant induced AChE inhibition in fish is commonly assessed in samples of muscle or brain tissue by using the method of Ellman et al. (1961) and often with modification for microplate readers as described by Bocquene´ and Galgani (1998). By these procedures, the enzymatic activity of AChE is measured from the changed absorbance per minute of the assay mixture due to production of the yellow coloured 5-thio-2-nitrobenzoic acid (TNB) anion, which is based on coupling of the following reactions:

Acetylthiocholine (ATC) → thiocholine + acetate

Thiocholine + DTNB → 5-thio-2-nitrobenzoic acid (TNB)

The conversion of DTNB (5,5'-dithiobis(2-nitrobenzoic acid), also called Ellman's reagent) to TNB is used as a measure of the hydrolysis of ATC into thiocholine. ATC is produced from hydrolysation of the neurotransmitter acetylcholine by AChE. AChE inhibitors will induce a decrease in the production of ATC and therefore a decrease in the production of TNB will be observed. The change in absorbance per minute was used to calculate the AChE activity:

$$\text{AChE activity (µmol ATC/min/mg protein)} = \frac{[\Delta A \times Vol_t \times 1000]}{\varepsilon \times \text{light path} \times Vol_s \times [\text{protein}]}$$

Where ΔA = change in absorbance (OD) per minute at 405 nm, corrected for spontaneous hydrolysis, Vol_t = total assay volume (0.380 mL), ε = extinction coefficient of TNB (M/ cm), light path = microplate well depth (1 cm), Vol_s = sample volume (in mL), and [protein] = concentration of protein in the enzymatic extract (mg/ mL).

Table 14: Interpretation guide for interpretation of acetylcholine esterase inhibition in human medicine.

Depression in cholinesterase activity	*Effect condition*
15 - 25 percent	Mild poisoning effect
25 - 35 percent	Moderate poisoning effect
35 - 50 percent, or more	Severe poisoning effect

Table 15: Assessment criteria for interpretation of acetylcholine esterase inhibition in different marine species as proposed by ICES working groups. Values are given for both background assessment levels (BAC) and environmental assessment criteria (EAC).

Fish species	**Unit**	**BAC**	**EAC**
Blue mussels (Mytilus edulis)	AChE activity; nmol.min-1 mg prot-1	30 [1*]	21 [1*]
		26 [1**]	19 [1**]
Blue mussels (Mytilus galloprovincialis)		29 [1+]	20 [1+]
Dab		15 [1++]	10 [1++]
Flounder		235 [2*]	165 [2*]
Dab		150 [2*]	105 [2*]
Red mullet		155 [2+]	109 [2+]
		75 [3++]	52 [3++]

[1] gills
[2] muscle tissue
[3] brain tissue
[*] French Atlantic waters
[**] Portuguese Atlantic waters
[+] French Mediterranean Waters
[++] Spanish Mediterranean Waters

2.4 Passive samplers (PS)

As required by the guidelines, passive samplers (PS) should be used at all offshore stations when caged organisms are used, this to quantify exposure of organisms, to monitor the dispersal of pollution/discharges to the water column environment and to validate dispersal models. The deployment of PS should include sampling at several stations and at several depths and the PS shall as a minimum be used to analyse the content of THC, PAH, NPD and alkylphenols (AP).

In the ICES - TIMES (Techniques in Marine Environmental Sciences) issue no. 52, Smedes and Booij provide advice on the use of silicone rubber passive samplers for the determination of freely dissolved non-polar contaminants in seawater (DOI: https://doi.org/10.17895/ices.pub.5077) (Smedes and Booij, 2012). The level of detail offered may be helpful to first-time users of passive samplers, who wish to implement passive sampling methods in their monitoring programmes, and to more experienced users to review their current methods. The aspects covered by these guidelines include pre-extraction, spiking with performance reference compounds, deployment, retrieval, extraction, clean-up, chemical analysis, and calculation of aqueous concentrations.

2.5 Integrated monitoring approach in WCM

Integrated environmental monitoring is when several monitoring parameters are assessed in several ecosystem compartments at the same time in order to determine holistic environmental impacts associated with specific environmentally relevant human activity/operation, such as the discharge of offshore PW to sea. The use of integrated monitoring approaches has received increasing attention following the European Commission driven development of 'Good Environmental Status' (GES) principles under the EU Marine Strategy Framework Directive (MSFD) (Directive 2008/56/EC), which established the aim that the seas and oceans within the EU member states (and collaborating states, such as Norway) should achieve or maintain the GES requirements (described by 11 quality descriptors) by the year 2020. The principles of GES and the 11 quality descriptors required for meeting GES are explained in the OSPAR Quality Status Report (QSR 2010). With regard to integrated pollution effect assessment methodology, a special emphasis is often put on quality descriptor 8 (concentrations of contaminants give no effects) and considerable efforts in expert working groups under the auspices of ICES and OSPAR has been put into evaluations of which ecotoxicological effect markers (biomarkers) that can be considered as “ready for use” for assessing or monitoring the descriptor 8 “no effect of contaminant” GES requirement under MSFD and for Integrated Monitoring of Contaminants in general (ICES, 2011; Davies and Vethaak, 2012; Lyons et al., 2017; Vethaak et al., 2017). In this context, the development of internationally accepted assessment criteria for biological effects measurements using Background Assessment Criteria (BAC) and Ecotoxicological Assessment Criteria (EAC) (OSPAR, 2013) is key important. Presently, a revision of the biological effects monitoring guidelines from ICES is under preparation by members from ICES and OSPAR, and the next Quality Status Report for the OSPAR commission (QSR2023) is also underway.

The multifactorial scope of the WCM program opens for the involvement of an integrated monitoring strategy but the offshore monitoring guidelines do not presently provide any guidance to how this can be or should done. Particularly, it will be important to provide guidance on selection of sentinel organisms, target substances, and biological effects parameters as well as to provide user-friendly tools that enables a transparent integration and interpretation of the obtained monitoring data. Unfortunately, to discuss these quite complex issues and the needs associated with the guidelines falls largely beyond the scope of the present report.

2.6 Candidate monitoring parameters for WCM

The appendix of the offshore monitoring guideline provides a list of parameters (and methods) that presently are under consideration for being included as required/recommended parameters for the WCM program in the future, provided that further development and /or validation confirms their suitability. These candidate parameters will not be addressed in any detail by the present study, but it can still be worthwhile to provide an overview of them as methodological issues are parts of the validation work that will be needed. These parameters are:

- Automatized counting of MN, comet, and other chromosome damage in mussel and fish cells
- Acetylcholin-esterase (AChE) inhibition in mussels
- Hormone disruptive effects in mussel and fish assessed with qPCR
- Oxidative stress in fish
- Concentration of alkylphenol metabolites in fish bile assessed with GC-MS
- Inclusion of new sentinel organisms, e.g.: Calanus spp., Lumpfish (egg and larvae), king scallops (Bamber, 2021), sponges, earlier life stages of various species (fish, invertebrates), planktonic organisms
- Characterizing possible differences in responses between different *Mytilus*-subspecies and their hybrids
- Peroxisomal proliferation in *Mytilus* spp. and fish
- Cellular energy allocation (CEA) in mussels
- ALA-D enzyme activity in fish, red blood cells
- Vitellogenin induction in fish
- eDNA tools
- Improved detection of chemicals released in the water column
- In situ fluorescence of water and sediment for measurement of aromatic hydrocarbons from PW and old drill cuttings
- Automatization of biological measurements by image analysis (digital) and flow cytometry
- Improved measurements for health status in blue mussels
- Morphological changes of fish (haddock, cod)
- Histopathology of blue mussels and fish
- Filtering activity of blue mussels

3 Discussions of some relevant topics

In this report we have evaluated the available method descriptions and interpretation basis for biomarkers in marine ecotoxicity monitoring with special attention to the Norwegian WCM studies. Although much progress has been made recently on how to evaluate and monitor the possible impact of marine contamination situations by using biomarkers tools in sentinel organisms, it remains an open question whether the current strategies and parameters that are required/recommended in the WCM part of the Norwegian offshore guidelines represent the optimal choices for addressing the kind of offshore discharge situations addressed by these surveys. In the last section of this report, we address some key discussion topics for this area of study and identify also some remaining topics for future work.

3.1 Organisations and sources of biomarker method information

Standardisation/harmonisation of operational and analytical protocols is key for the diverse process of validating biomarkers in different bioindicator organisms to become trustworthy and informative tools/parameters that can be employed in pollution effect monitoring and in ecotoxicological investigations in general. The organisation and performance of such method quality activities has been addressed by different expert organisations and advisory bodies at national or transnational levels.

In Europe, the OSPAR Commission for the Protection of the Marine Environment of the North-East Atlantic (OSPAR) and ICES (the International Council for the Exploration of the Sea) are two major intergovernmental organisations which have been and still are particularly important for the long-term follow-up of method related needs of environmental monitoring methods and operations in marine systems, including biological effect measures and biomarkers. How OSPAR and ICES are expected to cooperate with each other and to execute their different administrative tasks and financial responsibilities is outlined in a Memorandum of Understanding document (OSPAR and ICES, 2006).

In brief, OSPAR is the mechanism by which 15 European governments & the EU cooperate to protect the marine environment of the North-East Atlantic. OSPAR started in 1972 with the Oslo Convention against dumping and was broadened to cover land-based sources of marine pollution and the offshore industry by the Paris Convention of 1974. These two conventions were unified, updated and extended by the 1992 OSPAR Convention. The new annex on biodiversity and ecosystems was adopted in 1998 to cover non-polluting human activities that can adversely affect the sea. Beside many other responsibilities, OSPAR issues publications comprising background documents, data reports and the results of evaluations and assessments of data reported to OSPAR by its Contracting Parties (i.e., the involved nations). OSPAR's Coordinated Environmental Monitoring Programme (CEMP) aims to deliver comparable data from across the OSPAR Maritime Area, which can be used in assessments to address the specific questions raised in OSPAR's Joint Assessment and Monitoring Programme, (JAMP). JAMP and CEMP have developed many different guidelines that provide detailed documentation of agreed monitoring and assessment methods, see inventory of guidelines https://www.ospar.org/work-areas/cross-cutting-issues/cemp.

ICES is a broad and transnational network of thousands of scientists working together in expert groups, workshops, and committees to address many kinds of human interference with marine ecosystems for meeting conservation, management, and sustainability goals. Of particular importance to the present report is the Working Group on Biological Effects of Contaminants (ICES-WGBEC) which examines the biological effects of contaminants in the marine environment, improves understanding and research of such pollutants, and aids international research and monitoring activities. Among the key tasks of WGBEC are to develop strategies for integrated monitoring of biological effects, providing advice on appropriate methods for monitoring, and increasing the fundamental understanding of ecotoxicological processes and how anthropogenic contaminants interact with environmental processes and marine systems. WGBEC initiates transnational research and monitoring and provides guidance to international organizations and conventions such as OPSAR, HELCOM, and AMAP, including in relation to the Marine Strategy Framework Directive (MSFD). ICES is also home to the ICES Techniques in Marine Environmental Sciences (TIMES), which are open access and peer-reviewed manuals and protocols of relevance to marine environmental research and monitoring. The TIMES manuals include many biomarker SOPs which have been developed and quality assured by experts in their respective fields (see also the links to selected biomarker SOPs in Table 1).

The Biological Effects Quality Assurance in Monitoring Programmes (BEQUALM) project is an example of an international biomarker Quality Assurance (QA) programme that with the scope of developing appropriate quality standards for a wide range of biological effects techniques and devise a method for monitoring compliance of laboratories generating data from these techniques for national and international monitoring programmes. The key goal of BEQUALM was to develop a QA system for biological effects techniques which would be self-financing on the basis of fees recovered from participants. A series of intercalibration exercises and training workshops have been conducted to develop the QA infrastructure, including measurements of: water and sediment bioassays, metallothionein, ALA-D activity, DNA adducts, P4501A activity, gastropod imposex/intersex, lysosomal stability (mussels and fish), fish liver histopathology and external disease measurement, chlorophyll-a and phytoplankton assemblage analysis and benthic community analysis. Further descriptions of the BEQUALM activities are available here (http://www.bequalm.org/about.htm). Given the ambitious goals of BEQUALM, the overall scientific outcome of the program may appear as small as judged from the limited number of biomarker publications indexed from Web of Science that refer to this program as the source for standardised method protocols, i.e., (Davies and Minchin, 2002; Koehler, 2004, 2006; Köhler and Ellesat, 2008; Lang et al., 2017; Straumer et al., 2020). However, the use of a broader search engine (Google) that covers also non-article literature returns a factor of ten more report units of various kinds referring to the BEQUALM program as a key source of QA relevant info on biomarkers.

3.2 Role of reference materials in biomarker QA and interpretation

A “biomarker response” can be defined as the situation when an observed biomarker signal (in a pollutant exposed group of a sentinel organism) deviates significantly from the level measured in a comparable non-stressed reference group. In general, the *Zero Effect Level* of a biomarker can be defined based on its appearance in a group of non-contaminated/non-stressed specimens, the latter representing the *Normal Range of Variation* (NRV) of the biomarker parameter in this sentinel species and under the specified study conditions used. In general, having a good insight about the signal range that can be considered as within NRV is important as a basis for the interpretation of biomarker signals in unknown samples in environmental monitoring. The importance of using NRV data as a criterion for ecological significance in environmental monitoring and was discussed by Kilgour et al. (1998). A simple strategy for using biomarker NRV as a measure for supporting quality assurance and signal interpretation in contaminant effect studies and contaminant effect monitoring surveys is illustrated below (Figure 18).

It is known that the NRV for biomarkers can often be influenced by different confounding factors, e.g. factors that can modulate the biomarker but that which are not part of the contamination issue studied. Moreover, it is important to clarify NRV for the selected target species at conditions which are relevant for the kind of study addressed, e.g. by using biologically and statistically comparable reference groups and by ensuring that the most relevant confounding factors are controlled for. Biomarker NRV can be influenced significantly by outlier data, e.g. cases when unaccountable high biomarker signals are detected in one (or a few) individuals of control or reference specimens. This happens both in laboratory exposure studies, but especially in field studies. With respect to the field situations, the outliers can often be explained by the action of some sort of confounding factor (e.g. by immigration of specimens in from a neighbouring area with different physical conditions or different pollution level). When outliers are included in the control basis data material, the NRV may expand significantly. How to best deal with this problem is a matter for discussion. However, it can be argued that the NRV should primarily visualise a *typical* NRV, e.g. as representing the variation

range detected in more than 95% of the non-exposed controls. For normally distributed biomarker data this range can be estimated statistically as two standard deviations of the mean.

When some sort of pollutant exposure triggers a biomarker response, this can be detected as an upregulation (induction) of the biomarker signal level in the case of a positive response mechanism. Correspondingly, a down-regulation (decrease or inhibition) of the biomarker signal level away from a certain normal control level is observed in cases of negative response mechanisms. EROD induction and DNA adduct formation are two examples of up-regulative responses, whereas lysosomal destabilisation and ACHE inhibition are two examples of down-regulative responses. Effect assessments using either of these two (positive & negative) effect approaches indeed represents most of the normal use of biomarker tools in pollution effect studies.

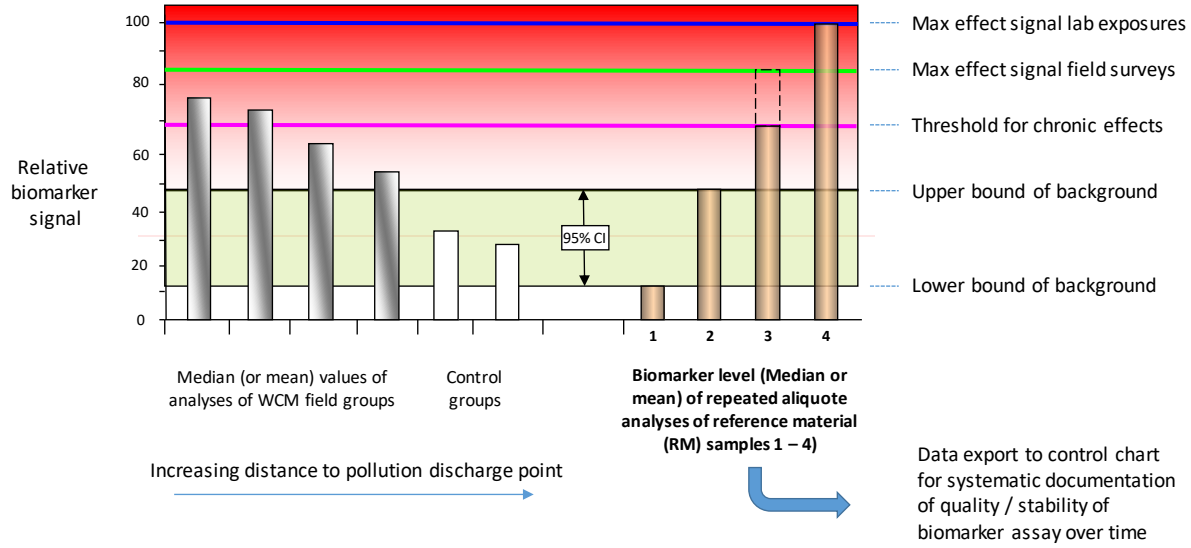

Figure 18: A strategy for using reference samples to visualize the Normal Range of Variation (NRV) when interpreting group means (or medians) of biomarker data from effect studies or effect monitoring surveys.

However, the simple fact that some biomarkers are regulated up and others down at a pollution effect is often confusing to many non-trained, and this underlines the need for straightforward interpretation (and visualisation) criteria, such as those shown in Figure 18, when biomarker results are presented. In characterisation studies of biomarkers, dose-response effects are normally studied in laboratory exposed animals; including typically also very high exposure concentrations of active chemical agents. Thus, the strongest biomarker responses are often observed in laboratory studies, and the maximum response level in such samples may therefore be considered to visualise the upper range of expected response for the biomarker parameter in question, and which provide a context for relative comparisons to biomarker responses observed in monitoring surveys.

3.3 Feasibility of making biomarker reference samples

Several biomarkers are measured by means of semi-quantitative assays, e.g. the semi-quantitative ELISA assays and fixed fluorescence screening of biliary PAH metabolites. Obviously, information on simple judgement criteria, such as NRV, is easiest to establish for quantitative biomarker measures; since the basis data in such cases can be collected from other laboratories and also obtained from scientific publications and reports. However, there are several pitfalls also for quantitative

biomarkers. International QA studies addressing ring-tests and inter-calibration of biomarker analyses have shown that different laboratories may obtain different measured values in biomarker assays of the same samples; based on various forms of differences of their analytical assays. Comparability of biomarker results between studies and between laboratories the analytical quality and standardisation of the biomarker assay used by each analytical laboratory is crucial. The need of an international effort to improve the quality control of different biomarker assays and to improve the general comparability of biomarker measures between laboratories is since long generally acknowledged among biomarker scientists. In connection with this; a significant bottleneck is the lack of proper standards for most biomarker assays.

The general lack of standards for the different biomarker assays can be based on a set of technical or practical reasons. However, to get around this problem, biomarker laboratories normally produce one or more reference sample in bulk; which is subdivided and stored in many aliquots under appropriate and stable storage conditions (e.g. at -80˚C). By means of such reference samples the stability of the biomarker assay over time and between different analytical personnel can be controlled. Whenever this kind of bulk reference samples is made by a biomarker expert laboratory the aliquot samples can of course also be shared between laboratories. Indeed, much of the inter-comparability need discussed above can be ensured by the by such collaborative actions. Indeed, reference samples that can be made available on request to analytical laboratories would also promote the process on international harmonisation of biomarker assays.

Optimally, such reference samples should be well characterised with respect to all relevant measures (stability of biomarker molecules, concentration in sample, etc.) prior to being offered to analytical laboratories. Experience have shown that testing, characterisation and internationally agreement of candidate reference samples can be achieved rather effectively in connection with international biomarker workshops or ring tests, or in international collaboration projects that primarily address more general processes of method improvements, harmonisations and standardisations. The results from such ring-test based activities can, as long as the activity is performed according to a pre-described quality assured system, form the basis of a quality certificate that can be issued by an international certification bureau, which also would provide interested laboratories with the certified reference materials (CRMs) on request.

When CRMs are made, the response levels can be designed in accordance with specific response level criteria. For example, it would be highly convenient if the biomarker levels in a set of available CRMs were comparable to the previously suggested interpretation criteria levels for this biomarker. Technically, such a set of graded reference samples is relatively straightforward to put together for most biomarker parameters, provided the existence in bulk of a maximum response sample. The reference samples beside this highest reference are made simply through a specific mixing with a comparable sample with minimum response levels. In comparison to the development of pure standards, this CRM set approach has several advantages. One major factor is the very high cost normally related to the production of pure standards, but another important factor is biomarker stability since some biomarker molecules are more stable when kept in their original matrix. For example, the latter is the case for conjugated aromatic pollutant metabolites in fish bile.

3.4 Biomarker uncertainty

For chemical detection of pollutants, criteria exist for ranging abiotic and biotic samples as being *non-polluted* or as having various degree of contamination. It is an aim to establish similar judgement criteria for biomarkers. However, this task is normally more challenging with biomarkers. Lack of biomarker standard materials and certified reference materials are general problems but also the

confounding influence on the biomarker level from various factors other than pollutant exposure complicates the judgement. Such factors can be nutritional status, gender, sexual development stage of the study animal, but also specific factors related to the surrounding environment where the study organism lives, e.g. water temperature and salinity.

The non-pollutant influence must be controlled for when a biomarker is used as a tool for pollutant effect studies. Some confounding factors are more generally relevant than others and some of the most relevant ones are summarised below (Table 16). In biomarker-based field campaigns (but also in laboratory studies), it is important that various steps are considered during the study planning process in order to minimise the influence of these general factors. Obviously, those biomarkers that are less influenced by confounding factors and that display a large (or characteristic) difference between the *normal* condition and a *stressed* condition are the ones where proper judgement criteria are easiest to decide. Influence on biomarker signals have been described related to species, gender, sexual maturation, ambient temperature adaptation, season and other factors. In marine field studies, the confounding factors may sometimes exert significant influence on the outcome of biomarker response levels detected in feral specimens that have been collected for the analyses. This can be a problem when mobile organisms such as fish are used. Optimally, the influence of all such non-pollutant factors should be well understood prior to using a given biomarker in any serious field monitoring activity. With such basis knowledge in hand much of the potential influence of confounding factors can be controlled for.

Table 16: Overview of factors which may exert confounding influence on the use and performance of biomarker tools in marine organisms.

Factor	**Problem**	**Suggested actions**
Different analytical methods	Lowered comparability of analytical result between studies or laboratories	Establish communication between labs. Share protocols and implement the use of similar reference material and/or share reference samples.
Spawning of study animals	Significant biological, biochemical, histological and anatomical alterations related to the ultimate phase of sexual development and spawning	Avoid using study animals in late phase of sexual development unless the involvement of such stages is an essential part of the effect study
Mobility of feral study animals	Increased variability within study sites in field studies	Impossible to fully control for. Avoid use of highly migratory fish species and studies during migration (spawning) season.

3.5 Biomarker relationships to exposure dose and time

Mechanistic studies of pollutant effects and investigations of dose-response relationships for biomarkers are usually done in the laboratory, either *in vitro* (exposed cells) or *in vivo*. Studies in animals may often utilise high concentrations of selected model pollutants (e.g. benzo[a]pyrene) administered by un-natural routes (e.g. intraperitoneal injections) and the exposure time is often short. Although high concentration/short-term exposure studies can yield valuable effect information, it can often be difficult to extrapolate the findings to responses observed under true field conditions. Furthermore, it is known that enzymes involved in compensating for exposure to

toxic compounds (detoxification and repair systems) can be induced (but also inhibited) depending on the concentration and type of contaminant and also the exposure time. In exposed animals, the biotransformation of the pollutant chemicals can start immediately following onset of the exposure. This metabolism can sometimes be boosted (induction), leading to a more efficient biotransformation and detoxification, and resulting in an enhanced clearance of the toxic agent from the organism. After end of exposure, the responsive system will typically return to normal after a certain recovery time. For enzymatic biomarkers in fish, this recovery time can last for some days and up to a few weeks, whereas some other biomarkers may require longer recovery periods.

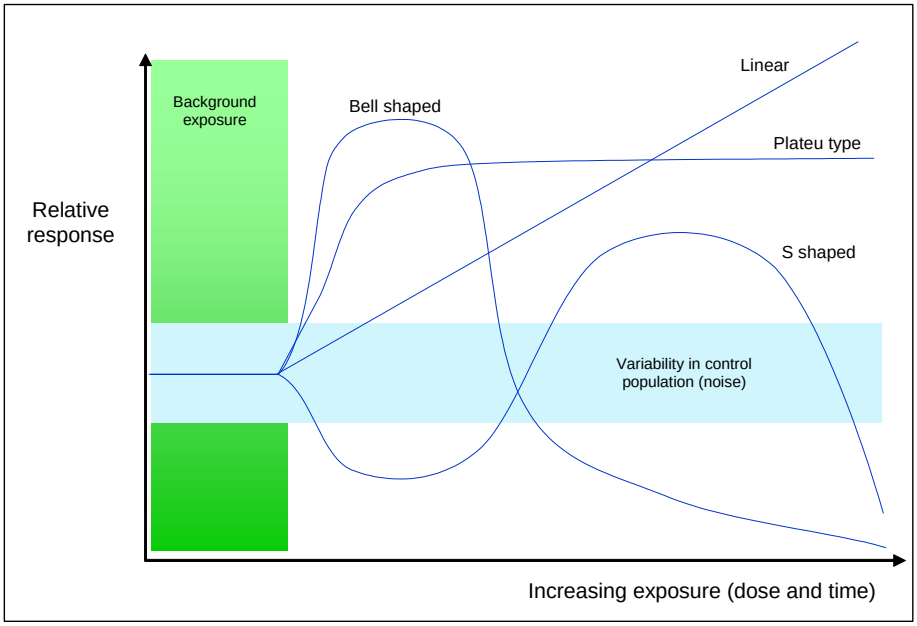

Figure 19: Theoretical linear and non-linear dose-response relationships for biomarker responses. The blue coloured areas represent the natural variability around the control level.

For biomarkers that have “positive” response direction, the response signal may often reach a plateau level at high pollutant exposure concentrations. Moreover, in connection with interpretation of biomarker results it is essential to be aware of the possibility that some biomarker system can be overwhelmed at very high pollutant exposure concentrations. This is the case in connection with enzymatic biomarkers that are induced by certain pollutant classes. Such non-linear dose-response phenomena are challenging when such biomarkers are used as tools for biomonitoring in highly polluted environments, or in laboratory studies using very high pollutant exposure concentrations. In Figure 19, this situation of non-linear dose-response relationship at high doses is illustrated; at relatively low concentrations the biomarker signal increase with the increasing exposure concentration before a plateau level is reached, and above a critical concentration (or dose) the signal is reduced again. For enzyme systems which are characterised by a certain constitutive normal level the signal level can at higher concentrations sometimes be reduced to below the control level, leading to an opposite result than anticipated. High-concentration depression of response signals is most relevant for detoxifying enzymatic biomarkers (e.g. CYP1A-EROD, Catalase, GST), but possibly also for other biomarkers that in different ways can be influenced by enzymatic processes (e.g. induction of DNA repair enzymes at high exposure concentrations may affect the level of DNA strand breaks).

Although non-linear dose-response relationship sometimes may complicate the interpretation of biomarker responses, this issue can normally be handled. Biology itself is seldom linear, and it is

often a matter of understanding the toxicology and biochemistry involved. Moreover, biomarkers will typically have a certain “response window” – a specific range in pollutant concentration and exposure time in which the biomarker signal can be detected and interpreted in a trustworthy and straightforward manner. Optimally, these criteria should be characterised in each study species used for biomarker studies, but sometimes, however, it would be sufficient if this basis information is described for a species that is closely related to the study species used.

It can sometimes be challenging to interpret results from biomarker analyses. Each biomarker may have a ‘window’ in concentration and time where it responds. The response may be linear or non-linear. Some biomarkers are responsive for a large range of concentrations, while others may respond in a narrower concentration range. Some biomarkers give a signal following only a few hours of exposure, others may need long-term exposure to give a signal. Some biomarkers are specific for certain compounds; like PAH metabolites in bile of fish. Other biomarkers are non-specific; alkaline unwinding and comet assay are non-specific genotoxicity biomarkers, while NRRT is a non-specific biomarker for cellular stress. Hence, a “no response” condition does not necessarily mean a “no exposure” situation or “no effect on fitness”. Theoretical response curves can be used for making testable effect hypotheses, and they can be useful when designing experiments for future effect studies where the aim is to learn more about biomarkers as early warnings of reduced fitness.

Biomarkers with non-linear dose-response curves may become significantly inhibited at higher concentrations. The main challenge in the interpretation of results from biomarkers with non-linear dose-response curves is to be able to tell when no response means too low exposure to give a signal, and when it indicates too high exposure – close to inhibition of the biomarker signal. It can be possible to interpret results from this type of biomarkers in laboratory studies, including frequent sampling and several exposure concentrations. With more detailed knowledge of the dose-response curves and possible time-related variation of response, this type of biomarkers can also be used in animals caged at different distance from the discharge point. When native animals are used for biomonitoring, biomarkers which respond only in a limited range of relevant concentrations and are time-dependent can only be used in addition to biomarkers with more linear responses and that respond in a wider range of concentrations and time.

Some biomarker signals may remain constant when the animal is continuously exposed to the same pollutant concentration. Other biomarker signals, and effects on fitness, may not be constant with increasing exposure time. Often experts in the biomarker field recommend the use of a suite of biomarkers for biomonitoring. When we search for useful biomarkers for biomonitoring studies of produced water discharges, we need biomarkers that respond at contaminant concentrations that realistically are encountered in the offshore water column at some distance from PW discharge points. Some biomarkers will not be sensitive enough to give a signal unless the animals are caged close to the platform or the fish have been swimming in the plume for the last couple of days. Other biomarkers give a signal at low concentration but may be inhibited at higher concentrations.

The assessment of short-term toxicity of pollutant chemicals is most often based on effect estimations in laboratory tests related to lethal concentration (LC) of test organisms or a dose response curve can be calculated from the LC test results. LC is the concentration of the test material that kills a specified percentage of the test organisms over the observation time. An example of how the LC is expressed would be a "48-hr LC_{50}." This is the concentration of the test sample that resulted in death of 50% of the organisms after a 48-hour exposure. A dose response expresses the response of the organism to a toxicant based upon body weight or dilution. The curve plots percent response verses log dose of the chemical. LC based studies generally operate at high exposure concentrations and it is widely acknowledged that LC data have only limited value in describing possible environmental impact of pollutants in field situations. Sub-lethal effect concentrations (EC) of pollutant exposure can be similarly described and here the biomarker approach is more relevant. In

tests of sub-lethal effects of pollutant chemicals, biomarkers can be used to identify effects of the exposure estimated in terms of a *no observable effect concentration* (NOEC) and a *lowest observable effect concentration* (LOEC); based on whether there is a statistical difference between unexposed control animals and the test animals. As many biomarker assays improve methodologically and as new biomarker parameters are developed these concentration values have been getting lower and lower. From a most conservative approach, any difference from the normal biomarker range can be regarded as undesirable, but in many regards this as an irrelevant standpoint as the exposed (and impacted) organism may not necessarily suffer any real harm from the effect condition. More recently, effect studies therefore aim at identifying the dose corresponding to the lowest observable adverse effect level (LOAEL) and the next lower dose, the no observed adverse effect level (NOAEL) for each effect and species. These two doses bracket the threshold dose/concentration - the dose/concentration below which no observable adverse effects occur. An adverse effect in this connection can for example be related to impaired reproduction or increased disease frequency (e.g. tumour development) (Figure 20). For biomarkers at this level of excellence the critical values for later harmful effects can be estimated, and the biomarker can act as an early warning tool for Environmental Risk Assessment studies of new chemicals or new field recipient situations. Very few biomarker parameters, if any, can presently be regarded as having fully reached this level. The linking of biomarker responses to fitness is discussed in more details below.

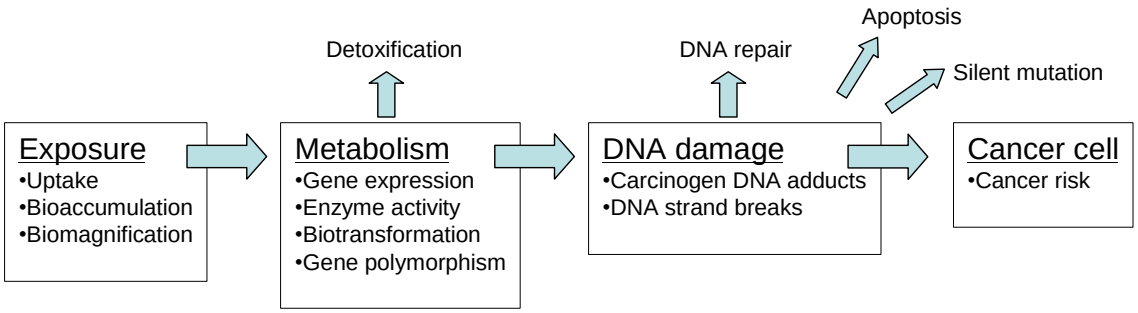

Figure 20: Exposure of vertebrate animals to carcinogens leads to a sequential row of events that eventually may result in the manifestation of a cancerous condition.

4 Suggestions for future research

4.1 Development of biomarker Certified Reference Materials

Several transnational projects and agencies have addressed quality assurance, international harmonisation and compilation of guidelines and protocols of biomarker analyses. These endeavours require much time and efforts and therefore also funding for system developments and maintenance. The relevance of such standardising initiatives is generally acknowledged among ecotoxicologists as intercalibration of analyses and inter-comparison of analytical results lead to improved biomarker quality and harmonised biomarker interpretation tools, which subsequently is a benefit of all parties. The availability of relevant and certified reference materials (CRMs) for biomarker analyses has been a significant bottleneck in the effort for ensuring better international inter-calibration and standardisation of biomarker analyses. So far, very few national or international programmes have succeeded in producing CRMs for biomarker analyses. A candidate CRM must be

made in bulk amounts, then it proceeds through a long process of characterisation and validation analyses before it eventually is certified and finally made available for interested laboratories through an international reference material bureau.

4.2 International harmonisation of biomarker interpretation

A collection of standard results background panels for assisting simple biomarker results interpretations, e.g. as those used in Figure 18, should be made with basis in the competence of an international group of expert laboratories. The potential outcome of such collaborative activities would be a harmonised interpretation basis which by all could be used as a standard background panel for biomarker data presentations. By using standard spreadsheet software this file can be circulated between biomarker expert laboratories and be progressively improved with time. In this way the biomarker interpretation file will be adapted to the state of the art within biomarker science. One biomarker expert laboratory at the time should have a special responsibility of upgrading this biomarker interpretation file as new and relevant information becomes available and also to ensure an effective dissemination of upgraded versions of the interpretation file to all interested laboratories and users, e.g. by means of a web-page based dissemination through the internet.

4.3 Biomarker response index

Some biomarker parameters are down-regulated when responding (e.g. lysosomal stability in invertebrate haemocytes or acetylcholine esterase inhibition in muscle cells) whereas others are up-regulated (e.g. EROD induction in fish). For some biomarkers a rather modest deviance from NRV is a strong response, e.g. such as a two-fold increase in hepatic GST activity, whereas for others a strong response is characterised with a much greater difference between stressed and unstressed conditions, e.g. DNA adduct formation in fish exposed to carcinogenic PAHs. To facilitate the interpretation of biomarker-based effect data one solution is to transform the measured biomarker values into a uniformed and relative response unit, e.g. by using a scale from 0 (normal condition) to 10 (strongest measurable response). In the case of a multi-biomarker study, an interesting side of this approach is the possibility of summing up response levels of many effect parameters into one single number, which popularly can be used as a *reason to concern index*, summarising the total effect pattern of the situation studied. This highly simplified approach could indeed have an interesting potential, but also several clear drawbacks. It would for example be more difficult to judge the quality of the reported data, partly since out of normal scale results can be hidden behind the relative units.

4.4 Link of biomarkers to fitness in exposed sentinels

Apart from a few biomarkers, the linkage between biomarker responses as an early level effect and later impaired biotic fitness is in general poorly described. Relating to individual organisms, the term *fitness* can be described as the individuals' ability and efficiency to exploit its environment and available resources in order to bring its genes further to the next generation. An improved insight about fitness relations may lead to an alteration of the ecological relevance of some biomarkers. Pollutant effects on impairment reproductive fitness have gained much attention, especially in studies related to human health, e.g. (Gaspari et al., 2003; Horak et al., 2003b, a). Studies to find links between pollutant responses on molecular biomarkers and later effects on impaired reproductive

capacity have so far failed to produce solid evidence. When the link between a specific biomarker response and later impaired fitness is discerned, and these findings also are sufficiently acknowledged in the ecotoxicological community, proper reference materials which illustrate critical threshold values of biomarker responses should be manufactured and made available at an international basis. A similar approach, as used for the more basal interpretation of guidelines above, can be adopted for fitness relevant effect levels of the biomarker signals. This level of expertise would clearly represent an advanced stage of biomarkers utilised as tools for environmental impact and environmental risk assessments.

In addition to having an interpretation basis for each biomarker in the WCM program, biomarker signals should sometimes be regarded in combinations, especially if several effect parameters together can provide information about the severity of an impact. Some combinations of responses are already known today. These will be described to form important contributions to the current biomarker interpretation basis. Eventually, this will be used to determine a biomarker-based strategy for future monitoring. One should be aware that biomarker-based exposure and effect monitoring potentially represents a cost-efficient approach compared to present monitoring methods, when applied in a similar fashion as a human health diagnostic tool. This is done by screening for possible effects, and in the cases where early warning signals are registered; the problem can be scrutinized in more detail with other methods. The monitoring methods should have high sensitivity to their relevant discharge situations to enable early warning of possible adverse effects.

Various biomarkers are used as tools within human epidemiology, environmental toxicology, cancer diagnosis and other human medicine disciplines. Much of the experience related to biomarkers in human medicine is actually directly applicable also for the development and use of pollutant responsive biomarkers in animals. One good example in this respect is the method for measurement of DNA adducts with the 32P-postlabelling assay, which initially was developed for diagnostic use in humans. Also, for the developments of judgement criteria for various biomarkers, there can be clear similarities and synergistic potentials between human health research and ecotoxicology research. Seen the other way around; discoveries in animal ecotoxicology research have also contributed to insight in human toxicology and health issues.

The analogy of animal biomarkers being in principle closely related to biomarkers used as effect parameters in human health tests helps in explaining the value of addressing these kinds of effect tools. Among biomarker researchers, there is an aim of developing non-destructive biomarkers, or effect parameters that are measured in samples obtained without killing the study animal.

5 Summary and conclusions

A combined access to internationally harmonised biomarker SOPs and biomarker reference materials (RMs) will facilitate biomarker method validations and method improvement activities and increase the value of biomarker use in effect monitoring and other ecotoxicological studies. Various suites of chemical markers and biomarkers in blue mussels and fish have repeatably been used in WCM surveys conducted at the NCS but a paucity of harmonised biomarker SOPs and biomarker interpretation involving RMs have constrained the information value of the large body of monitoring results from the WCM program.

Biomarkers have a valuable role to play in providing ecotoxicologically and ecologically valuable information offshore monitoring surveys and especially for informing on the validity of risk assessment predictions for discharges from offshore installations, ensuring that the risk predictions made associated with different discharge operations are accurate. Laboratory based biomarker

studies have the potential to further increase the relevance of these field surveys. Chemical monitoring of water bodies offshore can be time consuming and expensive. Simple and rapid exposure biomarker methods could provide an initial screen which acts to detect the presence of possible hazardous chemicals. The involvement of a suitable suite of biomarkers has the potential to gauge the health of native and caged organisms in the field and can provide a check on the water quality of the region. Significant deviations from the expected normal variability of health and fitness relevant biomarkers, characterised by means of biomarker reference materials can act as triggers for broader effect investigation on offshore biota. The use of WCM relevant biomarkers in laboratory-based effect assessments of simulated effluents and sediment contamination situations would provide more sensitive effect tests than those based on lethality and may also offer the opportunity for better understand the mechanisms behind putative ecotoxicity effects.

It is unlikely that more than a handful of individual biomarker techniques will prove suitable for the use in the WCM program. For EROD/ CYP1A biomarkers in fish, increased levels in enzyme activity or protein concentration signifies an increase in the biotransformation process associated with increased organic pollutant load. The consequences of increased biotransformation rate for sentinel specimens have not yet been fully assessed, although at present there is no obvious link to a long-term detrimental impact. Still, these particular biomarkers are useful for detecting possible exposure to organic chemicals and are useful in field monitoring situations where the risk of organic chemical spills poses a continuous risk and also in laboratory-based evaluations of these effluents. Genotoxic biomarkers, such as comet and DNA adduct assays, indicate possible damage to the study organism from adverse chemical exposures and such markers should potentially also be considered together with markers of DNA repair capabilities in sentinel study organisms. Biliary metabolites in fish provide a means for assessing the degree of recent exposure of fish to organic contaminants and are thus ideal for field monitoring purposes, perhaps acting as a screen in the same fashion as the EROD test. Lysosomal destabilisation assays provide an effect category that has been used successfully in many WCM field surveys and have also been used in laboratory-based exposure experiments.

There are three major areas where the measurement of biomarkers can make a valuable contribution to the WCM program; namely, laboratory-based studies, transplant operations (presently mussel caging) and evaluation of the health of local native fish populations. Laboratory based work permits investigation of the relationships between chemical exposure and biological responses in a controlled environment. The laboratory setting provides the opportunity to examine the linkage between biomarkers at relatively low levels of organisation, such as molecular and cellular responses, and higher-level population consequences as indicated by impacts on Darwinian fitness parameters (fecundity, fertilisation success etc). Once a clear linkage has been established then the biomarker measurement can be measured in the field and operate as a reliable surrogate for the ecologically important information needed for effective environmental management. This type of information is very valuable for environmental risk assessment procedures, providing data that is far more ecologically relevant than the acute toxicity test data that typically currently informs the process. Secondly, biomarkers used in transplant caging studies is presently involving mussels from a putatively clean areas that are transplanted in cages in the water column at different distances to offshore discharge situations for a predetermined period. In this active monitoring approach, biomarkers can be employed to determine the degree of exposure as well as biological responses. It is likely that both biomarkers of exposure and those of effect will have a role to play in this approach to future WCM monitoring. Thirdly, biomarkers in native fish collected offshore have a role to play in assessment of the health of fish stocks in areas downstream from oil and gas production installations. However, the key to the successful adoption of this form of monitoring lies in fully describing the normal background variation of the addressed biomarkers so that significant deviations from normal variability can be recognised and any links to long-term negative impacts can be identified.

The biomarkers addressed in this report constitute parameters required or recommended for offshore water column monitoring by the present edition of the national guidelines (Miljødirektoratet, 2020). The selection of monitoring parameters has evolved with time and revisions and was considerably reduced in the most recent edition. However, it remains uncertain whether the present procedures are optimal as there is a general paucity of data that confirm appropriate suitability for the special conditions that typically occur offshore. The collection, evaluation and discussion of biomarker SOPs in the present report can serve to provide insight on how different biomarkers are being measured and whether these biomarkers and analytical procedures have been shown to be fit for the purpose. Secondly, the possibilities for an improvement of the interpretation basis of these biomarkers is addressed, both regarding measures for validating their general suitability, their analytical quality performance and for establishing objective judgement criteria of the measured biomarker signals. Better knowledge regarding typical normal variability range of biomarkers (in non-polluted situations) is needed as well as better documentation of response intensities that can be expected in offshore relevant field scenarios. We suggest prioritising the development of open and internationally harmonised analytical procedures for all key biomarkers involved in the WCM program combined with development of adequate biomarker reference materials and associated and suitable quality-assurance and quality-control schemes for analytical determination in appropriate sentinels and samples. These reference materials should be made accessible (i.e., without huge costs) for any group that perform the analyses and interpretations of the different biomarker parameters. There is a substantial amount of laboratory-based research yet to be conducted before the full potential of this approach can be realised.

6 References

Ariese, F., Beyer, J., Jonsson, G., Porte Visa, C., Krahn, M.M., 2005. Review of analytical methods for determining metabolites of polycyclic aromatic compounds (PACs) in fish bile, *ICES-TIMES Techniques in Marine Environmental Sciences* **no 39**, 41 pp.

Ariese, F., Beyer, J., Wells, D., Maier, E.A., Gawlik, B.M., Lamberty, A., 2004. The certification of the mass fractions of PAH metabolites (1-hydroxyphenanthrene, 1-hydroxypyrene, and 3-hydroxy[a]pyrene in two fish bile materials: BCR 720, sediment-exposed flounder bile and BCR 721, oil-exposed plaice bile, RF-project 684827. European Commission, Directorate-General Joint Research Centre, Institute for Reference Materials and Measurements, Geel, Belgium, p. 50.

Au, D.W.T., 2004. The application of histo-cytopathological biomarkers in marine pollution monitoring: a review, *Marine Pollution Bulletin* **48**, 817-834.

Avlasevich, S., Bryce, S., De Boeck, M., Elhajouji, A., Van Goethem, F., Lynch, A., Nicolette, J., Shi, J., Dertinger, S., 2011. Flow cytometric analysis of micronuclei in mammalian cell cultures: past, present and future, *Mutagenesis* **26**, 147-152.

Baird, W.M., Mahadevan, B., 2004. The uses of carcinogen-DNA adduct measurement in establishing mechanisms of mutagenesis and in chemoprevention, *Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis* **547**, 1-4.

Bakke, T., Klungsøyr, J., Sanni, S., 2013. Environmental impacts of produced water and drilling waste discharges from the Norwegian offshore petroleum industry, *Mar Environ Res* **92**, 154-169.

Bamber, S., 2021. Introduction of additional invertebrate species to the Water column monitoring program (with a focus on the king scallop Pecten maximus), Stavanger, Norway, p. 41 pp.

Barsiene, J., Lazutka, J., Syvokiene, J., Dedonyte, V., Rybakovas, A., Bagdonas, E., Bjornstad, A., Andersen, O.K., 2004. Analysis of micronuclei in blue mussels and fish from the Baltic and North Seas, *Environmental Toxicology* **19**, 365-371.

Baussant, T., 2004. Biological effects of chronic levels of dispersed oil in North Sea and Arctic conditions: Biomarker responses and early life stage tests in two Mollusc species. Version 2. Akvamiljø, Stavanger, Norway.

Bechmann, R.K., Taban, I.C., 2001. The Comet Assay: Results from pilottests and from the BEEP and the BIOSEA prosject. Akvamiljø as, Stavanger, Norway.

Bernet, D., Schmidt, H., Meier, W., Burkhardt-Holm, P., Wahli, T., 1999. Histopathology in fish: proposal for a protocol to assess aquatic pollution, *Journal of Fish Diseases* **22**, 25-34.

Beyer, J., 2004. Frøy biomarker survey 2003, Projects 695018 & 695031. Akvamiljø as, Stavanger, Norway, p. 40.

Beyer, J., Goksøyr, A., Hjermann, D.O., Klungsøyr, J., 2020. Environmental effects of offshore produced water discharges: A review focused on the Norwegian continental shelf, *Marine Environmental Research* **162**, 20.

Beyer, J., Green, N.W., Brooks, S., Allan, I.J., Ruus, A., Gomes, T., Bråte, I.L.N., Schøyen, M., 2017. Blue mussels (*Mytilus edulis* spp.) as sentinel organisms in coastal pollution monitoring: A review, *Marine Environmental Research* **130**, 338-365.

Beyer, J., Jonsson, G., Porte, C., Krahn, M.M., Ariese, F., 2010. Analytical methods for determining metabolites of polycyclic aromatic hydrocarbon (PAH) pollutants in fish bile: A review, *Environmental Toxicology and Pharmacology* **30**, 224-244.

Beyer, J., Jonsson, G., Sundt, R., Westerlund, S., Larsen, B.K., Baussant, T., Hylland, K.D., Tollefsen, K.-E., Balk, L., Skarphedinsdottir, H., Liewenborg, B., Jenssen, B.M., Mascini, M., Bagni, G., Laschi, S., Bulukin, E., Barsiene, J., Cavcic, A., Stokke, T.U., Karlsen, K.A., Jørgensen, K.B., 2005. Pollutant exposure and effects in fish related to the discharge of produced water in the North Sea oil industry (poster presentation), NFR PROOF annual meeting, Solstrand, Norway.

Bignell, J., Cajaraville, M.P., Mariogómez, I., 2012. Background document: histopathology of mussels (*Mytilus* spp.) for health assessment in biological effects monitoring, in: Davies, I.M., Vethaak, D. (Eds.), Integrated marine environmental monitoring of chemicals and their effects. ICES, pp. 111-120.

Bignell, J.P., Caraville, M., Feist, S.W., Villalba, A., Marigomez, I., in prep. Histopathology of mussels *Mytilus* sp. for health assessment in biological effects monitoring, *ICES Techniques in Marine Envrionmental Sciences*.

Bignell, J.P., Dodge, M.J., Feist, S.W., Lyons, B., Martin, P.D., Taylor, N.G.H., Stone, D., Travalent, L., Stentiford, G.D., 2008. Mussel histopathology: and effects of season, disease and species, *Aquatic Biology* **2**, 1-15.

Bjørnstad, A., Larsen, B.K., Taban, I.C., Bechmann, R.K., Torgrimsen, S., Andersen, O.K., 2002. Assessment of DNA damage in Mussels, Crabs and Fish. Poster and platform presentation at the BEEP Workshop, Athens 6th - 9th November 2002.

Bocquené, G., Galgani, F., 1998. Biological effects of contaminants: Cholinesterase inhibition by organophosphate and carbamate compounds, ICES Techniques in Marine Environmental Sciences. ICES, Copenhagen, Denmark.

Bolognesi, C., Fenech, M., 2012. Mussel micronucleus cytome assay, *Nat Protoc* **7**, 1125-1137.

Bolognesi, C., Hayashi, M., 2011. Micronucleus assay in aquatic animals, *Mutagenesis* **26**, 205-213.

Brookes, P., Lawley, P.D., 1964. Evidence for the binding of polynuclear aromatic hydrocarbons to the nucleic acids of mouse skin: Relation between carcinogenic power of hydrocarbons and their binding to deoxyribonucleic acid, *Nature* **202**, 718-784.

Brooks, S., Pampanin, D., Harman, C., Dunaevskava, E., 2013. The Water Column Monitoring Programme 2013: Determining the biological effects of two offshore platforms on local fish populations. Norwegian Institute for Water Research, p. 61.

Brooks, S., Pampanin, D., Harman, C., Grung, M., 2015. Water Column Monitoring 2014: Determining the biological effects of an offshore platform on local fish populations. Norwegian Institute for Water Research, Norwegian Institute for Water Research, p. 79.

Brooks, S., Pampanin, D., Harman, C., Aarab, N., Sundt, R.C., 2011a. Water Column Monitoring 2011. Norwegian Institute for Water Research, p. 64.

Brooks, S.J., Farmen, E., 2013. The Distribution of the Mussel Mytilus Species Along the Norwegian Coast, *Journal of Shellfish Research* **32**, 265-270.

Brooks, S.J., Harman, C., Grung, M., Farmen, E., Ruus, A., Vingen, S., Godal, B.F., Barsiene, J., Andreikenaite, L., Skarpheoinsdottir, H., Liewenborg, B., Sundt, R.C., 2011b. Water Column Monitoring of the Biological Effects of Produced Water from the Ekofisk Offshore Oil Installation from 2006 to 2009, *Journal of Toxicology and Environmental Health-Part A* **74**, 582-604.

Bucheli, T.D., Fent, K., 1995. Induction of cytochrome P450 as a biomarker for environmental contamination in aquatic ecosystems, *Crit. Rev. in Env. Sci. Technol.* **25**, 201-268.

Børseth, J.F., Tollefsen, K.E., 2004. Water Column Monitoring 2003 - Summary Report. RF-Akvamiljø, p. 27 pp.

Collier, T.K., Varanasi, U., 1991. Hepatic activities of xenobiotic metabolizing enzymes and biliary levels of xenobiotics in English sole (*Parophrys vetulus*) exposed to environmental contaminants, *Archives of Environmental Contamination and Toxicology* **20**, 463-473.

Cossa, D., 1989. A review of the use of *Mytilus spp* as quantitative indicators of cadmium and mercury contamination in coastal waters, *Oceanologica Acta* **12**, 417-432.

D'Agostini, F., La Maestra, S., 2021. Micronuclei in Fish Erythrocytes as Genotoxic Biomarkers of Water Pollution: An Overview, in: DeVoogt, P. (Ed.), Reviews of Environmental Contamination and Toxicology, Vol 258. Springer International Publishing Ag, Cham, pp. 195-240.

Davies, I., Vethaak, D., 2012. Integrated marine environmental monitoring of chemicals and their effects, ICES Cooperative Research. ICES - International Council for the Exploration of the Sea, Copenhagen Denmark, p. 277.

Davies, I.M., Minchin, A., 2002. Quality assurance of imposex and intersex measurements in marine snails, *Journal of Environmental Monitoring* **4**, 788-790.

deZwaan, A., Eertman, R.H.M., 1996. Anoxic or aerial survival of bivalves and other euryoxic invertebrates as a useful response to environmental stress - A comprehensive review, *Comparative Biochemistry and Physiology C-Pharmacology Toxicology & Endocrinology* **113**, 299-312.

Duinker, A., Haland, L., Hovgaard, P., Mortensen, S., 2008. Gonad development and spawning in one and two year old mussels (Mytilus edulis) from Western Norway, *Journal of the Marine Biological Association of the United Kingdom* **88**, 1465-1473.

Eertman, R.H.M., Groenink, C.L.F.M., Sandee, B., Hummel, H., Smaal, A.C., 1995. Response of the blue mussel *Mytilus edulis* L. following exposure to PAHs or contaminated sediments, *Marine Environmental Research* **39**, 169-173.

Eggermont, M., Cornillie, P., Dierick, M., Adriaens, D., Nevejan, N., Bossier, P., Van den Broeck, W., Sorgeloos, P., Defoirdt, T., Declercq, A.M., 2020. The blue mussel inside: 3D visualization and description of the vascular-related anatomy of Mytilus edulis to unravel hemolymph extraction, *Sci Rep-Uk* **10**, 16.

Ehrhardt, M., Klungsøyr, J., Law, R.J., 1991. Hydrocarbons: Review of methods for analysis in sea water, biota, and sediments, *ICES Techniques in Marine Environmental Sciences*, 47 pp.

Ek-Huchim, J.P., Arcega-Cabrera, F., May-Tec, A.L., Amendola-Pimenta, M., Ceja-Moreno, V., Rodriguez-Canul, R., 2022. Red Blood Cell Cytotoxicity Associated to Heavy Metals and Hydrocarbons Exposure in Flounder Fish from Two Regions of the Gulf of Mexico, *Bull Environ Contam Toxicol* **108**, 78-84.

Ellman, G., Courtney, K., Andres, V., Featherstone, R., 1961. A new and rapid colorimetric determination of acetylcholinesterase activity, *Biochem. Pharmacol.* **7**, 88-95.

Feist, S.W., Lang, T., Stentiford, G.D., Köhler, A., 2004. Biological effects of contaminants: Use of liver pathology of the European flatfish dab (*Limanda limanda* L.) and flounder (*Platichthys flesus* L.) for monitoring, ICES Techniques in Marine Environmental Sciences, p. 42 pp.

Fenech, M., 2000. The in vitro micronucleus technique, *Mutation Research-Fundamental and Molecular Mechanisms of Mutagenesis* **455**, 81-95.

Fenech, M., Jarvis, L.R., Morley, A.A., 1988. Preliminary studies on scoring micronuclei by computerised image analysis, *Mutation Research* **203**, 33-38.

Fu, H.L., Xia, Y.J., Chen, Y.S., Xu, T., Xu, L., Guo, Z.L., Xu, H.M., Xie, H.Q., Zhao, B., 2018. Acetylcholinesterase Is a Potential Biomarker for a Broad Spectrum of Organic Environmental Pollutants, *Environmental Science & Technology* **52**, 8065-8074.

Garmendia, L., Izagirre, U., Cajaraville, M.P., Marigomez, I., 2011a. Application of a battery of biomarkers in mussel digestive gland to assess long-term effects of the Prestige oil spill in Galicia and the Bay of Biscay: Lysosomal responses, *Journal of Environmental Monitoring* **13**, 901-914.

Garmendia, L., Soto, M., Ortiz-Zarragoitia, M., Orbea, A., Cajaraville, M.P., Marigomez, I., 2011b. Application of a battery of biomarkers in mussel digestive gland to assess long-term effects of the Prestige oil spill in Galicia and Bay of Biscay: Correlation and multivariate analysis, *Journal of Environmental Monitoring* **13**, 933-942.

Garmendia, L., Soto, M., Vicario, U., Kim, Y., Cajaraville, M.P., Marigomez, I., 2011c. Application of a battery of biomarkers in mussel digestive gland to assess long-term effects of the Prestige oil spill in Galicia and Bay of Biscay: Tissue-level biomarkers and histopathology, *Journal of Environmental Monitoring* **13**, 915-932.

Gaspari, L., Chang, S.S., Santella, R.M., Garte, S., Pedotti, P., Taioli, E., 2003. Polycyclic aromatic hydrocarbon-DNA adducts in human sperm as a marker of DNA damage and infertility, *Mutation Research-Genetic Toxicology and Environmental Mutagenesis* **535**, 155-160.

Gedik, C.M., Ewen, S.W.B., Collins, A.R., 1992. Single-Cell Gel-Electrophoresis Applied to the Analysis of Uv-C Damage and Its Repair in Human-Cells, *Int. J. Radiat. Biol.* **62**, 313-320.

Giusti, L., Zhang, H., 2002. Heavy metals and arsenic in sediments, mussels and marine water from Murano (Venice, Italy), *Environmental Geochemistry and Health* **24**, 47-65.

Goksøyr, A., 1991. A Semiquantitative Cytochrome P4501A1 ELISA - a Simple Method for Studying the Monooxygenase Induction Response in Environmental Monitoring and Ecotoxicological Testing of Fish, *Science of the Total Environment* **101**, 255-262.

Gupta, R.C., Reddy, M.V., Randerath, K., 1982. ^{32}P-Postlabeling analysis of non-radioactive aromatic carcinogen-DNA adducts, *Carcinogenesis* **3**, 1081-1092.

Hamer, B., Jaksic, Z., Pavicic-Hamer, D., Peric, L., Medakovic, D., Ivankovic, D., Pavicic, J., Zilberberg, C., Schroder, H.C., Muller, W.E.G., Smodlaka, N., Batel, R., 2008. Effect of hypoosmotic stress by low salinity acclimation of Mediterranean mussels *Mytilus galloprovincialis* on biological parameters used for pollution assessment, *Aquatic Toxicology* **89**, 137-151.

Hansson, T., Thain, J., Martínez-Gómez, C., Hylland, K., Gubbins, M., Balk, L., 2017. Supporting variables for biological effects measurements in fish and blue mussel, *ICES Techniques in Marine Environmental Sciences*, 27 pp.

Harman, C., Brooks, S., Sundt, R.C., Meier, S., Grung, M., 2011. Field comparison of passive sampling and biological approaches for measuring exposure to PAH and alkylphenols from offshore produced water discharges, *Marine Pollution Bulletin* **63**, 141-148.

Harman, C., Langford, K., Sundt, R.C., Brooks, S., 2014. Measurement of naphthenic acids in the receiving waters around an offshore oil platform by passive sampling, *Environmental Toxicology and Chemistry* **33**, 1946-1949.

Hauser-Davis, R.A., Lopes, R.M., Ziolli, R.L., 2019. Inihibition of mullet (M-liza) brain acetylcholinesterase activity by in vitro polycyclic aromatic hydrocarbon exposure, *Marine Pollution Bulletin* **140**, 30-34.

Hellou, J., Law, R.J., 2003. Stress on stress response of wild mussels, Mytilus edulis and Mytilus trossulus, as an indicator of ecosystem health, *Environmental Pollution* **126**, 407-416.

Horak, S., Polanska, J., Widlak, P., 2003a. Bulky DNA adducts in human sperm: relationship with fertility, semen quality, smoking, and environmental factors, *Mutation Research-Genetic Toxicology and Environmental Mutagenesis* **537**, 53-65.

Horak, S., Polanska, J., Widlak, P., 2003b. High levels of bulky DNA adducts in human sperm correlate with impaired fertility, *Acta Biochimica Polonica* **50**, 197-203.

Hylland, K., Ruus, A., Børseth, J.F., Bechmann, R., Barsiene, J., Grung, M., Tollefsen, K., Myhre, L.P., 2006. Biomarkers in monitoring - a review. Norwegian Institute for Water Research, Oslo, Norway, p. 115.

Hylland, K., Ruus, A., Sundt, R.C., Feist, S., Marigomez, I., Balk, L., Abrahamsson, A., Barsiene, J., 2005. Water Column Monitoring 2004 – Summary report. Report by NIVA and RF for Statiol ASA on behalf of the OLF WCM coordination group, p. 59 pp. + appendix.

Hylland, K., Tollefsen, K.E., Ruus, A., Jonsson, G., Sundt, R.C., Sanni, S., Røe Utvik, T.I., Johnsen, S., Nilssen, I., Pinturier, L., Balk, L., Baršiene, J., Marigòmez, I., Feist, S.W., Børseth, J.F., 2008. Water column monitoring near oil installations in the North Sea 2001-2004, *Marine Pollution Bulletin* **56**, 414-429.

ICES, 2011. Report of the Study Group on Integrated Monitoring of Contaminants and Biological Effects (SGIMC). The International Council for the Exploration of the Sea (ICES), Copenhagen, Denmark, p. 265 pp.

ICES, 2021. Working Group on Biological Effects of Contaminants (WGBEC). ICES Scientific Reports. The International Council for the Exploration of the Sea (ICES), Copenhagen, Denmark, p. 90 pp.

Inoue, K., Waite, J.H., Matsuoka, M., Odo, S., Harayama, S., 1995. Interspecific variations in adhesive protein sequences of Mytilus edulis, M-galloprovincialis, and M-trossulus, *Biological Bulletin* **189**, 370-375.

Jonsson, G., Beyer, J., Wells, D., Ariese, F., 2003. The application of HPLC-F and GC-MS to the analysis of selected hydroxy polycyclic hydrocarbons in two certified fish bile reference materials, *Journal of Environmental Monitoring* **5**, 513-520.

Jonsson, G., Sundt, R.C., Aas, E., Beyer, J., 2004. An evaluation of two fluorescence screening methods for the determination of chrysene metabolites in fish bile, *Chemosphere* **56**, 81-90.

Kang, J.J., Fang, H.W., 1997. Polycyclic aromatic hydrocarbons inhibit the activity of acetylcholinesterase purified from electric eel, *Biochem. Biophys. Res. Commun.* **238**, 367-369.

Kilgour, B.W., Somers, K.M., Matthews, D.E., 1998. Using the normal range as a criterion for ecological significance in environmental monitoring and assessment, *Ecoscience* **5**, 542-550.

Klungsøyr, J., Wilhelmsen, S., Westrheim, K., Saetvedt, E., K.H., P., 1988. The GEEP Workshop: Organic Chemical Analyses, *Marine Ecology Progress Series* **46**, p. 19-26.

Koehler, A., 2004. The gender-specific risk to liver toxicity and cancer of flounder (*Platichthys flesus* (L.)) at the German Wadden Sea coast, *Aquatic Toxicology* **70**, 257-276.

Koehler, A., 2006. Diagnostic criteria of liver anomalies in relevant indicator fish species on the basis of BEQUALM, *Marine Environmental Research* **62**, S372-S372.

Köhler, A., Ellesat, K., 2008. Nuclear changes in blood, early liver anomalies and hepatocellular cancers in flounder (*Platichthys flesus* L.) as prognostic indicator for a higher cancer risk?, *Marine Environmental Research* **66**, 149-150.

Lang, T., Feist, S.W., Stentiford, G.D., Bignell, J.P., Vethaak, A.D., Wosniok, W., 2017. Diseases of dab (*Limanda limanda*): Analysis and assessment of data on externally visible diseases., macroscopic liver neoplasms and liver histopathology in the North Sea, Baltic Sea and off Iceland, *Marine Environmental Research* **124**, 61-69.

Large, A.T., Shaw, J.P., Peters, L.D., McIntosh, A.D., Webster, L., Mally, A., Chipman, J.K., 2002. Different levels of mussel (*Mytilus edulis*) DNA strand breaks following chronic field and acute laboratory exposure to polycyclic aromatic hydrocarbons, *Marine Environmental Research* **54**, 493-497.

Lee, R.F., Steinert, S., 2003. Use of the single cell gel electrophoresis/comet assay for detecting DNA damage in aquatic (marine and freshwater) animals, *Mutation Research-Reviews in Mutation Research* **544**, 43-64.

Lowe, D.M., Moore, M.N., Evans, B.M., 1992. Contaminant impact on interactions of molecular probes with lysosomes in living hepatocytes from dab *Limanda limanda*, *Mar Ecol Progr Ser* **91**, 135-140.

Lyons, B.P., Bignell, J.P., Stentiford, G.D., Bolam, T.P.C., Rumney, H.S., Bersuder, P., Barber, J.L., Askem, C.E., Nicolaus, M.E.E., Maes, T., 2017. Determining Good Environmental Status under the Marine Strategy Framework Directive: Case study for descriptor 8 (chemical contaminants), *Marine Environmental Research* **124**, 118-129.

Mamaca, E., Bechmann, R.K., Torgrimsen, S., Aas, E., Bjornstad, A., Baussant, T., Le Floch, S., 2005. The neutral red lysosomal retention assay and Comet assay on haemolymph cells from mussels (*Mytilus edulis*) and fish (*Symphodus melops*) exposed to styrene, *Aquatic Toxicology* **75**, 191-201.

Martínez-Gómez, C., Bignell, J., Lowe, D., 2015. Lysosomal membrane stability in mussels, ICES Techniques in Marine Environmental Sciences - ICES TIMES International Council for the Exploration of the Sea, Copenhagen, Denmark, p. 46 pp.

Martinez-Gomez, C., Robinson, C.D., Burgeot, T., Gubbins, M., Halldorsson, H.P., Albentosa, M., Bignell, J.P., Hylland, K., Vethaak, A.D., 2017. Biomarkers of general stress in mussels as common indicators for marine biomonitoring programmes in Europe: The ICON experience, *Marine Environmental Research* **124**, 70-80.

Meier, S., Karlsen, Ø., Le Goff, J., Sørensen, S., Sørhus, E., Pampanin, D.M., Donald, C.E., Fjelldal, P.G., Dunaevskaya, E., Romano, M., Caliani, I., Casini, S., Bogevik, A.S., Olsvik, P.A., Myers, M., Grøsvik, B.E., 2020. DNA damage and health effects in juvenile haddock (*Melanogrammus aeglefinus*) exposed to PAHs associated with oil-polluted sediment or produced water, *PLOS ONE* **15**, 34 pp.

Miljødirektoratet, Norwegian Environment Agency, 2011. Guidelines for offshore environmental monitoring on the Norwegian continental shelf, in: Iversen, P.E., Vik Green, A.M., Juel Lind, M., Rønn Hedegaard Petersen, M., Bakke, T., Lichtenthaler, R., Klungsøyr, J., Grafert, T., Natvig, H., Ersvik, M.. Norwegian Climate and Pollution Agency, Oslo, Norway, p. 50.

Miljødirektoratet, Norwegian Environment Agency, 2015. Guidelines for environmental monitoring of petroleum activities on the Norwegian continental shelf, in: Iversen, P.E., Juel Lind, M., Ersvik, M., Rønning, I., Skaare, B.B., Green, A.M.V., Bakke, T., Lichtenthaler, R., Klungsøyr, J., Hylland, K. (Eds.). Norwegian Environment Agency, Oslo, Norway, p. 65.

Miljødirektoratet, Norwegian Environment Agency, 2020. Guidelines for environmental monitoring of petroleum activities on the Norwegian continental shelf. Miljødirektoratet, p. 70 pp.

Moore, M.N., Lowe, D., Köhler, A., 2004. Biological effects of contaminants: Measurement of lysosomal membrane stability, ICES Techniques in Marine Environmental Sciences, p. 31 pp.

Moschino V. and Marin, G.M. (2006) Seasonal changes in physiological responses and evaluation of "well-being" in the Venus clam *Chamelea gallina* from the Northern Adriatic Sea., *Comparative Biochemistry and Physiology a-Molecular & Integrative Physiology*, **Vol. 145,** Issue 4, Pages 433-440.

Nash, R.D.M., Valencia, A.H., Geffen, A.J., 2006. The origin of Fulton's condition factor - Setting the record straight, *Fisheries* **31**, 236-238.

NIVA, 2017. High-throughput screening of mussel speciation by high-resolution melt analysis (confidential SOP). Norwegian Institute for Water Research Oslo, Norway, p. 7 pp.

NIVA, 2020. Micronucleus (MN) assay (confidential SOP). Norwegian Institute for Water Research (NIVA), p. 13 pp.

NIVA, 2022. Per and Polyfluorinated substances (PFAS) in fish blood (confidential SOP). Norwegian Institute for Water Research (NIVA), p. 2 pp.

Olivares-Rubio, H.F., Espinosa-Aguirre, J.J., 2021. Acetylcholinesterase activity in fish species exposed to crude oil hydrocarbons: A review and new perspectives, *Chemosphere* **264**.

Orban, E., Di Lena, G., Nevigato, T., Casini, I., Marzetti, A. and Caproni, R., 2002. Seasonal changes in meat content, condition index and chemical composition of mussels (*Mytilus galloprovincialis*) cultured in two different Italian sites, *Food Chemistry*, **Vol. 77,** Issue 1, Pages 57-65.

Oropesa, A.L., Perez-Lopez, M., Hernandez-Moreno, D., Garcia, J.P., Fidalgo, L.E., Lopez-Beceiro, A., Soler, F., 2007. Acetylcholinesterase activity in seabirds affected by the Prestige oil spill on the Galician coast (NW Spain), *Science of the Total Environment* **372**, 532-538.

OSPAR, 2012. JAMP Guidelines for the Integrated Monitoring and Assessment of Contaminants and their effects, in: Commission, O. (Ed.), Agreement 2012-09, p. 22.

OSPAR, 2013. Background document and technical annexes for biological effects monitoring, Update 2013, Monitoring and Assessment Series. OSPAR commission, London, UK, p. 239.

OSPAR, ICES, 2006. Memorandum of understanding between the OSPAR commission and the international council for the exploration of the sea, p. 6 pp.

OSPARCOM, 2004. OSPAR Guidelines for Monitoring the Environmental Impact of Offshore Oil and Gas Activities. OSPAR COMMISSION, p. 19.

Pampanin, D.M., Brooks, S., Børseth, J.F., Harman, C., Gomiero, A., Farmen, E., Aarab, N., Nerland, I.L., Westerlund, S., Grung, M., Lucas, C., Strålberg, E., 2013. Water Column Monitoring 2012 - Troll C platform. International Research Institute of Stavanger (IRIS) & Norwegian Institute for Water Research (NIVA), p. 91 pp.

Pampanin, D.M., Brooks, S., Grøsvik, B.E., Sanni, S., 2019. Water Column Monitoring 2017 - Environmental monitoring of petroleum activities on the Norwegian continental shelf 2017. NORCE, Stavanger, Norway, p. pp 92.

Pampanin, D.M., Brooks, S.J., Grøsvik, B.E., Le Goff, J., Meier, S., Sydnes, M.O., 2017. DNA adducts in marine fish as biological marker of genotoxicity in environmental monitoring: The way forward, *Marine Environmental Research* **125**, 49-62.

Richardson, D.M., Gubbins, M.J., Davies, I.M., Moffat, C.F., Pollard, P.M., 2004. Effects of feeding status on biliary PAH metabolite and biliverdin concentrations in plaice (Pleuronectes platessa), *Environmental Toxicology and Pharmacology* **17**, 79-85.

Riisgård, H.U., Larsen, P.S., Pleissner, D., 2014. Allometric equations for maximum filtration rate in blue mussels Mytilus edulis and importance of condition index, *Helgoland Marine Research* **68**, 193-198.

Roman, D., Locher, F., Suter, W., Cordier, A., Bobadilla, M., 1998. Evaluation of a new procedure for the flow cytometric analysis of in vitro, chemically induced micronuclei in V79 cells, *Environmental and Molecular Mutagenesis* **32**, 387-396.

Shaw, J.P., Large, A.T., Livingstone, D.R., Doyotte, A., Renger, J., Chipman, J.K., Peters, L.D., 2002. Elevation of cytochrome P450-immunopositive protein and DNA damage in mussels (*Mytilus edulis*) transplanted to a contaminated site, *Marine Environmental Research* **54**, 505-509.

Singh, N.P., McCoy, M.T., Tice, R.R., Schneider, E.L., 1988. A simple technique for quantification of low levels of DNA damage in individual cells, *Exp. Cell Res.* **175**, 184-191.

Smedes, F., Booij, K., 2012. Guidelines for passive sampling of hydrophobic contaminants in water using silicone rubber samplers, *ICES Techniques in Marine Environmental Sciences* **52**, 20 pp.

Straumer, K., Kraugerud, M., Feist, S.W., Ahvo, A., Lehtonen, K., Lastumaki, A., Ljones, M., Tornes, J.A., Lang, T., 2020. The use of Atlantic hagfish (*Myxine glutinosa*) as a bioindicator species for studies on effects of dumped chemical warfare agents in the Skagerrak. 1: Liver histopathology, *Marine Environmental Research* **161**, 10.

Strizak, Z., Ivankovic, D., Profrock, D., Helmholz, H., Cindric, A.M., Erk, M., Prange, A., 2014. Characterization of the cytosolic distribution of priority pollutant metals and metalloids in the digestive gland cytosol of marine mussels: Seasonal and spatial variability, *Science of the Total Environment* **470**, 159-170.

Sundt, R., Brooks, S., Grøsvik, B., Pampanin, D., Farmen, E., Harman, C., Meier, S., 2012. Water column monitoring of offshore produced water discharges. Compilation of previous experience and suggestions for future survey design. International Research Institute of Stavanger (IRIS), Stavanger, Norway, p. 123 pp.

Sundt, R.C., Ruus, A., Grung, M., Pampanin, D.M., Baršienė, J., Skarphéðinsdóttir, H., 2006. Water Column Monitoring 2006 - Summary report. IRIS, Akvamiljø, NIVA.

Sunila, I., Williams, L., Russo, S., Getchis, T., 2004. Reproduction and pathology of blue mussels, *Mytilus edulis* (L) in an experimental longline in Long Island Sound, Connecticut, *Journal of Shellfish Research* **23**, 731-740.

Taban, I.C., Bechmann, R.K., Torgrimsen, S., Baussant, T., Sanni, S., 2004. Detection of DNA damage in mussels and sea urchins exposed to crude oil using comet assay, *Marine Environmental Research* **58**, 701-705.

Taban, I.C., Bechmann, R.K., Torgrimsen, S., Sundt, R.C., Bjørnstad, A., Larsen, B.K., Andersen, O.K., 2003. DNA damage in mussels following exposure to three potential endocrine disruptors using the comet assay and the alkaline unwinding assay. Poster and platform presentation at BEEP Workshop, Barcelona,4.-6. December 2003.

Thain, J., Fernandez, B., Martinez-Gomez, C., 2019. Biological effects of contaminants: stress on stress (SOS) response in mussels, *ICES Techniques in Marine Environmental Sciences*.

Tice, R.R., Agurell, E., Anderson, D., Burlinson, B., Hartmann, A., Kobayashi, H., Miyamae, Y., Rojas, E., Ryu, J.C., Sasaki, Y.F., 2000. Single cell gel/comet assay: Guidelines for in vitro and in vivo genetic toxicology testing, *Environmental and Molecular Mutagenesis* **35**, 206-221.

Vaidya, O.C., Rantala, R.T.T., 1996. A comparative study of analytical methods: Determination of heavy metals in mussels (Mytilus edulis) from eastern Canada, *International Journal of Environmental Analytical Chemistry* **63**, 179-185.

Vethaak, A.D., Davies, I.M., Thain, J.E., Gubbins, M.J., Martinez-Gomez, C., Robinson, C.D., Moffat, C.F., Burgeot, T., Maes, T., Wosniok, W., Giltrap, M., Lang, T., Hylland, K., 2017. Integrated indicator framework and methodology for monitoring and assessment of hazardous substances and their effects in the marine environment, *Marine Environmental Research* **124**, 11-20.

Zhang, X.J., Yang, L., Zhao, Q., Caen, J.P., He, H.Y., Jin, Q.H., Guo, L.H., Alemany, M., Zhang, L.Y., Shi, Y.F., 2002. Induction of acetylcholinesterase expression during apoptosis in various cell types, *Cell Death and Differentiation* **9**, 790-800.

Östling, O., Johanson, K.J., 1984. Microelectrophoretic study of radiation-induced DNA damages in individual mammalian cells, *Biochem. Biophys. Res. Commun.* **123**, 291-298.

Aas, E., Bechmann, R.K., Larsen, B.K., Barsiene, J., Lazutka, J., Sanni, S., 2002a. Biomarkers for genotoxic effects in marine organisms. Project report 2001-2002. Akvamiljø, Stavanger, Norway.

Aas, E., Beyer, J., Goksøyr, A., 2000. Fixed wavelength fluorescence (FF) of bile as a monitoring tool for polyaromatic hydrocarbon exposure in fish: an evaluation of compound specificity, inner filter effect and signal interpretation, *Biomarkers* **5**, 9-23.

Aas, E., Jonsson, G., Sundt, R., Westerlund, S., Sanni, S., 2002b. PAH metabolites and metals in bile from cod caged in the North Sea serve as indicators of pollution. ICES, p. 6 pp.

NIVA: Norges ledende kompetansesenter på vannmiljø

Norsk institutt for vannforskning (NIVA) er Norges viktigste miljøforskningsinstitutt for vannfaglige spørsmål, og vi arbeider innenfor et bredt spekter av miljø, klima- og ressursspørsmål. Vår forskerkompetanse kjennetegnes av en solid faglig bredde, og spisskompetanse innen mange viktige områder. Vi kombinerer forskning, overvåkning, utredning, problemløsning og rådgivning, og arbeider på tvers av fagområder.

Norsk institutt for vannforskning

Økernveien 94 • 0579 Oslo
Telefon: 02348 • Faks: 22 18 52 00
www.niva.no • post@niva.no