ICES/OSPAR SGIMC REPORT 2010

ICES ADVISORY COMMITTEE

ICES CM 2010/ACOM:30

Report of the Joint ICES/OSPAR Study Group on Integrated Monitoring of Contaminants and Biological Effects (SGIMC)

25–29 January

Copenhagen, Denmark

ICES International Council for the Exploration of the Sea
CIEM Conseil International pour l'Exploration de la Mer

International Council for the Exploration of the Sea
Conseil International pour l'Exploration de la Mer

H. C. Andersens Boulevard 44–46
DK-1553 Copenhagen V
Denmark
Telephone (+45) 33 38 67 00
Telefax (+45) 33 93 42 15
www.ices.dk
info@ices.dk

Recommended format for purposes of citation:

ICES. 2010. Report of the Joint ICES/OSPAR Study Group on Integrated Monitoring of Contaminants and Biological Effects (SGIMC), 25–29 January, Copenhagen, Denmark. ICES CM 2010/ACOM:30. 211 pp. https://doi.org/10.17895/ices.pub.19280624

Contents

Executive summary

SGIMC met for five days in January 2010 at ICES HQ. The main tasks undertaken were:

- Continuation of the deveopment of backround documents and assessment criteria for biological effects measurements;
- Review and further develop OSPAR frameworks for integrated monitoring of contaminants and their effects;
- Review the need for technical Annexes to OSPAR Guidelines and to draft as necessary.

The main outcomes of the meeting were:

- A series of 13 documents completed updated documents, many with proposals for assessment criteria, recommended for adoption by OSPAR;
- Further develop proposals for a joint training and data interpretation workshop with MEDPOL on lysosomal stability to be held in 2010 in Italy;
- Updated work programme for 2010–2011, with a view to a final meeting of the SG in 2011.

1 Opening of the meeting

The meeting was opened at ICES Headquarters at 09.15 h on Monday 25 January 2010. The meeting was co-chaired by Ian Davies (OSPAR) and Dick Vethaak (ICES). The list of participants is given in Annex 1.

2 Adoption of the Agenda

The Agenda was adopted without amendment. The Terms of Reference for the meeting (see Annex 2) require SGIMC to report by 1 March to ACOM and OSPAR.

3 Preparation of the Report

The Report of SGIMC was drafted during the meeting, and circulated by the chairmen for comment and adopted at the meeting accordingly. The complete Report would be sent to ACOM and OSPAR before 1 March 2010.

4 Review progress made by Aberdeen WKIMC Workshop

Agenda item:

Review progress made by Aberdeen WKIMC Report and to receive Background Documents and draft assessment criteria from SGIMC Aberdeen Workshop planned for October 2009 on:

1) EROD;

2) Bile metabolites;

3) DNA adducts;

4) DR-CALUX;

5) Contaminant-linked fish disease;

6) (Liver neoplasm) and to assess their usefulness in integrated assessments.

SGIMC 2010 received a series of draft updated Background Documents from the Aberdeen WKIMC Workshop. SGIMC reviewed the documents, together with comments made by WGBEC 2010, made amendments where necessary, and recommends that the documents in Annexes as listed below be transferred to OSPAR for incorporation into the OSPAR JAMP Background Document on Biological Effects Monitoring Techniques.

Table 4.1 Background documents reviewed and developed at the Aberdeen WKIMC Workshop, October 2009.

DOCUMENT	LOCATION	STATUS
Background document on Cytochrome P4501A activity (EROD) – Chapter 3	Annex 4	Recommended for adoption by OSPAR
Background document on Lysosomal stability as a global health status indicator in biomonitoring (Chapter 4)	Annex 5	Recommended for adoption by OSPAR
Background document on DNA adducts (Chapter 11)	Annex 6	Recommended for adoption by OSPAR
Background document on PAH metabolites in bile (Chapter 5)	Annex 7	Recommended for adoption by OSPAR
Background document on Externally visible fish diseases, macroscopic liver neoplasms and liver histopathology (Chapter 6)	Annex 8	Recommended for adoption by OSPAR
Background document on Water Bioassays	Annex 9	Recommended for adoption by OSPAR

Cyto P4501A (EROD activity) background document

Martínez-Gómez, C. (IEO, Spain), Maes, T. (Cefas, UK) and Devriese, L. (ILVO, Belgium) reviewed and updated EROD Assessment Criteria (AC) established at the SKIMC Workshop ICES/OSPAR meeting in 2009. New Assessment Criteria values were established for Atlantic cod (*Gadus morhua*), flounder (*Platichthys flesus*), dab (*Limanda limanda*) and plaice (*Pleuronectes platessa*) by using EROD data submitted in ICES database to date (made available during SGIMC 2010 by the ICES DataCentre (Marilyn Sorenson)) and data provided during ICES/OSPAR WKIMON IV (Table EROD1).

Table EROD1. EROD Assessment criteria in fish target species used in biomonitoring programmes around European waters. EROD BRs established are restricted to the sampling conditions and the size length of the specimens used. The values of the assessment criteria must be considered as provisional and should be updated and revised when more data comes available.

EROD Assessment Criteria S9 Fraction	Sampling season	Bottom water temperature range	Size length cm	Sex	Background Response Range EROD activity (pmol/min/mg prot) 90P	Elevated Response Range EROD activity (pmol/min/mg prot) 90P	N
Dab (*Limanda Limanda*)	August–November	[10–18 °C]	12–25	Females	≤178	>178	556
				Males	≤147	>147	571
European flounder (*Platichtys flesus*)	August–November	[10–18 °C]	20–25	Females and/or males	≤24	>24	65
Plaice (*Pleuronectes platessa*)	January	[5–10°C]	18.5–22.5	Males	≤10	>10	116
EROD Assessment Criteria Microsomal fraction	Sampling season	Bottom water temperature range	Size length cm	Sex	Background Response Range EROD activity (pmol/min/mg prot) 90P	Elevated Response Range EROD activity (pmol/min/mg prot) 90P	N
Dab (*Limanda Limanda*)	August–November	[10–18 °C]	20–30	Females and/or males	≤780	>780	53
Cod (*Gadus morhua*)	August–November	[10–18 °C]	30–45	Females and/or males	≤145	>145	198
Plaice (*Pleuronectes platessa*)	September	[7–10 °C]	40–60	Females and/or males	≤255	>255	64
Four spotted megrim (*Lepidorhombus boscii*)	September–October	[11.7–12.7 °C]	18–22	Females and/or males	≤13	>13	317

EROD Assessment Criteria S9 Fraction	Sampling season	Bottom water temperature range	Size length cm	Sex	Background Response Range EROD activity (pmol/min/mg prot) 90P	Elevated Response Range EROD activity (pmol/min/mg prot) 90P	N
Dragonet (*Callionymus lyra*)	September–October	[12.0–12.8 °C]	15–22	Females and/or males	≤202	>202	159
Red mullet (*Mullus barbatus*)	April	[13.3–15.3 °C]	12–18	Males	≤208	>208	40

Establishment of EROD ACs

Data were split into subgroups based on species and the used sub-fractions in assessing hepatic EROD activities (microsomes and S9 fractions). EROD data were not considered when total number of data available was less than 30. Basically, background levels were calculated for those months and stations which contracting parties consider being reference stations (i.e. no known local sources of contamination or those areas which were not considered unequivocally as reference sites but as those less influenced from human and industrial activity.) Table EROD2.

Table EROD2. Reference stations used for the establishment of EROD Assessment Criteria.

SPECIES	REFERENCE STATION	COUNTRY
Cod	15F Ullerø area, 23B Karihavet area, 98B2 Austnesfjorden, 10B Varangerfjorden, 36B2 Austnesfjorden	Norway
Dab	Le Parfond	France
	15F Ullerø area, 36F Faerder area	Norway
	Belgian CS, Falls, Mouth of Thames, Plaice box, South and West Falls	Belgium
Flounder	St. Andrews bay	UK
	21F Akrafjord	Norway
Plaice	10F Skogeroy, 98F2 Husholmen	Norway
Red Mullet	Guardamar, Valencia	Spain
Dragonet	Asturias W, Asturias E,	Spain
Four-spotted megrim	Galicia S, Finisterre, Galicia N, Asturias W, Asturias E	Spain

Differences on EROD responses were tested for each of the species considered (t-test of the mean or Mann–Whitney test). Only for dab and red mullet, sex differences were demonstrated ($p<0.05$). For red mullet only males' data were considered for ACs due to the sampling time (pre-pawning period). For dab, ACs were calculated for each one of the genders. For the rest of species tested (cod, flounder, plaice, dragonet, four-spotted megrim) female and male data were combined to one group for further statistical analysis.

Background levels were calculated as the empirical 90% quantile. The 90% quantile (P90) separates the upper 10% of all values in the group from the lower 90%. The rationale for this decision was that elevated EROD levels would lie above the P90 quantile, whereas the majority of values below P90 belong to unexposed or non-responding individuals. EROD ACs established are restricted to the sampling conditions and the size length of the specimens used. The values of the assessment criteria must be considered as provisional and should be updated and revised when more data becomes available for these species.

The use of a conversion factor to transform data obtained in different fractions and therefore made them comparable was discussed by the group. Finally, it was concluded that considering data available in the literature and/or in databases managed during the meeting, the establishment of a conversion factor is not realistic, the main reason being that data in different fractions were not obtained for same specimens.

Recommendation: Countries conducting biomonitoring programmes should analyse EROD activity in both fractions (S9 and microsomal) in the same individuals in a small number of samples to contribute to the establishment of this conversion factor.

SGIMC further noted that dragonet and red mullet are ideal species for intercalibraton due to their widely spread distribution in Atlantic and Mediterranean waters.

Background responses (BR) had been developed by in ICES/OSPAR WKIMON III meeting in 2007. Basically, median values were calculated for those months and stations which contracting parties consider being reference stations (in terms of no known local sources of contamination and as those less influenced from human and industrial activity). For those medians, an overall median was calculated as the background EROD induction level. Using the same criteria, BRs were calculated for the new species. In addition to the BRs, the empirical 90% quantile was also calculated for all species. The 90% quantile (P90) separates the upper 10% of all values in the group from the lower 90%. The rationale for this decision was that elevated EROD levels would lie above the P90 quantile, whereas the majority of values below P90 belong to unexposed or non-responding individuals.

BRs for saithe, herring and haddock were estimated by using the same dataset that was available at the joint ICES/OSPAR WKIMON III meeting in 2007.

BRs for haddock in S9 fraction has been estimated using IRIS database (BioSea Project) made available by Total E&P Norge and Eni Norge AS (Norway), using data from Barents Sea as a reference station.

BRs for plaice, dragonet, four-spotted megrim and red mullet has been estimated using data recently submitted to ICES database and available from the Spanish Institute of Oceanography (IEO, Spain) and from Marine Scotland (Scotland, UK) during the workshop. Part of these results derived from the mid-term monitoring performed by the Spanish Institute of Oceanography (IEO) after the Prestige oil spill (Martínez-Gómez *et al*., 2009). Only dataset obtained along the northern Iberian shelf in autumn 2004 and autumn 2005 from selected areas were used from each species. For *L. boscii*, data from Basque country and Cantabria were excluded. For *C. lyra*, only data from Asturias W and Asturias E were considered. For red mullet, background values were derived from the results obtained in Valencia and Santa Pola areas (SE Spain), within the framework of the biomonitoring programme undertaken by the IEO in 2006 (MEDPOL project), under the responsibility of the Spanish Ministry of Environment, to contribute to MED POL Programme. For plaice, background values were derived from the results obtained in the station Broad Bay (Isle of Lewis, Scotland, UK), within the national CSEMP monitoring programme undertaken by the Marine Scotland from 2001 to 2009. Because significant sexual differences were observed in these two datasets, only males were considered.

EROD values can be expressed in different units, in relation to the sub-fraction used to assess the activity. The most commonly used sub-fractions in assessing hepatic metabolism are microsomes and S9 fractions. Based on existing studies in different organisms, a conversion factor of 2.5–3 was proposed to convert data between S9 and microsomal protein content, although this has to be confirmed with more data and further research.

Table EROD3. EROD Background responses in fish target species used in biomonitoring programmes around European waters.

EROD Assessment Criteria † sub-fraction S9 * microsomes sub-fraction	Sampling season	Bottom water temperature range	Size length	Sex	EROD BR activity	Upper limit of EROD BR activity	N
					Median (pmol/min/mg prot)	P90 (pmol/min/mg prot)	
		°C	cm				
Dab (*Limanda Limanda*)	August–November	[10–18]	12–25	Females and/or males	< 30†	< 152†	1034
European flounder (*Platichthys flesus*)	August–November	[10–18]	20–25	Females and/or males	< 14†	< 24†	30
Cod (*Gadus morhua*)	August–November	[10–18]	30–45	Females and/or males	< 78*	< 151*	74
Four spotted megrim (*Lepidorhombus boscii*)	September–October	[11.7–12.7]	18–22	Females and/or males	< 12*	< 13*	317
Dragonet (*Callionymus lyra*)	September–October	[12.0–12.8]	15–22	Females and/or males	< 144*	<202*	159
Red mullet (*Mullus barbatus*)	April	[13.3–15.3]	12–18	Males	< 85*	<208*	40
Plaice (*Pleuronectes platessa*)	January	[18.5–22.5]	7–10	Males	<3.71†	<9.49†	116
Haddock (*Melanogrammus aeglefinus*)	August	[5–10]	33–55	Females and/or males	< 72† / <215*	<162† /<421*	20/23
Saithe (*Pollachius virens*)	September	[5–10]	40–100	Females and/or males	<57†	<142†	21
Herring (*Clupea harengus*)	November	[5–10]	22–33	Females and/or males	<10†	<23†	24

Assessment criteria for PAH metabolites in fish calculated from reference sites

PAH metabolites are subcellular biomarkers which are contaminant specific markers of PAH exposure. However, PAH metabolites do not indicate directly a population relevant effect. Assessment Criteria for PAH metabolites such as BAC (Background Assessment Criteria) which are usually derived from reference sites describe the threshold which indicates a significant difference to background values (green -> yellow). EAC (Environmental Assessment Criteria) are usually derived from toxicological data and indicate a significant risk for the organism. In the "traffic light system" data exceeding EAC are displayed in red.

4.1.1 Monitoring data available from vTI Germany

Ulrike Kammann and Werner Wosniok (Germany) presented a set of monitoring data produced by vTI (Germany) comprising 1-hydroxypyrene and 1-hydroxyphenenthrene data analysed in bile samples of dab, flounder, cod and haddock caught mainly in the North Sea and the Baltic. In total more than 2000 individual fish have been analysed between 1998 and 2007. The analytical method used was HPLC-fluorescence. All fish were caught in August or September, so no seasonal variation is reflected in the data. No time-trends were visible in any of the regions during the time period covered (Kammann, 2007). Comparing the areas North Sea, German Bight and Baltic regional differences in PAH-metabolites obviously are larger than species differences (Figure 1). The regional distribution of the data of three species is shown in Figure 2.

We propose that levels of PAH metabolites from these three fish species may be combined when results are to be compared on a big regional scale.

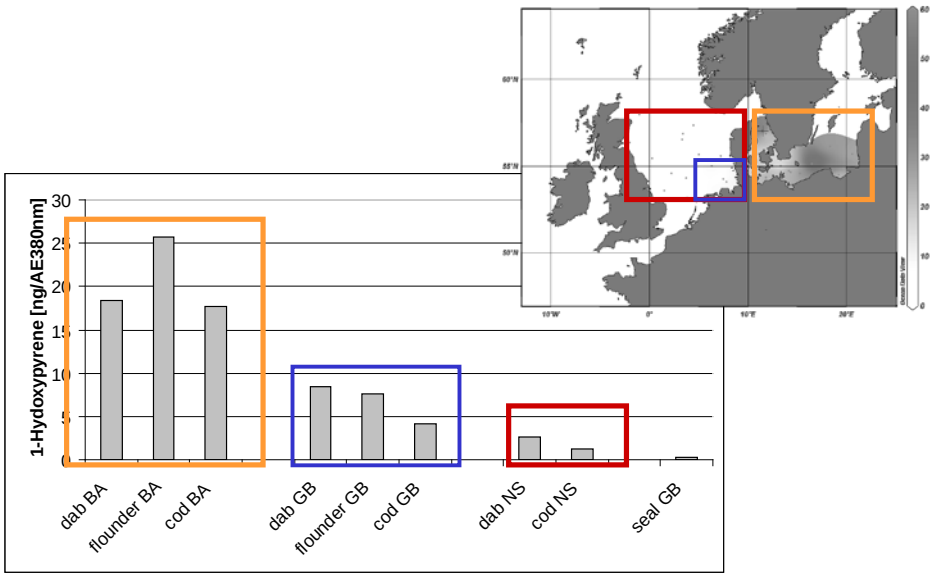


Figure 1. 1-Hydroxypyrene in bile fluids of dab, flounder and cod caught between 1998 and 2007 and grouped by the regions Baltic (BA), German Bight (GB) and North Sea (NS).

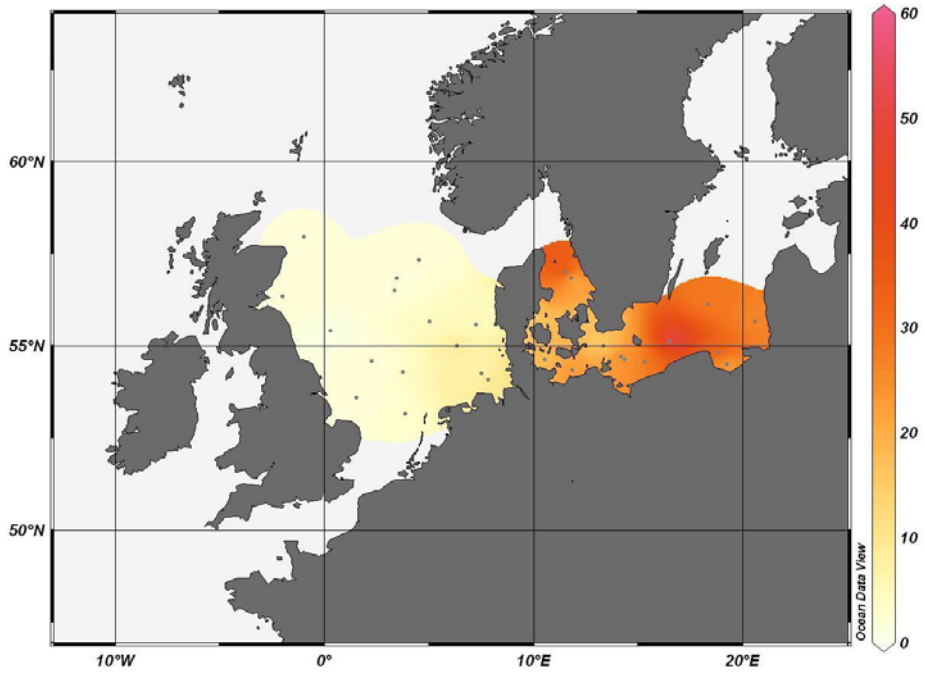


Figure 2. 1-Hydroxypyrene in bile fluids of dab, flounder and cod caught between 1998 and 2007 covering North Sea and Baltic Sea.

4.1.2 Background assessment criteria calculated from reference sites

The recommended way to calculate BACs is to use the 90th percentile of reference site data. We used two areas as reference sites: Iceland and Barents Sea and obtained BACs close to values presented by SGIMC 2009 (except flounder). Data from additional reference sites may improve the quality of the BAC in future. We applied the BAC for dab, cod and flounder using either species-specific BAC (dab value for flounder) or an overarching BAC for all species under investigation. The resulting regional assessment had a plausible appearance and showed a clear differentiation between the areas North Sea, German Bight and Baltic, as mentioned above. Two ways of presenting the results were compared: (1) a single colour for the region depending on whether the station mean exceeds the BAC or not (Figure 3) and (2) two colours representing the proportions of the fish above and below the BAC (Figure 4). The version presented in Figure 4 provides more information which is needed during the process of data aggregation in integrated monitoring, while data displayed as in Figure 3 displays the results in a simple "traffic light system" as it may be desired at the end of the integration process. Both ways of presenting are based on the same data and BACs.

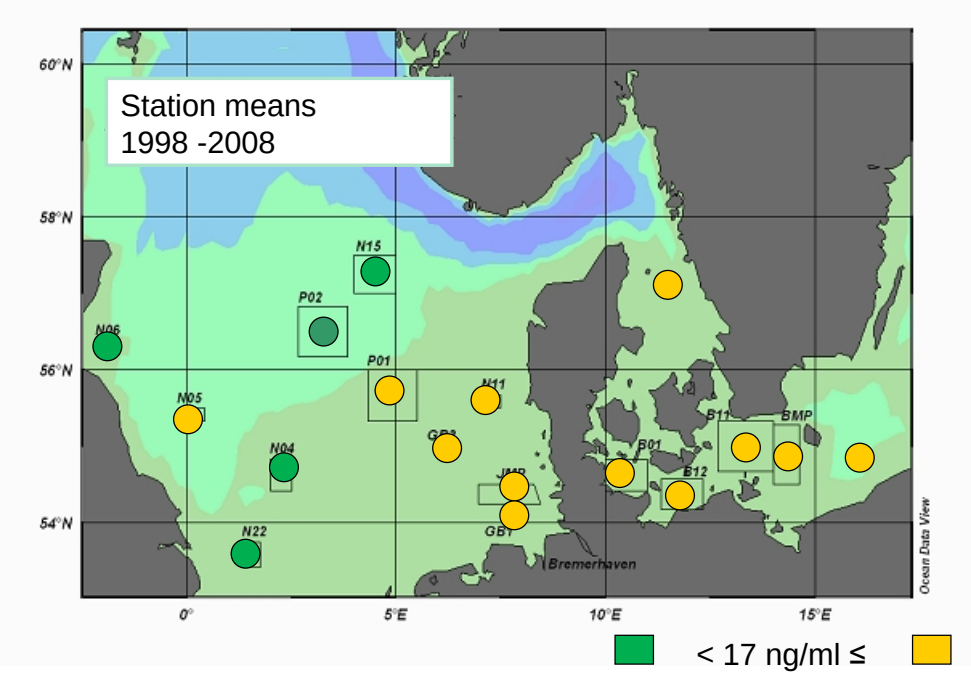


Figure 3. 1-Hydroxypyrene in bile fluids of dab, flounder and cod caught between 1998 and 2007 categorized by the species overarching BAC of 17 ng/ml. Mean values per station are compared with BAC.

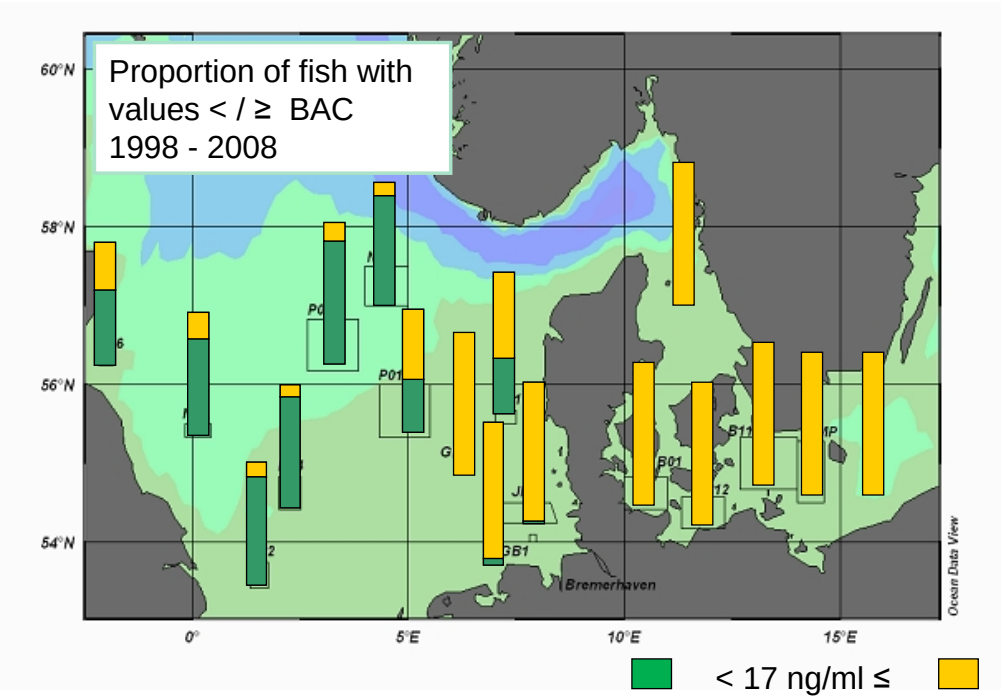


Figure 4. 1-Hydroxypyrene in bile fluids of dab, flounder and cod caught between 1998 and 2007 categorized by the species overarching BAC of 17 ng/ml. Proportion of single fish per station are categorized in relation to BAC.

The species overarching BAC are justified by the observation that different species from the same areas exhibit similar values of PAH-metabolites as shown above. The second argument is the fact that the BAC calculated for single species (13, 15 and 21 ng/ml; compare Table 1) are not distinguishable in the light of analytical variance of

the method which has a CV of 15% for HPLC-Fluorescence (calculated from a reference material analysed in vTI over a 3-year period). The overarching BAC might be of advantage when no species-specific BAC is available because of the lack of reference areas.

Comparing different ways of BAC calculation

BAC can be calculated from reference site data using the 90% percentile of all data considered as reference. If no reference site is available, BAC can be calculated using the 10% percentile of a dataset which is believed to contain low values close to reference values as well as non-reference data. Comparing both ways of BAC calculation for the present cod data shows that the 10% approach underestimates the BAC value (Figure 5). In general also an overestimation might occur, if a dataset contains very few or no near-reference values.

We conclude that a BAC can only be calculated reliably using data from reference areas. However, by identifying more or new reference areas the quality of a BAC can be improved continuously.

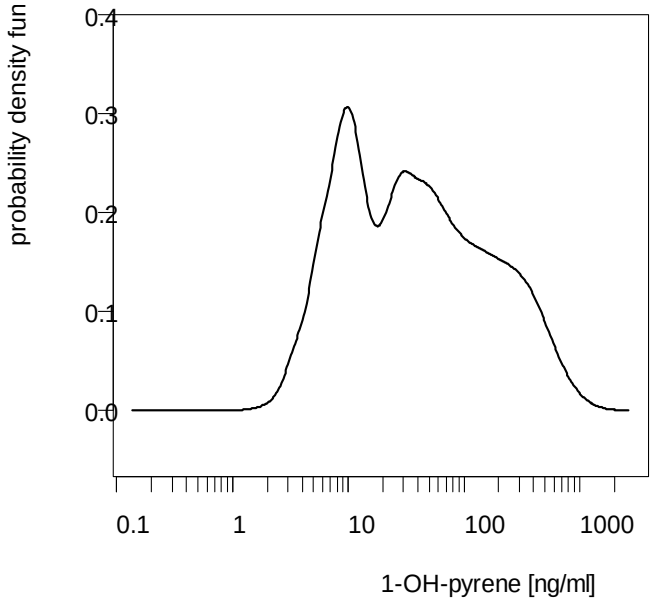


Figure 5. Distribution of 1-hydroxypyrene data in bile fluids of cod from reference (green) and non-reference regions (blue). BACs based on 10% of all and 90% of reference data are indicated.

Environmental assessment Criteria calculated from toxicological data

Even if PAH metabolites are a marker of exposure, high levels of metabolites can be linked to deleterious effects in fish so that EACs can be identified. During SGIMC 2010, Steinar Sanni (Norway) made data available from toxicological experiments linking oil exposure and PAH metabolites in fish with DNA-adducts and fitness data (Morton *et al*., unpublished data; Skadsheim *et al*. 2004; Skadsheim *et al*., 2009), where the latter serves as the effect quantity for the calculation of the EAC presented in Table 1. Table 1 is not yet complete and needs additional input. Also the two analytical methods HPLC-F and GC/MS need to be compared.

Table 1. Biological assessment criteria (BAC) and Environmental assessment criteria (EAC) for two PAH metabolites, different fish species and methods. Data partly taken from WKIMC 2009

Biological Effect	Fish species	BAC [ng/ml] HPLC-F	EAC [ng/g] GC/MS
Bile metabolit 1-hydroxypyrene	dab	16	
	cod	21	483
	flounder	16 4)	
	haddock	13	
	dab, cod, haddock	17	
	turbot		909
	halibut		745
Bile metabolit 1-hydroxyphenanthrene	dab	3,7	
	cod	2.7	518
	flounder	3.7 4)	
	haddock	0.8	
	dab, cod, haddock	2.4	
	turbot		1832
	halibut		262
Biological Effect	Fish species	BAC [µg/ml] syncronuos Fluor. 341/383	EAC [µg/ml] Fixed Fluor. 341/383
Bile metabolites of pyrene-type	dab	0.15	22 1)
	cod	1.1	35
	flounder	1.3	29 2)
	haddock	1.9	35 3)
	turbot		29
	halibut		22
	herring/sprat		16
AC based on 1)halibut, 2)turbot, 3)cod and 4)dab			

References

Boitsov, K.A. Kvestad, A.B. Kjersem, A. Goksøyr, A. Folkvord, J. Klungsøyr, A. Svardal, S. Meier. *in subm*. Development of Atlantic cod (*Gadus morhua*) exposed to produced water during early life stage. I. Effects on embryos, larvae, and juvenile fish.

Kammann U. 2007. PAH metabolites in bile fluids of dab (*Limanda limanda*) and flounder (*Platichthys flesus*)-spatial distribution and seasonal changes. Environ Sci Pollut Res 14(2) 102–108.

Skadsheim, A. 2004. Cod stocks exposed to crude oils: uptake, metabolism and biomarker effects. Technical report AM-2004/005, Akvamiljo, Randaberg, Norway.

Skadsheim, A., Sanni, S., Pinturier, L., Moltu, U.-E., Buffagni, M. and Bracco, L. 2009. Assessing and monitoring long-range-transported hydrocarbons as potential stressors to fish stocks. Deep Sea Research Part II. Vol. 56, Issues 21-22, pp. 2037–2043.

5 Background documents and assessment criteria

Agenda item:

To receive Background Documents and draft assessment criteria from ICES WGBEC on:

1) Acetyl cholinesterase
2) Mussel histopathology
3) Micronucleus and Comet assay
4) MT and ALA-D
5) Intersex in fish

Outcomes of document reviews

SGIMC 2010 received and reviewed a series of draft updated Background Documents from the ICES WGBEC, where they were available. SGIMC made amendments where necessary, and recommends that the documents in Annexes be taken forward for further work and completion in 2011, or are transferred to OSPAR for incorporation into the OSPAR JAMP Background Document on Biological Effects Monitoring Techniques, as indicated in Table 5.1 below.

Table 5.1. Background documents received from WGBEC 2010 and reviewed and further developed at SGIMC 2010.

DOCUMENT	LOCATION	STATUS
Background document on Acetylcholinesterase as a method for assessing neurotoxic effects in aquatic organisms	Annex 10	Recommended for adoption by OSPAR
Background document on mussel histopathology		Draft not yet available. Defer to WGBEC 2011.
Background document on Micronucleus assay as a method for assessing DNA damage in marine organisms	Annex 11	Recommended for adoption by OSPAR
Background document on Comet assay as a method for assessing DNA damage in aquatic organisms	Annex 12	Recommended for adoption by OSPAR, noting further technical development requirements
Background document on Assessment criteria for d-aminolevulinic acid dehydratase (ALA-D) measured in fish blood		Current document considered satisfactory. Assessment criteria need to be reviewed using 90% iles and threshold statistics.
Background document on Intersex in fish		Document not yet available. Deferred to WGBEC 2011.
Background document on hepatic metallothionein in fish		Current text satisfactory, but only refers to fish. Document should be extended to cover metallothionein in shellfish

Observations on background AChE activity in mussels

Kari Lehtonen made available to SGIMC data on AChE activity in *Mytilus edulis* samples from two background locations in the northern Baltic Sea. A regression (solid line) was fitted to the data, together with the 95% confidence band (dashed lines) for the regression line.

It can be concluded that there is a seasonal effect on AchE activity in mussels, because no straight horizontal line would fit into the confidence band over the whole range of

observations without intersecting the confidence band limits. The data suggest that maximum activities (September) are approximately twice the minimum activity reported (late autumn and spring).

The development of assessment criteria, particularly background levels of activity, will need to take account of such seasonal patterns.

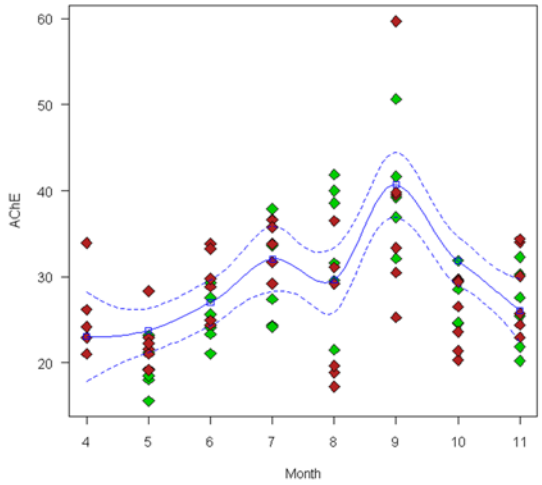


Figure 5.1. Loess smoother fitted to seasonal data on AchE activity in mussels from Granbusken (green), Sundholmen (red) in the Baltic Sea.

Reviews of integrated monitoring strategies and associated Background Documents

SGIMC reviewed the integrated monitoring strategies for fish and shellfish, and the availability and status of Background Documents for elements of the strategies. As a consequence of this review, a series of additional background documents were written, or existing documents revised, as shown in Table 5.2 below.

Table 5.2. Background documents reviewed and developed at SGIMC 2010.

DOCUMENT	LOCATION	STATUS
Background document on reproductive success in eelpout (*Zoarces viviparus*)	Annex 13	Recommended for adoption by OSPAR
Background document on Measurement of supporting metrics for fish: condition indices, GSI, HSI	Annex 14	Incomplete. Defer to SGIMC/WGBEC 2011
Background document on Stress on Stress (SoS) in bivalve molluscs	Annex 15	Recommended for adoption by OSPAR

6 Recap of the integrated monitoring framework

The integrated monitoring frameworks for fish, shellfish and contaminant-based monitoring were reviewed, and updated versions of the relevant Technical Annexes are included as Annexes 16, 17 and 18.

7 Status on JRC/ICES GES 8 Report

The European Commission had asked ICES and the Joint Research Centre to establish a series of expert Task Groups at assist them in interpreting the eleven Descriptors of Good Environmental Status in the Marine Strategy Framework Directive into practical monitoring activities, by giving guidance on assessment criteria and methodological standards. Descriptor 8 (Concentrations of contaminants are at levels not giving rise to pollution effects) is most relevant to SGIMC.

Ian Davies and Dick Vethaak are members of the Task Group for Descriptor 8. They reported on progress of the Task Group. The main points were:

- The TG will recommend monitoring procedures including both chemical measurements in biota, sediment and water, and biological effects measurements covering various levels of biological organization.
- The data will be interpreted using assessment criteria similar to those developed for Water Framework Directive (EQS) and in ICES/OSPAR (EAC). To meet GES, concentrations should not exceed assessment thresholds identified on the basis of toxicological data, and pollution effects should not exceed thresholds representing harm at organism, population, community and ecosystem levels, and concentrations and effects should not be increasing.
- The parameters to measure will not be defined, but the selection of contaminants, monitoring species and biological effects measurements should be made for each assessment region by the Member States (MS) with responsibility for implementation of MSFD in each region.
- Monitoring data should also be assessed against background conditions to provide early warning of deteriorating conditions.
- Integration procedures were not defined, but the work in ICES and OSPAR was recognized as a good foundation for development.
- Advantage should be taken of international experience in programme design and harmonization of sampling and analytical procedures.

The Report would be completed within the next few days and submitted to the GES Management Group.

8 To elaborate plans for the SGIMC training Workshop on Lysosomal Stability data quality and interpretation proposed to be held in Spain in 2010

Concepción Martínez-Gómez presented an update of the workshop plan originally to be held in Murcia Spain in summer 2010. At WGBEC 2010, however, it was decided to explore the possibilities of amalgamating the proposed ICES/OSPAR Workshop with a training workshop to be organized for MEDPOL by Prof Aldo Viarengo in September 2010 in Alessandria, Italy. This seemed opportune because of the presence of high level experts at the MEDPOL workshop and also in line with the requested harmonization of methods between OSPAR and MEDPOL. The Case for a joint ICES/OSPAR/MEDPOL Workshop to be held in association with a MEDPOL training course, including proposed Terms of Reference for the ICES/OSPAR component on lysosomal stability data quality and interpretation, were drafted during the meeting (see Annex 20). SGIMC recommended that the Case be sent to ICES and OSPAR for approval.

Meanwhile, Concepción Martínez-Gómez would continue communications with Prof Viarengo with a view to developing a coordinated programme for the Workshop. In addition to lysosomal stability, participants of Joint Workshop would also receive training on practical aspects of the stress on stress biomarker method. This method is included in the two-tiered approach of MEDPOL as well as in the core methods of the OSPAR integrated monitoring scheme for mussels, and requires harmonization.

9 Progress made on draft merged guidelines for the integrated monitoring and assessment of contaminants and their effects and associated Technical Annexes

To continue towards the completion of the draft merged Guidelines for the Integrated Monitoring and Assessment of Contaminants and their effects and drafting the necessary Technical Annexes, including an annex on survey design, in the light of advice and work packages developed intersessionally by the co-chairs, and emerging requirements in other fora, for example implementation of WFD and MSFD (OSPAR request (2008/8).

SGIMC drafted and/or reviewed a series of Technical Annexes as contributions to the OSPAR Guidelines for the Integrated Monitoring and Assessment of Contaminants and their effects. SGIMC addressed the integrated frameworks developed during the WKIMON series and produced in the WKIMON III and IV reports as well as published in the peer review literature (Thain J, Vethaak D, Hylland K. ICES J Mar Sci 1508). The integrated schemes for fish, mussels and sediment were revisited and the following changes made to take account of recommendations from SGIMC, WKIMC and WGBEC meetings:

a) Sediment integrated framework

- In vitro bioassays were added to the scheme and DR-CALUX referred to specifically

b) Mussel integrated scheme

- The chemical determinands to be measured in mussels were updated by changing to bold those compounds which are now included in the OSPAR CEMP (also blue text changed to bold).
- The columns were re-ordered left to right, subcellular response, tissue response, whole organism response.
- Scope for growth was down-rated as a 'core' component because very few laboratories are able to do/are doing the method as part of a monitoring programme. Stress on stress response of mussels gives similar information and is applied by a greater number of laboratories. It was considered that SFG should be used where possible but SoS used as an alternative. SoS was therefore included as a 'core' method.
- Solid lines were used in the scheme to indicate 'core' components and italics used to indicate where supporting documentation for certain methods is lacking.
- Gametogenesis was added as a separate method to mussel histopathology and not considered a core method.
- MXR and GST were removed from the scheme following WGBEC recommendations from 2008.
- COMET and MT were no longer considered 'core' methods for the integrated package and micronucleus assay was added as a 'core' method.

c) Fish integrated scheme

- As for mussel scheme, the columns were re-ordered and OSPAR CEMP components highlighted in bold. Italics were used to identify methods that were lacking supporting documentation and solid lined boxes indicate 'core' components.

- Organotins were removed from the chemical determinands to be measured in biota.
- External fish disease was added to the 'whole organism' response column.
- Liver nodules were made a core component of the programme and intersex downgraded from a 'core' component.
- ALA-D was removed from the integrated scheme.
- EROD/CYP1A was missing from the WKIMON IV report version. It was thought that this was in error so this is now included.

d) Revised documents

The outcomes of the reviews are included in Annexes to this report, as shown below (Table 9.1) together with recommendations concerning adoption of the revised documents.

Table 9.1. Reviews of Technical Annexes.

DOCUMENT	LOCATION	STATUS
Technical Annex on sampling and analysis for integrated chemical and biological effects monitoring in fish and shellfish	Annex 16	Recommended for adoption by OSPAR.
Technical Annex for Mussel (*Mytilus* sp.) OSPAR Integrated Monitoring	Annex 17	Recommended for adoption by OSPAR.
Technical Annex on recommended packages of chemical and biological methods for monitoring on a contaminant basis	Annex 18	Recommended for adoption by OSPAR.
Discussion document on survey design for integrated chemical and biological effects monitoring	Annex 19	Recommended for adoption by OSPAR.

Note: The discussion document on survey design for integrated chemical and biological effects monitoring included as Annex 19, is considered by SGIMC to be the best advice currently available on the subject, recognizing that it has not been possible to develop a full solution to the task. A full and statistically robust solution requires information regarding assessment thresholds, integrated assessment methods, etc. that have not yet been developed.

10 Progress made on integrating/aggregating data assessment strategies

10.1 Harmonisation: progress made in the HELCOM area

a) ICES SGEH activities

Kari Lehtonen reported on activities in the ICES Study Group for the Development of Integrated Monitoring and Assessment of Ecosystem Health in the Baltic Sea (SGEH) focuses its main activities on matters related to biological effects of contaminants in marine organisms in the Baltic Sea. Information on the effects of contaminants on biodiversity is also closely followed. To achieve the target of developing assessments of Ecosystem Health in the Baltic Sea links with groups dealing with fisheries and eutrophication impacts will be established with expected participation of experts having data and information relevant to SGEH. Important aspects are identification of links between SGEH work related to HELCOM, OSPAR, EU (with a special reference to MSFD) and other ICES EGs, especially SGIMC, WGBEC, WGIAB. In regard to the MSFD, suggested criteria and methodological standards for the descriptors will be discussed in this group in the next meeting. SGEH will also follow closely the outcome of the Quality Status Report of OSPAR. Progress made through the BONUS+ programme BEAST project and other similar activities in and outside the Baltic Sea will be reviewed. SGEH will meet in Gdynia (PL) in March 1–5, 2010, chaired by Kari Lehtonen (FI).

b) BONUS+ Programme BEAST project

Kari Lehtonen reported on the BEAST project (Biological Effects of Anthropogenic Chemical Stress: Tools for the Assessment of Ecosystem Health) was launched under the Baltic Sea BONUS+ Programme (2009–2011) and consists of 16 partners from all nine Baltic Sea countries. Detailed information on BEAST is available in the WGBEC Report 2009, at the BONUS+ website (http://www.bonusportal.org/research_projects) and at BEAST website (http://www.environment.fi/syke/beast). In short, BEAST consists of three thematic Work Packages (WP):

- WP1: Field studies and experiments in selected subregions of the Baltic Sea; basic research: testing and validation of biomarkers in Baltic Sea species and environmental conditions.
- WP2: Application and validation of methods in monitoring and assessment in the Baltic Sea; recommendations and practical guidelines for the integration of chemical-biological monitoring of hazardous substances in Baltic Sea monitoring programmes (mainly HELCOM).
- WP3: Developing tools for ecosystem health assessment in the Baltic Sea; testing and developing approaches (e.g. indices) for the assessment of Ecosystem Health in different subregions of the Baltic Sea.

Research activities in the three WPs are organized under five subregional Tasks, i.e. field and experimental studies in the Gulf of Bothnia, G. of Finland, G. of Riga, G. of Gdańsk and the Belt Sea.

BEAST WP2 and WP3 will closely follow OSPAR developments in regard to developments in integrated monitoring and assessment approaches to enable harmonization of methodologies between the sea areas.

c) HELCOM Thematic Assessment of hazardous substances in the Baltic Sea 2010.

HELCOM Thematic Assessment of hazardous substances in the Baltic Sea 2010.

Dr Jakob Strand presented the CHASE assessment tool, which has been used to prepare a HELCOM integrated thematic assessment of hazardous substances in the Baltic Sea. A draft report is available and the final version is to be published later this year (HELCOM 2010, Eds: Korpinen and Laamanen). The report is connected to the HELCOM Baltic Sea Action Plan, which identifies pollution by hazardous substances as one of the four main issues to improve the health of the Baltic Sea.

The CHASE assessment tool is used to create an integrated traffic light system. CHASE is used to integrate the status of contamination by individual chemicals and biological effects at specific sites or areas into a single status value called contamination ratio. Ultimately, the use of this integrative tool provides an overview of the status of contamination and biological effects by hazardous substances over the Baltic Sea.

The threshold levels used in CHASE were obtained from national legislation or international agreements or EU directives (e.g. EC Environmental Quality Standards, OSPAR Environmental Assessment Criteria). The use of national or international threshold levels ensures that this assessment is fully compatible with national legislation and the implementation of the European Union directives.

Altogether, 144 assessment units were analysed using CHASE, 40 of the assessment units were open sea areas and 114 were coastal assessment units.

CHASE includes four themes;

Contaminants close to natural levels (water, sediment, biota)

Fish safe to eat

Healthy wildlife

Radionuclides

CHASE calculates a "Contamination Ratio (CR)" for all variables within each theme: CR = measured value / threshold level

CR are integrated within each theme to a "Contamination sum", which is calculated as “sum of all Contaminant ratios/ (sqrt n)”

The contamination sum is assessed according to five assessment classes: High (0–0,5), Good (0.5–1), Moderate (1–5), Poor (5–10), and Bad (>10). The overall integrated assessment is based on the worst condition of the four themes. Assessments for at least two themes have to be included in the calculation before the overall assessment is performed.

Each of the CHASE classifications with moderate, poor or bad status were examined to determine whether there were common causes for the poor status. The three most common substances causing the degraded status of the marine environment (i.e. having highest CRs) were collated.

All common groups of hazardous substances: PCBs, dioxins, heavy metals, organometals, alkylphenols, phtalates, brominated substances, polycyclic aromatic hydrocarbons (PAH), DDTs and chlorinated pesticides were found among the decisive substances.

PCBs were among the top three substances in 39% of the cases,

Mercury in 22%,

Cadmium in 19%,

Tributyltin in 18% and

Lead in 18%.

Dioxins/dioxin-like PCBs, DDT and DDE had a high influence on the chemical status in ca. 10% of the cases.

The above-mentioned 'decisive substances' were mainly found in samples from biota (60%, fish, mussels and birds) and only secondarily in sediment samples (35%). Only five per cent of those samples were from water.

Biological effects measurements, such as imposex in marine snails and lysosomal membrane stability, micronuclei and PAH-metabolites in fish have also impact on the assessment in some areas, mainly in the southern part of the Baltic Sea, where data were available.

The assessment would be published as: HELCOM 2010. Hazardous substances in the Baltic Sea. An Integrated Thematic Assessment of hazardous substances in the Baltic Sea. Baltic Sea Environment Proceedings No. 120B. Samuli Korpinen and Maria Laamanen (Eds.).

SGIMC welcomed the work that was in progress in HELCOM, and expressed thanks for the early prepublication viewing of the analysis. The main points made during discussions were:

SGIMC appreciated the comprehensive work by HELCOM on an integrated thematic assessment of hazardous substances in the Baltic Sea which seems to be an important progress for the assessment of the environmental conditions in the Baltic Sea. It was recognized that the approach for the integration of chemical concentration levels and biological effects measurements are an important step forward, but it was also recognized that there is still room for improvement in future.

The extensive collation of standards/thresholds applied in a range of countries and contexts would be a useful resource for other groups.

The use of thresholds for both background assessment levels (e.g. OSPAR BAC) and thresholds that address the risk of adverse effects in marine organisms (e.g. EC WFD-EQS or OSPAR EAC) in the calculation of CRs based on chemical concentration levels can confuse the uniform interpretation of the meaning of CRs. It will not be fully clear if CR>1 reflect concentration levels above natural background levels or concentration levels that pose a risk for marine organisms. This will also have importance for the interpretation for the overall contamination sum used in the integrated assessment for the theme "Contaminants close to natural levels".

It appeared possible that different assessment thresholds for the same parameter could be used in different countries or assessment areas. This raised concerns for the comparability of the overall assessment across the Baltic area.

The use of contamination ratios (CRs) without any normalization could lead to distortions in the assessment. For example the scale of variation found in different biological effects measurements, which are included in the CHASE theme healthy wildlife, is known to differ greatly and in some case up to order of magnitudes for induction for some biomarkers. This will be reflected in the calculated CRs and subsequently also

the final score (i.e. sum of CRs) and thereby the overall assessment by the contamination sum. It was recommended to normalize the data in some way, so that the relative weightings of the various parameters are more equalised. This would lead to more comparable CRs between parameters, which then can be better combined to an overall contamination sum.

A consequence of using a square root function in the calculation of the Contamination Sum is that, all other things being equal, the more data (more parameters) that are available, the worse the final Contamination Sum and assessment. SGIMC recognizes that this is precautionary, but further documentation is needed to show that the square root functions is the best way to counter the sum of CRs increase with the number of data assessed compared with the "dilution" of higher CRs by adding more (low CR) data.

SGIMC looked forward to the publication of the final report, as a contribution towards the widespread efforts to develop robust multivariate environmental assessments.

10.2 Data integration initiatives in OSPAR

a) Traffic lights and integrated assessment

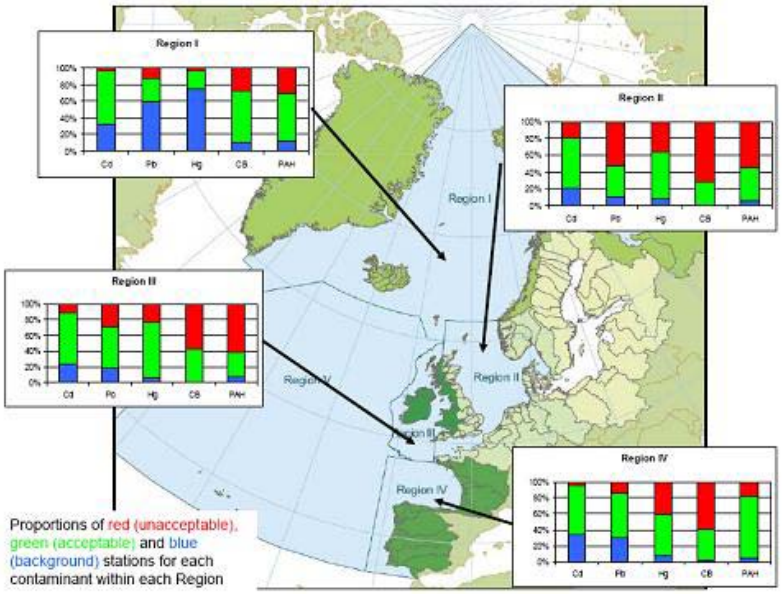
Ian Davies opened this agenda item by describing the process that had been developed by OSPAR MON for the aggregation and presentation of chemical data on the concentrations of contaminants in sediment and biota for the QSR 2010. The purpose of the work had been to present the monitoring data in relation to the OSPAR objectives for hazardous substances, i.e. that concentrations should be at or close to background, and that there should be no unexpected/unacceptable biological effects arising from hazardous substances.

The fundamental elements of the process are a coherent and consistent set of assessment criteria for hazardous substances in the monitoring matrices, tools for the statistical analysis of data, including normalization of concentrations and temporal trends assessments.

The assessment criteria (BACs and EACs) were used as the thresholds in a three colour traffic light system. Concentrations of each contaminant at each station in each matrix were compared with the thresholds and converted into assessments expressed as colours.

The OSPAR area was divided into a series of subareas for data integration, in which offshore areas were assessed separately from inshore areas (within 12 miles of the coast). Within each assessment area, the assessments for each contaminant, expressed as colours, were combined to give an overall assessment for each contaminant. This was expressed as a bar containing appropriate proportions of the three colours, so that as much information as possible was retained through to the final presentation. The information used at underlying stages was also retained to enable full traceability from the final assessment back to the original data.

An example of the final presentation is shown in Figure below.



Ian Davies also presented the approach to integration developed at the WKIMC meeting in Aberdeen in October 2009. This approach was also based around a three colour traffic light scheme, and assessment criteria corresponding to background levels of response of biological effects measurements (analogous to BACs) and to levels of biological response that represent harm to organisms (analogous to EACs). The WKIMC scheme had recognized biomarkers of exposure for which background responses could be defined but upper assessment thresholds could not, and biomarkers of effect for which both types of thresholds would be appropriate.

The WKIMC scheme had been discussed by WGBEC who had provided various comments. The WGBEC draft Report also included a proposal for an alternative approach.

There was extensive discussion of the WKIMC approach. The main points agreed were:

- Assessment schemes for biological effects based on two assessment thresholds (as described by WKIMC) seemed the best way forward at this time to address the objectives of the OSPAR hazardous substances strategy.
- Biological effects could be integrated with the MON assessment output as additional variables, i.e. presented in the same way as individual contaminants. This might become unwieldy if data were available for a large number of effects measurements.
- Alternative presentations were discussed, based upon either the integrated schemes for fish and shellfish, or integrated schemes for particular categories of contaminants. For example, a presentation based upon the PAH theme could be developed, or one based upon the metals theme. This approach would have the advantage of limiting the number of parameters to display, while presenting a coherent thematic assessment.

It was felt that the majority of the WGBEC comments were addressed by the conclusions above. Other relevant comments were:

Red responses (i.e. values above EACs for chemicals and above an upper assessment threshold for effects) were equivalent in the context of the OSPAR objectives. Differ-

ent approaches might be appropriate in assessments for other purposes, for example to derive expressions of ecosystem health.

Differences in the significance of expressions of effect at molecular/ cellular/ tissue/ organism levels were to a degree addressed through the two threshold assessment scheme. Other approaches might be appropriate to other purposes.

The use of statistical approaches such as PRIMER E does not assist in meeting OSPAR objectives.

The Study Group considered the level of significance that should be given to exceedence of BAC and EAC for different biological effects measurements in an integrated assessment:

Biomarkers for environmental pollution are assessed at different levels of biological organization. They range from biochemical parameters as indicators for specific contaminant exposure to pathological endpoints at subcellular and tissue level up to responses at the levels of behaviour, reproduction, health, and biodiversity. This approach has been developed to obtain information on the status of environmental deterioration due to pollution as well as on the progress of damage on the biological system scale (Broeg *et al.,* 2005; Dagnino *et al*., 2007). To finally extract this additional information from the dataset, one measure is the weighting of the various biomarkers according to their level of biological organization and their relevance for the maintenance animals' health, for the population structure, and finally the community. This course of action is aimed in giving biomarkers of exposure, early-, and late biological effects different weights for the final integrated assessment, considering that not all biomarker responses are of the same relevance for the maintenance of environmental integrity.

Based on the heuristic model of "syntropic ecotoxicology", Downs and Ambrose, 2001 showed that effects of contaminants proceed from the lower biological levels to the next higher one, if physiological responses of the organism to pollutant impacts are not able to handle the stress at lower levels, or to repair damage that already had occurred (Figure 10.1).

SGIMC included all biomarker which are mentioned in the fish and mussel schemes for an integrated assessment, core biomarker as well as promising ones (see above in this report). Proposed is a two dimensional grading: the first dimension is based on the assessment criteria (BAC, EAC) of each single biomarker, the second dimension represents the location of each biomarker on the scale of its relevance for the ecosystem health. This is shown in the Figure below.

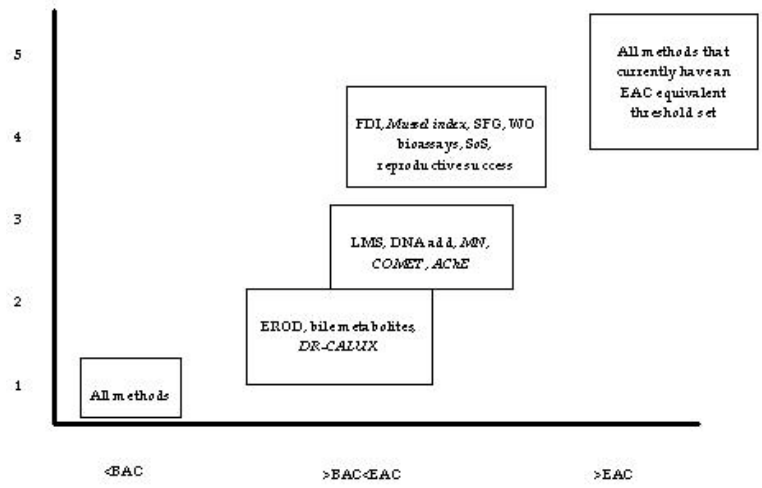


Figure 10.1. Draft integration/weighting of biological effects data in an integrated assessment framework. Methods in italics still require development of relevant assessment criteria, but these are in progress. There is some uncertainty over the validity of applying EAC equivalent assessment criteria for PAH bile metabolites (due to inhibition at highly contaminated field sites) and whether suitable EAC equivalent assessment criteria for whole organism bioassays can be derived in a manner that would allow them to be given equal weighting in an integrated assessment framework as other methods.

References

Broeg, K., v. Westernhagen, H., Zander, S., Körting, W., Koehler, A. 2005. The "bioeffect assessment index" (BAI): A concept for the quantification of effects of marine pollution by an integrated biomarker approach. Marine Pollution Bulletin 50, 495–503.

A. Dagnino, A., Allen, J.I., Moore, M.N., Broeg, K., Canesi, L., Viarengo, A. 2007. Development of an expert system for the integration of biomarker responses in mussels into an animal health index. Biomarkers 12(2) 155–172.

Downs, T. and R. Ambrose. 2001. Syntropic ecotoxicology: a heuristic model for understanding the vulnerability of ecological systems to stress. Ecosystem Health 7: 266–283.

11 Recommendations and proposals for OSPAR for adopting changes in BG documents, Technical Annexes and Assessment Criteria

SGIMC recommended that the following documents be recommended to OSPAR for adoption:

Document	Location	Status
Background document on Cytochrome P4501A activity (EROD) – Chapter 3	Annex 4	Recommended for adoption by OSPAR
Background document on Lysosomal stability as a global health status indicator in biomonitoring (Chapter 4)	Annex 5	Recommended for adoption by OSPAR
Background document on DNA adducts (Chapter 11)	Annex 6	Recommended for adoption by OSPAR
Background document on PAH metabolites in bile (Chapter 5)	Annex 7	Recommended for adoption by OSPAR
Background document on Externally visible fish diseases, macroscopic liver neoplasms and liver histopathology (Chapter 6)	Annex 8	Recommended for adoption by OSPAR
Background document on Water Bioassays	Annex 9	Recommended for adoption by OSPAR
Background document on Acetylcholinesterase as a method for assessing neurotoxic effects in aquatic organisms	Annex 10	Recommended for adoption by OSPAR
Background document on Comet assay as a method for assessing DNA damage in aquatic organisms	Annex 12	Recommended for adoption by OSPAR, noting further technical development requirements
Background document on Assessment criteria for d-aminolevulinic acid dehydratase (ALA-D) measured in fish blood		Current document considered satisfactory. Assessment criteria need to be reviewed using 90% iles and threshold statistics.
Background document on hepatic metallothionein in fish		Current text satisfactory, but only refers to fish. Document should be extended to cover metallothionein in shellfish
Background document on reproductive success in eelpout (Zoarces viviparus)	Annex 13	Recommended for adoption by OSPAR
Background document on Stress on Stress (SoS) in bivalve molluscs	Annex 15	Recommended for adoption by OSPAR
Technical Annex on sampling and analysis for integrated chemical and biological effects monitoring in fish and shellfish	Annex 16	Recommended for adoption by OSPAR.
Technical Annex for Mussel (Mytilus sp.) OSPAR Integrated Monitoring	Annex 17	Recommended for adoption by OSPAR.
Technical annex on recommended packages of chemical and biological methods for monitoring on a contaminant basis	Annex 18	Recommended for adoption by OSPAR.
Discussion document on survey design for integrated chemical and biological effects monitoring	Annex 19	Recommended for adoption by OSPAR.

Note: The discussion document on survey design for integrated chemical and biological effects monitoring included as Annex 19, is considered by SGIMC to be the best advice currently available on the subject, recognizing that it has not been possible to develop a full solution to the task. A full and statistically robust solution requires information regarding assessment thresholds, integrated assessment methods etc that have not yet been developed.

12 Update of Workplan for 2010/2011

Proposed Work Programme for SGIMC from January 2009 to January 2011.

Effect	Task	Responsible member	When	Report to	Status January 2010
EROD	Organising Workshop	Ian Davies	For Oct 2009	SGIMC 2010	Completed
EROD	Update Background Document and develop improved approach to Background Response assessment criteria	Ian Davies and others	Oct 2009	SGIMC 2010	Completed
PAH bile metabolites	Update Background Document	Dick Vethaak and Ketil Hylland	Mar 09	WGBEC 2009	Completed
PAH bile metabolites	To develop Background Response assessment criteria	Ketil Hylland (and Ian Davies)	Oct 2009	SGIMC 2010	Completed
DNA adducts	Update Background Document	Brett Lyons (and Ian Davies)	Oct 2009	SGIMC 2010	Completed
DNA adducts	To develop Background Response and EAC-equivalent assessment criteria	Brett Lyons (and Ian Davies)		SGIMC 2011	To be reviewed by SGIMC 2011
DR-CALUX	Prepare Background Document	Dick Vethaak	September 2010	SGIMC 2011	In progress; to be reviewed at SGIMC 2011
DR-CALUX	To develop Background Response and EAC-equivalent assessment criteria	Dick Vethaak	Oct 2009	SGIMC 2011	Preliminary AC available but need refinement. To be reviewed at SGIMC 2011
DR-CALUX	Complete TIMES series method document	Dick Vethaak	March 2010	SGIMC 2010	In progress; will be send to editor March 2010

Effect	Task	Responsible member	When	Report to	Status January 2010
Liver nodules (neoplasm)	To develop Background Response and EAC-equivalent assessment criteria	Thomas Lang and Dick Vethaak	Oct 2009	SGIMC 2010	Completed
Extraction procedures for bioassay methods	Complete TIMES series method document	Dick Vethaak + John Thain	Imminent	WGBEC 2011	In progress, to be reviewed at SGIMC 2011
VTG	Establish BAC in monitoring species	Ian Davies and Dick Vethaak	For SGIMC 2010	SGIMC 2010	Completed
VTG	Develop EAC equivalent for monitoring species	Ian Davies and Dick Vethaak	SGIMC 2011	SGIMC 2011	To be developed at SGIMC 2011
Intersex in fish	Review Background document	Steve Feist	WGBEC2011	SGIMC 2011	To be reviewed at SGIMC 2011
Intersex in fish	To develop Background Response and EAC-equivalent assessment criteria	Steve Feist (and Ian Davies)	SGIMC 2010	SGIMC 2011	To be reviewed by SGIMC 2011
Fish Disease Index.	No action required by SGIMC				No action required by SGIMC
Reproductive success (eelpout).	Review BG document and TA		SGIMC 2010	SGIMC 2010	Completed
Background document on Supporting parameters in fish: condition indices, SLI and SGI	develop BD	John Thain, Dick Vethaak	WGBEC 2011	SGIMC 2011	SGIMC 2011
Lysosomal stability (Neutral Red)	Organising training workshop Draft proposal Permission from ICES/OSPAR	Concepcion Martínez-Gómez	June 2010	SGIMC 2011	Plans developed for joint MEDPOL/OSPAR/ICES training Workshop to be held in Alexandria Sept 2010. Draft proposal completed
Acetyl cholinesterase	Update Background Document	Thierry Burgeot	WGBEC2011	SGIMC 2010	Completed

Effect	Task	Responsible member	When	Report to	Status January 2010
Acetyl cholinesterase	To develop Background Response assessment criteria	Thierry Burgeot	WGBEC 2011	SGIMC 2011	in progress
Mussel histopathology	ICES Times manuscript incl uding BAC in preparation	Steve Feist + Miren Cajaraville	WGBEC 2011, WGPDMO 2011	SGIMC 2011	will be send to SGICM 2011
Micronucleus assay + comet assay	Background document and draft BAC	Brett Lyon	WGBEC2011	SGIMC 2011	Comet assay completed; micronucl. assay completed
MT & ALA-D	Develop BC using recent data	Ketil Hylland	Feb 2011	SGIMC 2011	Needs expanded to cover mussels. Need assessment criteria for MT and ALA-D
New Chapter	In vitros YES/YAS, ER CALUX	JT / DV	WGBEC 2011	SGIMC	in progress
Chapter 8	Add Sed & SW elutriate bioassays for invert embryos. Further validate others as more data becomes available	Ricardo Beiras	2010	WGBEC 2011	To be reviewed at SGIMC 2011
Chapter 9	As above with copepods	As above	As above	WGBEC 2011	To be reviewed at SGIMC 2011
Chapter 10	Update BG and ass cri for Whole sediments with amphipods as more data becomes available	As above	As above	WGBEC 2011	To be reviewed at SGIMC 2011

13 Adoption of Report and closure of meeting

SGIMC agreed that the Report could be adopted, subject to circulation of full draft by the Co-chairs.

14 Any other business

There was no significant other business.

15 Closure of the meeting

It was agreed that the Co-chairs would lead on completion of the Meeting Report.

The meeting was closed at 1300 h on 29 January 2010.

Annex 1: List of participants

Name	Address	Phone/Fax	Email
Katja Broeg	Alfred-Wegener-Institut Foundation for Polar and Marine Research PO Box 12 0161 D-27515 Bremerhaven Germany		katja.broeg@awi.de
Kris Cooreman	Institute for Agricultural and Fisheries Research Ankerstraat 1 B-8400 Oostende Belgium	Phone +32 59569820 Fax +32 59330629	kris.cooreman@ilvo.vlaanderen.be
Ian M. Davies Co-chair	Marine Scotland Marine Laboratory PO Box 101 AB11 9DB Aberdeen UK	Phone +44 1224 295468 Fax +44 1224 295511	I.M.davies@marlab.ac.uk
Lisa Devriese	Institute for Agricultural and Fisheries Research Ankerstraat 1 B-8400 Oostende Belgium	Phone +32 Fax +32	Lisa.devriese@ilvo.vlaanderen.be
Michelle Giltrap	Trinity College Dublin Zoology Department Dublin 2 Ireland	Phone +353 Fax +353	giltrapm@tcd.ie
Matt J. Gubbins	Marine Scotland Marine Laboratory PO Box 101 AB11 9DB Aberdeen Torry United Kingdom	Phone +44(0)122 429 5681	m.gubbins@marlab.ac.uk
Claus Hagebro ICES Secretariat	International Council for the Exploration of the Sea H. C. Andersens Boulevard 44–46 DK-1553 Copenhagen V Denmark	Phone +45 3338 6754 Fax +45 3393 4215	claus@ices.dk
Ulrike Kammann	Johann Heinrich von Thünen-Institute Institute for Fishery Ecology Palmaille 9 D-22767 Hamburg Germany	Phone +49 4038905198	Ulrike.kammann@vti.bund.de
Kari Lehtonen	Finnish Environment Institute Marine Centre PO Box 140 FI-00251 Helsinki Finland	Phone +358 40 7030305 Fax +358 9 323 2970	kari.lehtonen@ymparisto.fi
Thomas Maes	Centre for Environment, Fisheries and Aquaculture Science Pakefield Road NR33 0HT Lowestoft Suffolk United Kingdom	Phone +44 1502 524433 Fax +44 1502 513865	thomas.maes@cefas.co.uk

Name	Address	Phone/Fax	Email
Concepción Martínez-Gómez	Instituto Español de Oceanografía Centro Oceanográfico de Murcia P.O. Box 22 E-30740 San Pedro del Pinatar Murcia Spain	Phone +34 Fax +34	cmartinez@mu.ieo.es
Johan Robbens	Institute for Agricultural and Fisheries Research Ankerstraat 1 B-8400 Oostende Belgium	Phone +32 Fax +32	johan.robbens@ilvo.vlaanderen.be
Steinar Sanni	International Research Institute of Stavanger PO Box 8046 N-4068 Stavanger Norway		Steinar.Sanni@iris.no
John Thain	Centre for Environment, Fisheries and Aquaculture Science Weymouth Laboratory The Nothe, Barrack Road DT4 8UB Weymouth Dorset United Kingdom	Phone +44 (0)1305 20 6617 Fax +44 (0)1305 20 6601	john.thain@cefas.co.uk
Dick Vethaak Co-chair	Ecosystem Analysis and Assessment (ESA) Unit Marine and Coastal Systems Rotterdamseweg 185 HD 2629 Delft The Netherlands	Phone +31 015-2858659/+31 065-1232412	dick.vethaak@deltares.nl
Werner Wosniok	University of Bremen Institut für Statistik, FB 3 PO Box 330 440 DE-28344 Bremen Germany	Phone +49 42 12 1834 71 Fax +49 42 12 18 8944	wwosniok@math.uni-bremen.de

Annex 2: Agenda

Draft Agenda and tentative timetable SGIMC 25–29 January 2010.

DATE	APPROX TIME	AGENDA ITEM	ISSUE	LEAD
Monday 25 Jan	09:30	1	Introduction by Co-chairs, tour de table	ID, DV
,,	09.45	2	Adoption of agenda	ID, DV
,,		3	Review progress made by Aberdeen WKIMC report and to receive Background Documents and draft assessment criteria from SGIMC Aberdeen Workshop planned for October 2009 on: EROD Bile metabolites DNA adducts DR-CALUX Contaminant-linked fish disease (liver neoplasm) and to assess their usefulness in integrated assessments.	ID, DV
,,		4	To receive Background Documents and draft assessment criteria from ICES WGBEC on: Acetyl cholinesterase Mussel histopathology Micronucleus and Comet assay MT and ALA-D Intersex in fish	JT/MG
,,		5	Recap of integrated monitoring framework: are we on the right track?; are the right biomarkers and bioassays included?	JT, DV
		6	Status on JRC/ICES GES 8 TG report	ID
,,		7	To elaborate plans for the SGIMC Workshop proposed to be held in Spain in summer/autumn 2010.	CMG
,,		8	Discussion and further work on integrating/ aggregating data assessment strategies (including feedback from WGBEC).	ID, JT
,,		9	To continue towards the completion of the draft merged Guidelines for the Integrated Monitoring and Assessment of Contaminants and their effects and drafting the necessary Technical Annexes, including an annex on survey design, in the light of advice and work packages developed intersessionally by the co-chairs, and emerging requirements in other fora, for example implementation of WFD and MSFD (OSPAR request (2008/8).	DV, MG TM, etc
26–28 Jan	plenary reporting to SG		Sub-group sessionss Subgroup(s) working on assessment criteria for specific methods Sub-group on finalizing the merged OSPAR guidelines.	 WW, etc DV, MG, TM, JT, etc
29 Jan		12	Recommendations and proposals for OSPAR for adopting changes in BG documents and AC.	ID, All
		13	Any other business.	ID, DV
28 Jan	15.00	14	Adoption of the Report and closure of the meeting.	ID, DV
			SGIMC will report by 1 March 2010 for the attention of ACOM.	

Annex 3a: SGIMC draft Terms of Reference for the 2011 meeting

2010/2/ACOM30 The Study Group on Integrated Monitoring of Contaminants and Biological Effects (SGIMC) (Co-chairs: Dr Ian M Davies, UK and Dick Vethaak, NL will meet at ICES Headquarters late January 2011 to:

a) To receive such new and revised Background Documents and draft assessment criteria as have been prepared intersessionally by ICES SGIMC and WGBEC and to assess their usefulness in integrated assessments;

b) To review comments on Background Documents and draft assessment criteria from OSPAR and ICES groups and to revised documents as necessary;

c) To complete the outstanding tasks in the SGIMC Work programme (Section 12 of the 2010 Report);

d) To complete the draft merged Guidelines for the Integrated Monitoring and Assessment of Contaminants and their effects and drafting the necessary Technical Annexes, in the light of advice and work packages developed intersessionally by the Co-chairs, and emerging requirements in other fora, for example implementation of WFD and MSFD;

e) To bring the work of SGIMC to a conclusion and to recommend how any outstanding tasks should be addressed.

Material and data relevant to the meeting must be available to the group no later than 14 days prior to the starting date.

SGIMC will report by 1 March 2011 for the attention of ACOM.

Supporting Information

Priority:	The current activities of this joint ICES/OSPAR Study Group are firmly directed at meeting the needs of both organizations for integration of chemical and biological effects measurements. The output also has strong application to the development of moinitoring stratgies and assessment criteria for Descriptor 8 of Good Environmental Status under Marine Strategy Framework Directive.
Scientific justification and relation to action plan:	Term of Reference a) Various outputs are expected to be delivered through the WGBEC and SGIMC processes in 2010, either through agenda items or TIMES series documents. SGIMC will take these outputs and consider how the proposed assessment criteria could be used in integrated assessment. Term of Reference b) SGIMC 2010 recommended a large number of dicuments for adoption by OSPAR. This task is to review the repsonse of ICES and OSPAR to this recommendation and to revise documents as necessary. Term of Reference c) Considerable progress has been made with the detailed work programme items. Outstanding items will be completed in 2011. Term of Reference d) This is an OSPAR request (2008/8). The details of this task witll be elaborated by the co-chairs intersessionally. The completion of the Guidelines and associated technical Annexes is an important step in the implementation of integrated monitoring. Term of Reference e) The 2011 meeting will be the third meeting of SGIMC. The SG should complete all outstanding tasks, and if this is not possible recommend how they should be addressed, with a view to recommending that the SG be terminated.
Resource requirements:	The research and monitoring programmes which provide the main input to this group are already underway. Additional activities are proposed through a series of Workshops. The additional resource required to undertake activities in the through meetings of SGIMC is negligible.
Participants:	Anticipated attendance is approximately 20 people. The developments under MSFD may attract visitors from outside the ICES.
Secretariat facilities:	None.
Financial:	No financial implications.
Linkages to advisory committees:	The output from this SG is directly linked to the needs of OSPAR (and probably the EU) for advice on assessment methods and criteria, and integration of chemical and biological monitring.
Linkages to other committees or groups:	There is valuable working relationship between SGIMC and other environmental expert groups, particualrly WGBEC, WGMS and MCWG. The collaboration is partly effected through joint memberships.
Linkages to other organizations:	The work of this group is closely aligned with the work of OSPAR (ASMO) and HELCOM, for example data assessments by OSPAR MON. It is likely that MSFD will lead to linkages with other regional Conventions.

Annex 3b: Recommendations

SGIMC recommended that the following documents be recommended to OSPAR for adoption:

DOCUMENT	LOCATION	STATUS
Background document on Cytochrome P4501A activity (EROD) – Chapter 3	Annex 4	Recommended for adoption by OSPAR
Background document on Lysosomal stability as a global health status indicator in biomonitoring (Chapter 4)	Annex 5	Recommended for adoption by OSPAR
Background document on DNA adducts (Chapter 11)	Annex 6	Recommended for adoption by OSPAR
Background document on PAH metabolites in bile (Chapter 5)	Annex 7	Recommended for adoption by OSPAR
Background document on Externally visible fish diseases, macroscopic liver neoplasms and liver histopathology (Chapter 6)	Annex 8	Recommended for adoption by OSPAR
Background document on Water Bioassays	Annex 9	Recommended for adoption by OSPAR
Background document on Acetylcholinesterase as a method for assessing neurotoxic effects in aquatic organisms	Annex 10	Recommended for adoption by OSPAR
Background document on Comet assay as a method for assessing DNA damage in aquatic organisms	Annex 12	Recommended for adoption by OSPAR, noting further technical development requirements
Background document on Assessment criteria for d-aminolevulinic acid dehydratase (ALA-D) measured in fish blood		Current document considered satisfactory. Assessment criteria need to be reviewed using 90% iles and threshold statistics.
Background document on hepatic metallothionein in fish		Current text satisfactory, but only refers to fish. Document should be extended to cover metallothionein in shellfish
Background document on reproductive success in eelpout (Zoarces viviparus)	Annex 13	Recommended for adoption by OSPAR
Background document on Stress on Stress (SoS) in bivalve molluscs	Annex 15	Recommended for adoption by OSPAR
Technical Annex on sampling and analysis for integrated chemical and biological effects monitoring in fish and shellfish	Annex 16	Recommended for adoption by OSPAR.
Technical Annex for Mussel (Mytilus sp.) OSPAR Integrated Monitoring	Annex 17	Recommended for adoption by OSPAR.
Technical annex on recommended packages of chemical and biological methods for monitoring on a contaminant basis	Annex 18	Recommended for adoption by OSPAR.
Discussion document on survey design for integrated chemical and biological effects monitoring	Annex 19	Recommended for adoption by OSPAR.

Note: The discussion document on survey design for integrated chemical and biological effects monitoring included as Annex 19, is considered by SGIMC to be the best advice currently available on the subject, recognizing that it has not been possible to develop a full solution to the task. A full and statistically robust solution requires information regarding assessment thresholds, integrated assessment methods, etc. that have not yet been developed.

That Annex 20 concerning a training Workshop on Lysosomal Stability quality and data interpretation proposed to be held in association with MEDPOL in Italy in 2010 be forwarded to appropriate ICES and OSPAR Groups for support.

Annex 4: Background Document on Cytochrome P4501A activity (EROD)

Version date: 29 January 2010

Introduction

The cytochrome P450 1A family of enzymes are responsible for the primary metabolism of planar polyaromatic hydrocarbons and PCBs and activate several procarcinogens such as benzo(a)pyrene. 7-ethoxyresorufin is a convenient artificial substratum which was developed as a safe sensitive assay by Burke and Meyer (1974). Thus the term "EROD" has been adopted as a measure of CYP1A activity in aquatic organisms (Stagg and McIntosh, 1998).

In addition to being substrata for biotransformation, planar compounds, such as PAHs, PCBs and Dioxins also induce synthesis of cytochrome P450 1A by binding to the cytosolic Ah (aryl hydrocarbon)-receptor/ARNT complex. Measurement of EROD activity is the tool used currently to quantify this induction. The induction of cytochrome P450 enzymes in fish liver was first suggested as an indicator of environmental contamination in the 1970s by Payne (1976) which has now gained widespread use (see e.g. Förlin *et al*., 1990; George *et al*., 1995; Goksøyr *et al*., 1991; Whyte *et al*., 2000) and standardized by ring testing (BEQUALM).

Dose-response

In a review Whyte *et al*. (2000) rank chemicals according to the level of EROD activity they induce in treated or exposed fish when compared with untreated or control fish. Contaminants that induce EROD less than tenfold above control levels are considered "weak" inducers, 10- to 100-fold are "moderate" inducers, and chemicals that elicit >100-fold induction are considered "strong" inducers. Dioxins, planar PCBs and PAHs (benzo[a]pyrene) are categorized as "strong" inducers. Over 25 studies have observed induction of hepatic EROD by benzo[a]pyrene in 15 species of fish (Whyte *et al*., 2000).

Relevance of other factors

Several endogenous and exogenous factors have been shown to affect hepatic EROD. The most important endogenous factors for most fish species are, gender, reproductive status and season, all of which can be controlled through sampling design. In addition, environmental temperature has been shown to affect EROD (Sleiderink *et al*., 1995; Lange *et al*., 1999). Seasonal cycles in EROD induction have been observed for i.e. rainbow trout (Förlin and Haux, 1990), flounder (Van Westernhagen *et al*., 1981; Hylland *et al*., 1998), plaice (George and Young, 1986) and salmon (Larsen *et al*., 1992), most likely due to both to changes in water temperature and reproductive cycles (which it is not really possible to separate in the field). The main age-related factors are time of exposure/accumulation, food selection and reproductive stage.

Several species have baseline EROD activities within the same order of magnitude among different studies/measurements and also show greater than tenfold EROD induction after contaminant exposure (Whyte *et al*., 2000). These are, however, mostly fresh-water species.

CYP1A expression is suppressed in spawning females due to interference of 17β-estradiol (E_2) (or xenoestrogen) with transcription of the gene. This may also lead to an underestimation of a PAH-type response of EROD activity, however, this hormone also controls the induction of vitellogenin (VTG; egg yolk protein) which is produced by the liver during gonadal recrudescene. Therefore interference of environmental estrogens on CYP1A induction can be assessed.

Dietary factors can be potentially important for the induction of CYP1A. Firstly, of course, AhR ligands can be presented to the organism through the food. Secondly, proper nutrition is a prerequisite for enzyme systems to function properly. Hylland *et al*., (1996) reported an elimination of EROD response (i.e. to control levels) in BaP-treated flounder deprived of food for one month.

Background responses

Baseline levels of EROD in seven marine species have been estimated from results derived from the joint ICES/OSPAR WKIMON III meeting (2007) and recent data submitted to ICES database (Table 1). The fish were from sites which contracting parties consider being reference stations (i.e. no known local sources of contamination or those areas which were not considered unequivocally as reference sites but as those less influenced from human and industrial activity. The datasets from which these values have been derived are described in Table 2. Further information on the baseline levels and dose response of EROD activity in experimental systems and field studies is given in Tables 3 and 4.

Assessment criteria

Background response ranges have been developed as described above and 90th percentiles of values from reference sites can be used to distinguish between 'background' and 'elevated' responses. As many factors are known to influence EROD activity (see above) and it is difficult to correct for all in the assessment of data, it is advisable to include an appropriate reference group in studies that include EROD as an endpoint. The information provided in Table 2 will also allow data to be assessed against the appropriate assessment criteria for fish species, gender, size, sampling season and bottom-water temperature.

Quality assurance

Cytochrome P4501A is possibly the most widely used biomarker. There have been three international intercalibrations for the method, both within BEQUALM. The intercalibrations have pinpointed variability relating to most steps in the analytical process, excepting possibly the enzyme kinetic analysis itself. It is imperative that laboratories have internal quality assurance procedures, e.g. use internal references samples with all batches of analyses.

Acknowledgement

The current review has been derived from an overview prepared for the Norwegian offshore companies through OLF (Hylland *et al*., 2006), the joint Workshop ICES/OSPAR WKIMON III (2007) and the Workshop SGIMC (2009).

Relevant literature (marine and fresh-water fish)

Bernhoft, A., Hektoen, H., Skaare, J. U., and Ingebrigtsen, K. 1994. Tissue Distribution And Effects On Hepatic Xenobiotic-Metabolizing Enzymes Of 2,3,3',4,4'-Pentachlorobiphenyl (Pcb-105) In Cod (*Gadus-morhua*) And Rainbow-Trout (*Oncorhynchus-mykiss*). Environmental Pollution 85, 351–359.

Beyer, J., Sandvik, M., Hylland, K., Fjeld, E., Egaas, E., Aas, E., Skare, J. U., and Goksoyr, A. 1996. Contaminant accumulation and biomarker responses in flounder (*Platichthys flesus* L) and Atlantic cod (*Gadus morhua* L) exposed by caging to polluted sediments in Sorfjorden, Norway. Aquatic Toxicology 36, 75–98.

Burke, M. D., and Mayer, R. T. 1974. Ethoxyresorufin: Direct fluorimetric assay of a microsomal O-dealkylation which is preferentially inducible by 3-methylchoalanthrene. Drug metabolism and Disposition 2, 583–588.

Cooreman, K., Roose, P. and Vyncke, W. 1993. EROD monitoring in dab from the Belgian continental shelf. ICES C.M./E:14.

Förlin L, Haux C. 1990. Sex differences in hepatic cytochrome P-450 monooxygenase activities in rainbow trout during an annual reproductive cycle. J. Endocrinol. 124:207–213.

Forlin, L., Lemaire, P., and Livingstone, D. R. 1995. Comparative-Studies of Hepatic Xenobiotic-Metabolizing and Antioxidant Enzymes in Different Fish Species. Marine Environmental Research 39, 201–204.

George, S. G., Christiansen, J. S., Killie, B., and Wright, J. 1995. Dietary Crude-Oil Exposure During Sexual-Maturation Induces Hepatic Mixed-Function Oxygenase (Cyp1a) Activity At Very-Low Environmental Temperatures In Polar Cod *Boreogadus-saida*. Marine Ecology-Progress Series 122, 307–312.

George, S. G., Wright, J., and Conroy, J. 1995. Temporal Studies Of The Impact Of The Braer Oilspill On Inshore Feral Fish From Shetland, Scotland. Archives Of Environmental Contamination And Toxicology 29, 530–534.

Goksoyr, A., Andersson, T., Buhler, D. R., Stegeman, J. J., Williams, D. E., and Forlin, L. 1991. Immunochemical Cross-Reactivity Of Beta-Naphthoflavone-Inducible Cytochrome P450 (P450ia) In Liver-Microsomes From Different Fish Species And Rat. Fish Physiology And Biochemistry 9, 1–13.

Goksoyr, A., Andersson, T., Hansson, T., Klungsoyr, J., Zhang, Y., and Forlin, L. 1987. Species Characteristics Of The Hepatic Xenobiotic And Steroid Biotransformation Systems Of 2 Teleost Fish, Atlantic Cod (Gadus-Morhua) And Rainbow-Trout (*Salmo-gairdneri*). Toxicology And Applied Pharmacology 89, 347–360.

Goksoyr, A., Beyer, J., Husoy, A. M., Larsen, H. E., Westerheim, K., Wilhelmsen, S., and Klungsoyr, J. 1994. Accumulation And Effects Of Aromatic And Chlorinated Hydrocarbons In Juvenile Atlantic Cod (*Gadus-morhua*) Caged In A Polluted Fjord (Sorfjorden, Norway). Aquatic Toxicology 29, 21–35.

Hasselberg, L., Meier, S., Svardal, A., Hegelund, T., and Celander, M. C. 2004. Effects of alkylphenols on CYP1A and CYP3A expression in first spawning Atlantic cod (*Gadus morhua*). Aquatic Toxicology 67, 303–313.

Hektoen, H., Bernhoft, A., Ingebrigtsen, K., Skaare, J. U., and Goksoyr, A. 1994. Response Of Hepatic Xenobiotic-Metabolizing Enzymes In Rainbow-Trout (*Oncorhynchus-mykiss*) And Cod (*Gadus-morhua*) To 2,3,7,8-Tetrachlorodibenzo-P-Dioxin (2,3,7,8-Tcdd). Aquatic Toxicology 28, 97–106.

Hylland, K., Ruus, A., Børseth, J.F., Bechmann, R., Barsiene, J., Grung, M., Tollefsen, K., Myhre, L.P. 2006. Biomarkers in monitoring - a review. NIVA-report 5205. 106 p.

Hylland K, Sandvik M, Skåre JU, Beyer J, Egaas E, Goksøyr A. 1996. Biomarkers in flounder (*Platichthys flesus*): an evaluation of their use in pollution monitoring. Mar. Environ. Res. 42:223–227.

Kloeppersams, P. J., and Benton, E. 1994. Exposure Of Fish To Biologically Treated Bleached-Kraft Effluent.2. Induction Of Hepatic Cytochrome-P4501a In Mountain Whitefish (*Prosopium-williamsoni*) And Other Species. Environmental Toxicology And Chemistry 13, 1483–1496.

Lange U, Goksoyr A, Siebers D, Karbe L. 1999. Cytochrome P450 1A-dependent enzyme activities in the liver of dab (*Limanda limanda*): kinetics, seasonal changes and detection limits. Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology 123:361–371.

Larsen, H.E., Celander, M. and Goksøyr, A. 1992. The cytochrome P450 system of Atlantic salmon (*Salmo salar*): II. Variations in hepatic catalytic activities and isozyme patterns during an annual reproductive cycle. Fish Physiol. Biochem. 10, 291–301.

Lindesjoo, E., Husoy, A. M., Petterson, I., and Forlin, L. 1996. Histopathological and immunohistochemical studies in roundnose grenadier (*Coryphaenoides rupestris*) in the Skagerrak, North Sea. Marine Environmental Research 42, 229–233.

Livingstone, D. R., Lemaire, P., Matthews, A., Peters, L. D., Porte, C., Fitzpatrick, P. J., Forlin, L., Nasci, C., Fossato, V., Wootton, N., and Goldfarb, P. 1995. Assessment Of The Impact Of Organic Pollutants On Goby (*Zosterisessor ophiocephalus*) And Mussel (*Mytilus-galloprovincialis*) From The Venice Lagoon, Italy - Biochemical-Studies. Marine Environmental Research 39, 235–240.

Martínez-Gómez, C., Fernández, B., Valdés, J., Campillo, J. A., Benedicto, J., Sánchez, F., and Vethaak, A. D. 2009. Evaluation of three-year monitoring with biomarkers in fish following the Prestige oil spill (N Spain). Chemosphere, 74(5): 613–620.

McDonald, S. J., Willett, K. L., Thomsen, J., Beatty, K. B., Connor, K., Narasimhan, T. R., Erickson, C. M., and Safe, S. H. 1996. Sublethal detoxification responses to contaminant exposure associated with offshore production platforms. Canadian Journal of Fisheries and Aquatic Sciences 53, 2606–2617.

Michel, X., Mora, P., Garrigues, P., Budzinski, H., Raoux, C., and Narbonne, J. F. 2001. Cytochrome P450 dependent activities in mussel and fish from coastal marine environment: Field studies on the French coast of the Mediterranean sea. Polycyclic Aromatic Compounds 18, 307–324.

Myers MS, Johnson LL, Collier TK. 2003. Establishing the causal relationship between polycyclic aromatic hydrocarbon (PAH) exposure and hepatic neoplasms and neoplasia-related liver lesions in English sole (*Pleuronectes vetulus*). Hum Ecol Risk Assess 9:67–94.

Payne, J. F. 1976. Field evaluation of benzopyrene hydroxylase induction as a monitor for marine petroleum pollution. Science 191, 945–946.

Peters, L. D., Nasci, C., and Livingstone, D. R. 1998. Variation in levels of cytochrome P4501A, 2B, 2E, 3A and 4A-immunopositive proteins in digestive gland of indigenous and transplanted mussel (*Mytilus galloprovincialis*) in Venice Lagoon, Italy. Marine Environmental Research 46, 295–299.

Ruus, A., Hylland, K., W., and Green, N. 2003. Joint Assessment and Monitoring Programme (JAMP). Biological Effects Methods, Norwegian Monitoring 1997–2001. Norwegian State Pollution Monitoring Programme Report no. 869/03. Ta-no. 1948/2003.

Ruus, A., Sandvik, M., Ugland, K. I., and Skaare, J. U. 2002. Factors influencing activities of biotransformation enzymes, concentrations and compositional patterns of organochlorine contaminants in members of a. marine food web. Aquatic Toxicology 61, 73–87.

Shaw, J. P., Large, A. T., Chipman, J. K., Livingstone, D. R., and Peters, L. D. 2000. Seasonal Variation in mussel (*Mytilus edulis*) digestive gland cytochrome P4501A-and 2E-

immunoidentified protein levels and DNA strand breaks (*Comet assay*). Marine Environmental Research 50, 405–409.

Sleiderink H, Beyer J, Everaarts JM, Boon J.P. 1995. Influence of temperature on cytochrome P450 1A in dab (*Limanda limanda*) from the southern North Sea: Results from field surveys and a laboratory study. Mar. Environ. Res. 39:67–71.

Sole, M., Peters, L. D., Magnusson, K., Sjolin, A., Granmo, A., and Livingstone, D. R. 1998. Responses of the cytochrome P450-dependent monooxygenase and other protective enzyme systems in digestive gland of transplanted common mussel (*Mytilus edulis* L.) to organic contaminants in the Skagerrak and Kattegat (North Sea). Biomarkers 3, 49–62.

Stagg, R., and McIntosh, A. 1998. Biological effects of contaminants: Determination of CYP1A-dependent mono-oxygenase activity in dab by fluorimetric measurement of EROD activity. ICES Techniques in Marine Environmental Science, 23, 16 pp.

Von Westernhagen, H. Rosenthal, H., Dethlefsen, V. Ernst, W., Harms, U. and Hansen, P. D. 1981. Bioaccumulating substances and reproductive success in Baltic flounder (*platichthys flesus*). Aquatic Toxicology 1: 85–99.

Whyte, J. J., Jung, R. E., Schmitt, C. J., and Tillitt, D. E. 2000. Ethoxyresorufin-O-deethylase (EROD) activity in fish as a biomarker of chemical exposure. Critical Reviews in Toxicology 30, 347–570.

Williams, R. T. 1974. Inter-species variations in the metabolism of xenobiotics. Biochem. Soc. T. 2, 359–377.

Williams, T. G., Lockhart, W. L., Metner, D. A., and Harbicht, S. 1997. Baseline studies in the Slave River, NWT, 1990–1994.3. MFO enzyme activity in fish. Science of the Total Environment 197, 87–109.

Aas, E., Baussant, T., Balk, L., Liewenborg, B., and Andersen, O. K. 2000. PAH metabolites in bile, cytochrome P4501A and DNA adducts as environmental risk parameters for chronic oil exposure: a laboratory experiment with Atlantic cod. Aquatic Toxicology 51, 241–258.

Table1. EROD Assessment criteria in fish target species used in biomonitoring programmes around European waters. EROD BRs established are restricted to the sampling conditions and the size length of the specimens used. The values of the assessment criteria must be considered as provisional and should be updated and revised when more data comes available.

EROD Assessment Criteria S9 Fraction	Sampling season	Bottom water temperature range	Size length cm	Sex	Background Response Range EROD activity (pmol/min/mg prot) 90P	Elevated Response Range EROD activity (pmol/min/mg prot) 90P	N
Dab (*Limanda limanda*)	August–November	[10-18 °C]	12–25	Females	≤178	>178	556
				Males	≤147	>147	571
European flounder (*Platichtys flesus*)	August–November	[10-18 °C]	20–25	Females and/or males	≤24	>24	65
Plaice (*Pleuronectes platessa*)	January	[5-10°C]	18.5–22.5	Males	≤10	>10	116
EROD Assessment Criteria Microsomal fraction	Sampling season	Bottom water temperature range	Size length cm	Sex	Background Response Range EROD activity (pmol/min/mg prot) 90P	Elevated Response Range EROD activity (pmol/min/mg prot) 90P	N
Dab (*Limanda limanda*)	August–November	[10-18 °C]	20–30	Females and/or males	≤780	>780	53
Cod (*Gadus morhua*)	August–November	[10-18 °C]	30–45	Females and/or males	≤145	>145	198
Plaice (*Pleuronectes platessa*)	September	[7-10 °C]	40–60	Females and/or males	≤255	>255	64
Four spotted megrim (*Lepidorhombus boscii*)	September–October	[11.7-12.7 °C]	18–22	Females and/or males	≤13	>13	317

EROD Assessment Criteria S9 Fraction	Sampling season	Bottom water temperature range	Size length cm	Sex	Background Response Range EROD activity (pmol/min/mg prot) 90P	Elevated Response Range EROD activity (pmol/min/mg prot) 90P	N
Dragonet (*Callionymus lyra*)	September–October	[12.0–12.8 °C]	15–22	Females and/or males	≤202	>202	159
Red mullet (*Mullus barbatus*)	April	[13.3–15.3 °C]	12–18	Males	≤208	>208	40

Table 2. Description of data used in setting background and elevated response ranges.

EROD Assessment Criteria † sub-fraction S9 * microsomes sub-fraction	Sampling season	Bottom water temperature range	Size length	Sex	EROD BR activity	Upper limit of EROD BR activity	N
					Median (pmol/min/mg prot)	P90 (pmol/min/mg prot)	
		°C	cm				
Dab (*Limanda limanda*)	August–November	[10–18]	12–25	Females and/or males	<30†	<152†	1034
European flounder (*Platichthys flesus*)	August–November	[10–18]	20–25	Females and/or males	<14†	<24†	30
Cod (*Gadus morhua*)	August–November	[10–18]	30–45	Females and/or males	<78*	<151*	74
Four spotted megrim (*Lepidorhombus boscii*)	September–October	[11.7–12.7]	18–22	Females and/or males	<12*	<13*	317
Dragonet (*Callionymus lyra*)	September–October	[12.0–12.8]	15–22	Females and/or males	<144*	<202*	159
Red mullet (*Mullus barbatus*)	April	[13.3–15.3]	12–18	Males	<85*	<208*	40
Plaice (*Pleuronectes platessa*)	January	[18.5–22.5]	7–10	Males	<3.71†	<9.49†	116
Haddock (*Melanogrammus aeglefinus*)	August	[5–10]	33–55	Females and/or males	<72†/<215*	<162† /<421*	20/23
Saithe (*Pollachius virens*)	September	[5–10]	40–100	Females and/or males	<57†	<142†	21
Herring (*Clupea harengus*)	November	[5–10]	22–33	Females and/or males	<10†	<23†	24

Table 3. Dose-response, background response and sensitivity in experimental studies with gadoid fish.

SPECIES	SUBSTANCE(S)	LOWEST-HIGHEST CONCS	EXPOSURE TIME	BASELINE/CONTROL (LEVEL/ACTIVITY)	INDUCTION (FOLD)	REFERENCE
Polar cod *Boreogadus saida* juvenile	Crude oil (Oseberg C)	~200 mg/kg (i.p. inj.)	10 and 21 d post inj.	~30 pmol/min/mg	~8 and ~2.5 (245 and 80 pmol/min/mg)	(George *et al.*, 1995)
Polar cod *Boreogadus saida* Male	Crude oil (Oseberg C)	~200 mg/kg (oral)	21 d post exposure	28 pmol/min/mg ± 6 (n=12)	~5 (132 ± 14 pmol/min/mg)	(George *et al.*, 1995)
Polar cod *Boreogadus saida* female	Crude oil (Oseberg C)	~200 mg/kg (oral)	21 d post exposure	8 pmol/min/mg ± 2 (n=14)	~5 (42 ± 6 pmol/min/mg)	(George *et al.*, 1995)
Polar cod *Boreogadus saida* juvenile	β-naphthoflavone	50 mg/kg (i.p. inj.)	21 d post inj.	~30 pmol/min/mg	~12.5 (380 pmol/min/mg)	(George *et al.*, 1995)
Cod, *Gadus morhua* Juvenile	2,3,7,8-TCDD	0.008 mg/kg oral dose twice, d 0 and d 4	9 and 17 d post exposure	55.4 (d 9) and 91.4 (d 17) pmol/min/mg	~4 and ~3 (230 and 277 pmol/min/mg)	(Hektoen *et al.*, 1994)
Cod, *Gadus morhua* Juvenile	PCB-105	10 mg/kg oral dose twice, d 0 and d 4	measure at d 9 and d 17	55.4 (d 9) and 91.4 (d 17) pmol/min/mg	1.5 and 1.2	(Bernhoft *et al.*, 1994)
Cod, *Gadus morhua* Juvenile	β-naphthoflavone	100 mg/kg (i.p. inj. at d 0 and d 4)	measure at d 7	84 pmol/min/mg ± 8 (n=5)	~13 (1074 ± 340 pmol/min/mg)	(Goksoyr *et al.*, 1987)
Cod, Gadus morhua	β-naphthoflavone	100 mg/kg (2 i.p. inj.)	measure 3-4 d after last injection	40 pmol/min/mg	~72 (2870 pmol/min/mg)	(Goksoyr *et al.*, 1991)
Cod, *Gadus morhua* juvenile	Crude oil (North Sea)	0.06 – 1 ppm	30 days	~2 pmol/min/mg	~ 2- 5.5 (~ 4 – 11 pmol/min/mg)	(Aas *et al.*, 2000)

Table 4. Dose-response, background response and sensitivity in field studies with gadoid fish.

SPECIES	SUBSTANCE(S)	LOWEST-HIGHEST CONCS	EXPOSURE TIME	BASELINE/CONTROL (LEVEL/ACTIVITY)	INDUCTION (FOLD)	REFERENCE
Rockling, *Ciliata mustella*	Crude oil (Gullfaks; M.V. Braer spill, Shetland)	85000 tons spill 129 ± 38 ng /g dry wt. of PAHs (selected 2- and 3-ring) detected in muscle.	3 months after spill	~160 pmol/min/mg ± 50	~9 (1480 pmol/min/mg)	(George *et al.*, 1995)
Roundnose grenadier, *Coryphaenoides rupestris*	i.a. PAHs and PCBs			260 ± 20 (Male) ~170 (Female) pmol/min/mg	~2 (530 ± 70 (male) and ~350 (female) pmol/min/mg)	(Lindesjoo *et al.*, 1996)
Hake, *Urophycis* spp.	Pollution (PAH) from oil platforms (Gulf of Mexico) <100m from platforms			10.9 ± 6.4 and 11.7 ± 10.5 pmol/min/mg (>3000 m from platforms)	<1 (10.6 ± 3.8 and 10.5 ± 7.1 pmol/min/mg)	(McDonald *et al.*, 1996)

Annex 5: Lysosomal stability as a global health status indicator in biomonitoring

Version date: 29 January 2010.

Background

Lysosomal functional integrity is a generic common target for environmental stressors in all eukaryotic organisms from yeast and protozoans to humans (Cuervo, 2004), that is evolutionarily highly conserved, and lysosomal membrane stability is a good diagnostic biomarker of individual health status (Allen and Moore, 2004; Bayne and Moore, 1998; Burlando *et al*., 2002; Cajaraville *et al*., 1995, 2000; Dondero *et al*., 2006b; Galloway *et al*., 2002; 2004; Hankard *et al*., 2004; Klionsky and Emr, 2000; Köhler *et al*., 1992; 2002; Lekube *et al*., 2000; Lowe, 1988; Lowe *et al*., 1982; 1992; 1995; 2006; Marigomez and Baybay-Villacorta, 2003; Moore, 1976; 1985; 1988; 1990; 2002; Moore *et al*., 2004a; Moore *et al*., 2006a,b,c; Nicholson and Lam, 2005; Svendsen and Weeks, 1995; Svendsen *et al*., 2004; Winston *et al*., 2002). Dysfunction of lysosomal processes has been mechanistically linked with many aspects of pathology associated with toxicity and degenerative diseases (Cuervo, 2004; Köhler, 2004; Köhler *et al*., 2002; Moore *et al*., 2006a, b). Recent studies have shown that lysosomal autophagy provides a second line of defence against oxidative stress (Cuervo, 2004; Moore *et al*., 2006c), and the capability to effectively up-regulate this process is probably a significant factor contributing to the ability of some organisms to tolerate stressful and polluted environments.

Lysosomal membrane stability has recently been adopted by UNEP as part of the first tier of techniques for assessing harmful impact in the Mediterranean Pollution programme (MEDPOL Phase IV). Other lysosomal biomarkers including lipofuscin in molluscs (age/stress pigment), and lysosomal neutral lipid (chemically induced lipidosis) in molluscs and fish have been adopted as part of the second tier assessment methods (Krishnakumar *et al*., 1994; Moore, 1988; Moore *et al*., 2004b).

This biomarker can also be used prognostically to predict liver damage and tumour progression in the liver of various fish species (Broeg *et al*., 1999a, b; Köhler *et al*., 2002; Köhler, 2004), and hepatopancreatic degeneration in molluscs (e.g. blue and green mussels, fresh-water bivalves and snails, periwinkles, oysters), coelomocyte damage in earthworms, as well as enhanced protein turnover (i.e. lysosomal autophagy) as a result of radical attack on proteins; and energetic status (i.e. scope for growth) as a predictive indicator of fitness of individuals within a population (Allen and Moore, 2004; Kirchin *et al*., 1992; Köhler *et al*., 2002; Moore *et al*., 2004a; 2006a; Nicholson and Lam, 2005; Svendsen and Weeks, 1995; Svendsen *et al*., 2004).

Lysosomes are known to accumulate many metals and organic xenobiotics. Adverse lysosomal reactions to xenobiotic pollutants include swelling, lipidosis (pathological accumulation of lipid), lipofuscinosis (pathological accumulation of age/stress pigment) in molluscs but not fish, and loss of membrane integrity (Köhler *et al*., 2002; Moore, 1988; Moore *et al*., 2006a, b; Viarengo *et al*., 1985a). Metals such as copper, cadmium and mercury will also induce lysosomal destabilization in mussels (Viarengo *et al*., 1981; 1985a, b), and if oxyradicals are generated then lipofuscinosis can also occur (1985b).

Lysosomal membrane integrity or stability in blue mussels is correlated with oxygen and nitrogen radical scavenging capacity (TOSC), protein synthesis, scope for growth

and larval viability (oysters *Crassostrea gigas*); and inversely correlated with DNA damage (incidence of micronuclei), lysosomal swelling, lipidosis and lipofuscinosis, which are characteristic of failed or incomplete autophagy (Dailianis *et al*., 2003; Kalpaxis *et al*., 2004; Krishnakumar *et al*., 1994; Moore *et al*., 2004a, b; 2006a; Regoli, 2000; Ringwood *et al*., 2004). In fish liver, lysosomal membrane stability is strongly correlated with a suppression of the activity of macrophage aggregates (Broeg, 2003; Broeg *et al*., 2005), lipidosis and lipofuscinosis (Broeg *et al*., 1999 a,b; Broeg *et al*., in preparation; Köhler, 2004).

Lysosomal stability and other lysosomal biomarkers such as lipofuscin are strongly correlated with mussel tissue concentration of PAHs, which are ubiquitous contaminants (Cajaraville *et al*., 2000; Krishnakumar *et al*., 1994; Moore, 1990; Moore *et al*., 2006a, b, c; Viarengo *et al*., 1992), as well as organochlorines and PCB congeners in liver of fish (Köhler *et al*., 2002). Lysosomal stability of various species of mussel and fish from different climate zones clearly reflects gradients of complex mixtures of chemicals in water and sediments (da Ros *et al*., 2002; Pisoni *et al*, 2004; Schiedek *et al*., 2006, Barsiene *et al*., 2006; Sturve *et al*., 2005), single pollution events and accidents (Einsporn *et al*., 2005; Broeg *et al*., 2002, Nicholson and Lam, 2005) and also serves for the discovery of new "Hot Spots" of pollution (Bressling, 2006; Moore *et al*., 1997, 1998a,b; 2004).

A conceptual mechanistic model has been developed linking lysosomal damage and autophagic dysfunction with injury to cells, tissues and the whole animal; and the complementary use of cell-based bioenergetic computational model of molluscan hepatopancreatic cells that simulates lysosomal and cellular reactions to pollutants has also been demonstrated (Allen and McVeigh, 2004; McVeigh *et al*., 2006; Lowe, 1988; Moore *et al*., 2006a, b, c). The integration of biomarker data can be achieved using multivariate statistics and then mapped onto a two dimensional representation of "health status space" (see below) by using lysosomal membrane stability as a measure of cellular well-being (Allen and Moore, 2004; Clarke, 1999; Dagnino *et al*., 2007; Dondero *et al*., 2006a; Lowe, 1988; Moore, 1988; Moore *et al*., 2006a). This is viewed as a crucial step towards the derivation of explanatory frameworks for prediction of pollutant impact on animal health.

Health status space is analogue to phase space in physics. For a system of n first-order ordinary differential equations, the 2n-dimensional space consisting of the possible values of x is known as its phase space. In its simplest form it is a two dimensional graph where any point can be described in terms of two numbers the x and y coordinates. The dimensions of multidimensional health status space are multiple contaminant and biomarker data, environmental variability, space and time. Principal component analysis (PCA) has been used to reduce the dimensionality of the problem to a simple two-dimensional representation (Allen and Moore, 2004; Lowe *et al*., 2006; Moore *et al*., 2006a).

Confounding factors

Lysosomal stability is an indicator of health status and will be affected by non-contaminant factors such as severe nutritional deprivation, severe hyperthermia and prolonged hypoxia (Moore *et al*., 1980; Moore *et al*., 2007). Processing for neutral red retention (NRR) in samples of molluscs adapted to low salinity environments should use either physiological saline adjusted to the equivalent ionic strength or else use ambient filtered seawater. The major confounding factor in respect of biomonitoring is the adverse effect of the final stage of gametogenesis and spawning in mussel,

which is a naturally stressful process (Bayne *et al*., 1978). In general, this period should be avoided any way for sampling purposes, as most physiological processes and related biomarkers are adversely affected (Moore *et al*., 2004b). However, for fish, spawning has only a minimal effect on lysosomal stability and does not mask harmful chemical induced damage to lysosomal membrane stability (Köhler, 1991).

Ecological relevance

Lysosomal integrity is directly correlated with physiological scope for growth (SFG) and is also mechanistically linked in terms of the processes of protein turnover (Allen and Moore, 2004; Moore *et al*., 2006a), and Ringwood *et al*. (2004) have also shown that lysosomal stability in parent oysters is directly correlated with larval viability. Finally, lysosomal stability is also directly correlated with diversity of macrobenthic organisms in an investigation in Langesund Fjord in Norway (Moore *et al*., 2006b).

Quality assurance

Intercalibration exercises for lysosomal stability techniques have been carried out in the ICES/UNESCO-IOC-GEEP Bremerhaven Research Workshop and UNEP-MEDPOL programme, and for the neutral red retention method in the GEF Black Sea Environment Programme (Köhler *et al*., 1992; Lowe *et al*., 1992; Moore *et al*., 1997; 1998a, b; Viarengo *et al*., 2000). The results from these operations indicated that both techniques could be used in the participating laboratories in an effective manner with insignificant inter-laboratory variability.

The standards used in this intercalibration involved digestive glands from marine mussels prepared at the University of Genoa/University of Eastern Piedmont, Alessandria (Italy). Comparisons of the cytochemical and the neutral red retention techniques have been performed in fish liver (ICES-IOC Bremerhaven Workshop, 1990) and in mussels experimentally exposed to PAHs (Lowe *et al*., 1995).

Background responses and assessment criteria

Health status thresholds for NRR and cytochemical methods for lysosomal stability have been determined from data based on numerous studies (Cajaraville *et al*., 2000; Moore *et al*., 2006a).

Lysosomal stability is a biophysical property of the bounding membrane of secondary lysosomes and appears to be largely independent of taxa. In all organisms tested to date, which includes protozoans, annelids (terrestrial and marine), molluscs (freshwater and marine), crustaceans (terrestrial and aquatic), echinoderms and fish, the absolute values for measurement of lysosomal stability (NRR and cytochemical method) are directly comparable. Furthermore, measurements of this biomarker in animals from climatically and physically diverse terrestrial and aquatic ecosystems also indicate that it is potentially a universal indicator of health status. For the cytochemical method animals are considered to be healthy if the lysosomal stability is $\geq$20 minutes; stressed but compensating if <20 but $\geq$10 minutes and severely stressed and probably exhibiting pathology if <10 minutes (Moore *et al*., 2006a). Similarly for the NRR method, animals are considered to be healthy if NRR is $\geq$120 minutes; stressed but compensating if <120 but $\geq$50 minutes and severely stressed and probably exhibiting pathology if <50 minutes (Moore *et al*., 2006a).

References

Allen, J.I. and Moore, M.N. 2004. Environmental prognostics: is the current use of biomarkers appropriate for environmental risk evaluation. Mar. Environ. Res., 58: 227–232.

Allen, J.I. and McVeigh, A. 2004. Towards computational models of cells for environmental toxicology. J. Mol. Histol., 35: 697–706.

Bayne, C.J. and Moore, M.N. 1998. Non-lymphoid immunologic defences in aquatic invertebrates and their value as indicators of aquatic pollution. In: EcoToxicology: Responses, Biomarkers and Risk Assessment (Ed. J.T. Zelikoff), pp. 243–261. Published for OECD by SOS Publications, Fair Haven, New Jersey.

Barsiene, J., Lehtonen, K. K., Koehler, A., Broeg, K., Vuorinen, P. J., Lang, T., Pempkowiak, J., Balk, L., Syvokiene, J., Dedonyte, V., Rybakovas, A., Repecka, R., Vuontisjärvi, H. and Kopecka, J. 2006. Biomarker responses in mussel (*Mytilus edulis*) and flounder (*Platichthys flesus*) in the Klaipeda-Butinge area (Baltic Sea). Mar. Poll. Bull., 53: 422–436.

Bressling J. 2006. Natural and toxic stress indicator in blue mussel *Mytilus edulis* along an offshore-inshore gradient (Weser estuary). Master's Thesis, University of Oldenburg, pp122.

Broeg, K., Köhler, A. and von Westernhagen, H. 2002. Disorder and recovery of environmental health monitoring by means of lysosomal stability in liver of European flounder (*Platichthys flexus* L.). Mar. Environ. Res., 54: 569–573.

Broeg, K., Zander, S., Diamant, A., Körting, W., Krüner, G., Paperna, I. and von Westernhagen, H. 1999a. The use of fish metabolic, pathological and parasitological indices in pollution monitoring. I. North Sea. Helgoland Mar. Res., 53: 171–194.

Broeg, K.,Zander, S.,Diamant, A.,Körting, W.,Krüner, G.,Paperna, I., and von Westernhagen, H. 1999b. The use of fish metabolic, pathological and parasitological indices in pollution monitoring I. Red Sea and Mediterranean Sea. Helgoland Mar. Res., 53: 195–208.

Broeg, K., Kaiser, W., Bahns, S. and Köhler, A. The liver of wrasse - morphology and function as a mirror of point source chemical impact. In preparation.

Burlando, B., Evangelisti, V., Dondero, F., Pons, G., Camakaris, J. and Viarengo, A. 2002. Occurrence of Cu-ATPase in Dictyostelium: possible role in resistance to copper. Biochem. Biophys. Res. Comm., 291: 476–483.

Cajaraville, M.P., Bebianno, M.J., Blasco, J., Porte, C., Sarasquete, C. and Viarengo, A. 2000. The use of biomarkers to assess the impact of pollution in coastal environments of the Iberian Peninsula: a practical approach. Sci. Tot. Environ., 247: 295–311.

Cajaraville, M.P., Abascal, I., Etxeberria, M. and Marigómez, I. 1995. Lysosomes as cellular markers of environmental pollution: time- and dose-dependent responses of the digestive lysosomal system of mussels after petroleum hydrocarbon exposure. Environ. Toxicol. Water Qual., 10: 1–8.

Clarke, K.R. 1999. Non-metric multivariate analysis in community-level ecotoxicology. Environ. Toxicol. Chem., 18: 117–127.

Cuervo, A.M. 2004. Autophagy: in sickness and in health. TRENDS in Cell Biology, 14, 70–77.

Dagnino A., Allen, J.I., Moore M., Canesi L. and Viarengo A. 2007. Integration of biomarker data into an organism health index: development of an expert system and its validation with field and laboratory data in mussels. Biomarkers, in press.

Dailianis, S., Domouhtsidou, G.P., Raftopoulou, E. Kaloyianni, M. and Dimitriadis, V.K. 2003. Evaluation of neutral red retention assay, micronucleus test, acetylcholinesterase activity and a signal transduction molecule (cAMP) in tissues of *Mytilus galloprovincialis* (L.), in pollution monitoring. Mar. Environ. Res,. 56: 443–470.

Da Ros, L., Meneghetti, F. and Nasci, C. 2002. Field application of lysosomal destabilisation indices in the mussel *Mytilus galloprovincialis*: biomonitoring and transplantation in the Lagoon of Venice (northeast Italy), Mar. Environ. Res., 54: 817–822.

Dondero, F., Dagnino, A., Jonsson, H., Capri, F., Gastaldi, L. and Viarengo, A. 2006a. Assessing the occurrence of a stress syndrome in mussels (*Mytilus edulis*) using a combined biomarker/gene expression approach. Aquatic Toxicol., 78S, S13–S24.

Dondero, F., Jonsson, H., Rebelo, M., Pesce, G., Berti, E., Pons, G. and Viarengo, A. 2006b. Cellular responses to environmental contaminants in amoebic cells of the slime mould *Dictyostelium discoideum*. Comp. Biochem. Physiol., 143C: 150–157.

Einsporn, S., Broeg, K. and Köhler, A. 2005. The Elbe flood 2002-toxic effects of transported contaminants in flatfish and mussel of the Wadden Sea. Mar. Pollut. Bull., 50: 423–429.

Galloway, T.S., Sanger, R.C., Smith, K.L., Fillmann, G., Readman, J.W., Ford, T.E. and Depledge, M.H. 2002. Rapid assessment of marine pollution using multiple biomarkers and chemical immunoassays. Environ.Sci.Technol., 2002, 36: 2219–2226.

Galloway, T.S., Brown, R.J., Browne, M.A., Dissanayake, A., Lowe, D., Jones, M.B. and Depledge, M.H. 2004. A multibiomarker approach to environmental assessment. Environ. Sci. Technol., 38: 1723–1731.

Hankard, P.K., Svendsen, C., Wright, J., Wienberg, C., Fishwick, S.K., Spurgeon, D.J. and Weeks, J.M. 2004. Biological assessment of contaminated land using earthworm biomarkers in support of chemical analysis. Sci. Total Environ. 330: 9–20.

Kalpaxis, D.L., Theos, C., Xaplanteri, M.A., Dinos, G.P., Catsiki, A.V. and Leotsinidis, M. 2004. Biomonitoring of Gulf of Patras, N. Peloponnesus, Greece. Application of a biomarker suite including evaluation of translation efficiency in *Mytilus galloprovincialis* cells. Environ. Res., 94: 211–220.

Kirchin, M.A., Moore, M.N., Dean, R.T. and Winston, G.W. 1992. The role of oxyradicals in intracellular proteolysis and toxicity in mussels. Mar. Environ. Res., 34: 315–320.

Klionsky, D.J. and Emr, S.D. 2000. Autophagy as a regulated pathway of cellular degradation. Science, 290: 1717–1721.

Köhler, A. 1991. Lysosomal perturbations in fish liver as indicator for toxic effects of environmental pollution. Comp. Biochem. Physiol., 100C: 123–127.

Köhler, A. 1990. Identification of contaminant-induced cellular and subcellular lesions in the liver of flounder (*Platichthys flesus* L.) caught at differently polluted estuaries. Aquat. Toxicol. 16: 271–294.

Köhler, A. and Pluta, H.J. 1995. Lysosomal injury and MFO activity in the liver of flounder (*Platichthys flesus* L.) in relation to histopathology of hepatic degeneration and carcinogenesis, Mar. Environ. Res., 39: 255–260.

Köhler, A., Alpermann, T., Lauritzen, B. and van Noorden, C.J.F. 2004b. Clonal xenobiotic resistance during pollution-induced toxic injury and hepatocellular carcinogenesis in liver of female flounder. Acta Histochem., 106: 155–170.

Köhler, A., Deisemann, H. and Lauritzen, B. 1992. Ultrastructural and cytochemical indices of toxic injury in dab liver. Mar. Ecol. Prog. Ser., 91: 141–153.

Köhler, A., Wahl, E. and Söffker, K. 2002. Functional and morphological changes of lysosomes as prognostic biomarkers of toxic liver injury in a marine flatfish (*Platichthys flesus* (L)). Environ. Toxicol. Chem., 21: 2434–2444.

Krishnakumar, P.K., Casillas, E. and Varanasi, U. 1994. Effect of environmental contaminants on the health of *Mytilus edulis* from Puget Sound, Washington, USA. I. Cytochemical measures of lysosomal responses in the digestive cells using automatic image analysis. Mar. Ecol. Prog. Ser., 106: 249–261.

Lehtonen, K., Schiedek, D., Koehler, A., Lang, T., Vuorrinen, P.J., Foerlin, L., Barsiene, J., Pempkowski, aJ., Gercken, J. 2006. The Beep Project in the Baltic Sea: Overview of results and outline for a regional biological effects monitoring strategy. Mar. Poll. Bull., 53: 523–537.

Lekube, X., Cajaraville, M.P. and Marigomez, I. 2000. Use of polyclonal antibodies for the detection of changes induced by cadmium in lysosomes of aquatic organisms. Sci. Total Environ., 247: 201–212.

Lowe, D.M. 1988. Alterations in the cellular structure of *Mytilus edulis* resulting from exposure to environmental contaminants under field and experimental conditions. Mar. Ecol. Prog. Ser., 46: 91–100.

Lowe, D. M., Fossato, V. U. and Depledge, M. H. 1995. Contaminant induced lysosomal membrane damage in blood cells of mussels *M. galloprovincialis* from the Venice Lagoon: an in vitro study. Mar. Ecol. Prog. Ser., 129: 189–196.

Lowe, D.M., Moore, M.N. and Clarke, K.R. 1981. Effects of oil on digestive cells in mussels: quantitative alterations in cellular and lysosomal structure. Aquatic Toxicol., 1: 213–226.

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

Lowe, D.M., Moore, M.N. and Readman, J.W. 2006. Pathological reactions and recovery of hepatopancreatic digestive cells from the marine snail *Littorina littorea* following exposure to a polycyclic aromatic hydrocarbon. Mar. Environ. Res., 61: 457–470.

McVeigh, A., Allen, J.I., Moore, M.N., Dyke, P. and Noble, D. 2004. A carbon and nitrogen flux model of mussel digestive gland epithelial cells and their simulated response to pollutants. Mar. Environ. Res., 58: 821–827.

Marigomez, I. and Baybay-Villacorta, L. 2003. Pollutant-specific and general lysosomal responses in digestive cells of mussels exposed to model organic chemicals. Aquatic Toxicol., 64: 235–257.

Moore, M.N. 1976. Cytochemical demonstration of latency of lysosomal hydrolases in digestive cells of the common mussel, *Mytilus edulis* and changes induced by thermal stress. Cell Tissue Res., 175: 279–287.

Moore, M.N. 1985. Cellular responses to pollutants. Mar. Poll. Bull., 16: 134–139.

Moore, M.N. 1988. Cytochemical responses of the lysosomal system and NADPH-ferrihemoprotein reductase in molluscs to environmental and experimental exposure to xenobiotics. Mar. Ecol. Prog. Ser., 46: 81–89.

Moore, M.N. 1990. Lysosomal cytochemistry in marine environmental monitoring. Histochem. J., 22: 187–191.

Moore, M.N. 2002. Biocomplexity: the post-genome challenge in ecotoxicology. Aquatic Toxicol., 59: 1–15.

Moore, M.N., Allen, J.I. and McVeigh, A. 2006a. Environmental prognostics: an integrated model supporting lysosomal stress responses as predictive biomarkers of animal health status. Mar. Environ. Res. In press.

Moore, M.N., Allen, J.I., McVeigh, A. and Shaw, J. 2006c. Lysosomal and autophagic reactions as predictive indicators of environmental impact in aquatic animals. Autophagy, 2: 217–220.

Moore, M.N., Allen, J.I. and Somerfield, P.J. 2006b. Autophagy: role in surviving environmental stress. Marine Environmental Research, 62 Suppl. 1: S420–S425.

Moore, M.N., Depledge, M.H., Readman, J.W. and Leonard, P. 2004a. An integrated biomarker-based strategy for ecotoxicological evaluation of risk in environmental management. Mutation Res., 552: 247–268.

Moore, M.N., Koehn R.K. and Bayne, B.L. 1980. Leucine aminopeptidase (aminopeptidase-1), N-acetyl-β-hexosaminidase and lysosomes in the mussel *Mytilus edulis* L., in response to salinity changes. J. Exp. Zool., 214: 239–249.

Moore, M.N., Lowe, D. and Köhler, A. 2004b. Measuring lysosomal membrane stability. ICES Techniques in Marine Environmental Sciences, No. 36. ICES, Copenhagen, 31 p.

Moore, M. N., Lowe, D. M., Wedderburn, R. J., Wade, T., Balashov. G., Büyükgüngör, H., Daurova, Y., Denga, Y., Kostylev, E., Mihnea, P., Moncheva, S.Tabagari, S., Ciocan, C., Özkoc, H., Depledge, M. H. 1998a. International Mussel Watch (UNESCO/IOC) in the Black Sea: a pilot study for biological effects and contaminant residues. In: Environmental degradation of the Black Sea, (Eds. S. Beşiktepe *et al*.), pp. 273–289. Kluwer Academic Publishers, Amsterdam.

Moore, M.N., Topping, G., Minhea, P.-E., Kiryanov, S. and Mikhailov, V. 1998b. Design of a monitoring programme for the Black Sea: contaminant levels and biological effects. In: Black Sea Pollution Assessment (Eds. Mee, L. and Topping,G.), 293–299. UN Publications, New York.

Moore, M.N., Viarengo, A., Donkin, P. and Hawkins, A.J.S. 2007. Autophagic and lysosomal reactions to stress in the hepatopancreas of blue mussels. Aquatic Toxicology, In press.

Moore, M. N., Wade, T. L., Lowe, D. M., Wedderburn, R. J., Balashov, G., Büyükgüngör, H., Daurova, Y., Denga, Y., Kostylev, E., Mihnea, P., Moncheva, S., Tabagari, S., Ciocan, C., Özkoc, H. and Depledge, M. H. 1997. The UNESCO/IOC "Black Sea Mussel Watch Pilot Study": biological effects and contaminant residues. In: Black Sea Pollution Assessment (Eds. Mee, L. and Topping,G.), pp. 279–292. UN Publications, New York.

Nicholson, S. and Lam, P.K.S. 2005. Pollution monitoring in Southeast Asia using biomarkers in the mytilid mussel *Perna viridis* (Mytilidae: Bivalvia). Environ. Internat., 31: 121–132.

Pisoni, M., Cogotzi, L., Frigeri, A., Corsi, I., Bonacci, S., Iacocca, A., Lancini, L., Mastrototaro, F., Focardi, S., Svelto *et al*. 2004. DNA adducts, benzo(a)pyrene monooxygenase activity, and lysosomal membrane stability in *Mytilus galloprovincialis* from different areas in Taranto coastal waters (Italy). Environ. Res., 96: 163–175.

Regoli, F. 2000. Total; oxyradical scavenging capacity (TOSC) in polluted and translocated mussels: a predictive biomarker of oxidative stress. Aquatic Toxicol.: 50, 351–361.

Ringwood, A.H., Hoguet, J., Keppler, C. and Gielazyn, M. 2004. Linkages between cellular biomarker responses and reproductive success in oysters, *Crassostrea virginica*. Mar Environ. Res., 58: 151–155.

Schiedek, D., Broeg, K., Barsiene, J., Lehtonen, K. K., Gercken, J., Pfeifer, S., Vuontisjärvi, H., Vuorinen, P. J., Köhler, A., Balk, L., Schneider, R. 2006. Biomarker responses and indication of contaminant effects in blue mussel (*Mytilus edulis*) and eelpout (*Zoarces viviparus*) from the western Baltic Sea. Mar. Poll. Bull., 53: 387–405.

Sturve, J., Balk, L., Berglund, A., Broeg, K., Böhmert, B., Koehler, A., Massey, S., Parkkonen, J., Savva, D., Stephensen, E., Förlin, L. 2005. Effects of dredging in Goteborg harbour assessed by biomarkers in eelpout (*Zoarces viviparus*). Environ. Toxicol. Chem., 24: 1951–1961.

Svendsen, C. and Weeks, J. M. 1995. The use of a lysosome assay for the rapid assessment of cellular stress from copper to the freshwater snail *Viviparus contectus* (Millet). Mar. Poll. Bull. : 31, 139–42.

Svendsen, C., Spurgeon, D.J., Hankard, P.K. and Weeks, J.M. 2004. A review of lysosomal membrane stability measured by neutral red retention: is it a workable earthworm biomarker? Ecotox. Environ. Safety, 57: 20–29.

Viarengo, A., Lafaurie, M., Gabrielides, G.P., Fabbri, R., Marro, A. and Romeo, M. 2000. Critical evaluation of an intercalibration exercise undertaken in the framework of the MEDPOL biomonitoring program. Mar. Environ. Res., 49: 1–18.

Viarengo, A., Moore, M.N., Mancinelli, G., Mazzucotelli, A. and Pipe, R.K. 1985a. Significance of metallothioneins and lysosomes in cadmium toxicity and homeostasis in the digestive gland cells of mussels exposed to the metal in the presence and absence of phenanthrene. Mar. Environ. Res., 17, 184–187.

Viarengo, A., Moore, M.N., Pertica, M., Mancinelli, G., Accomando, R. 1992. A simple procedure for evaluating the protein degradation rate in mussel (*Mytilus galloprovincialis* Lam.) tissues and its application in a study of phenanthrene effects on protein catabolism. Comp. Biochem. and Physiol., 103B: 27–32.

Viarengo, A., Moore, M.N., Pertica, M., Mancinelli, G., Zanicchi, G. and Pipe, R.K. 1985b. Detoxification of copper in the cells of the digestive gland of mussel: the role of lysosomes and thioneins. Sci. Tot. Environ., 44: 135–145.

Viarengo, A., Zanicchi, G., Moore, M.N. and Orunesu, M. 1981. Accumulation and detoxication of copper by the mussel *Mytilus galloprovincialis*. Aquatic Toxicol., 1: 147–157.

Winston, G.W., Adams, S.M., Benson, W.H., Gray, L.E., Matthews, H.S., Moore, M.N., and Safe, S. 2002. Biological bases of similarities and differences. In: "Interconnections between Human Health and Ecological Integrity", (Eds. R.T. Di Giulio and W.B. Benson), pp. 43–65. Society of Environmental Toxicology and Chemistry (SETAC), Pensacola, Florida.

Annex 6: DNA adducts

Version date: 29 January 2010.

Background

In the chemical carcinogenesis model the initiating step is the covalent modification of DNA by a carcinogen (Miller and Miller, 1981). The measurement of covalent structures formed between genotoxic compounds and DNA, termed DNA adducts, can be utilized as a biological marker of exposure to genotoxic contaminants. DNA adducts can be removed by cellular repair processes or by cell death, but during chronic exposures they often reach steady state concentrations. DNA adducts have several important features which make them suitable as biomarkers of genotoxic exposure:

a) It is a quantifiable measurement of the biologically effective dose of a contaminant reaching a critical cellular target and therefore a useful epidemiological biomarker for detecting exposure to environmental genotoxins.

b) DNA adduct levels integrate multiple toxicokinetic factors such as uptake, metabolism, detoxification, excretion and DNA repair in target tissues.

c) DNA adducts are relatively persistent once formed (may last several months) and therefore they provide an assessment of chronic exposure accumulated over many weeks rather than a few days, as afforded by other PAH biomarkers such as EROD induction or the presence of bile metabolites.

d) Studies from North America have shown that risk factors for certain lesions can be generated by correlating the level of DNA damage with lesion occurrence, thus allowing the use of a relatively simple biomarker in predicting risk.

Polycyclic aromatic hydrocarbons (PAHs) are a ubiquitous and large group of environmental contaminants, some of which are known to cause genetic toxicity through the formation of DNA adducts. Over the past 25 years a growing body of research has investigated the uptake, bioaccumulation and metabolism of PAHs and there is now extensive experimental and field based evidence supporting their role in the initiation and progression of chemical carcinogenesis. Numerous field studies in both North America and Europe have established a correlation between PAH sediment concentrations and the prevalence of hepatic tumours in fish (Malins *et al*., 1985; Myers *et al*., 1991; Baumann, 1998). For example, liver and skin neoplasia in brown bullheads (*Ictaluvus nebulosus*) from the Black River, Ohio (USA) have been shown to be strongly correlated with PAH sediment contamination (Baumann, 1998). Further work carried out in Puget Sound (USA) has also found positive correlations between hepatic lesions including neoplasia (hepatocellular carcinomas and cholangiocellular carcinomas) and foci of cellular alteration (pre-neoplastic lesions) in English sole (*Parophrys vetulus*) and sediment PAH contamination (Malins *et al*., 1985). Therefore, the measurement of DNA adduct levels in marine organisms is an important step in assessing risk from exposure to environmental carcinogens and mutagens.

Of the techniques currently available for the detection of DNA adducts the most sensitive method for the detection of a wide range of compounds chemically bound to DNA is the ^{32}P-postlabelling assay (Gupta *et al*., 1982). The method possesses a number of advantages that make it suitable for the assessment of DNA adduct induced by

environmental genotoxins (for a review see Beach and Gupta, 1992; Phillips, 1997; Phillips, 2005). The technique is applicable to any tissue sample from which DNA can be isolated and is also extremely sensitive, capable of detecting one adducted nucleotide in 10^9–10^{10} undamaged nucleotides from 5–10 μg DNA. In addition, providing the adduct is amenable to the labelling reaction and subsequent thin layer chromatography, its prior characterization is not required. It is this last feature that makes the assay particularly appropriate to aquatic biomonitoring, because it is suitable for the analysis of the diverse array of adducts induced by complex mixtures of environmental chemicals. It is important to note that ^{32}P-postlabelling is only semi-quantitative as not all DNA adducts are labelled with the same efficiency and the various enrichment and chromatograph steps involved will preferentially select certain adducts. However, the assays sensitivity, coupled with the assays ability to detect a wide range of genotoxic compounds (e.g. PAHs), has led to its wide spread use in environmental biomonitoring programmes using both vertebrate and invertebrate sentinel organisms (van der Oost *et al*., 1994; Ericson *et al*., 1998; Lyons *et al*., 1999; Akcha *et al*., 2004; Lyons *et al*., 2004b; Balk *et al*., 2006), following exposure to specific environmental genotoxins (Ericson *et al*., 1999; Lyons *et al*., 1999) and to compounds present in organic extracts from PAH contaminated sediments (Stein *et al*., 1990; French *et al*., 1996).

Ecological relevance and validation for use in the field

The field validation of a biomarker of exposure, such as DNA adducts is essential in establishing their credentials when used in routine monitoring programmes. In North America the technique has been widely used (>30 marine and fresh-water species) and guidelines for implementation are published in an ICES Times technical document (Reichert *et al*., 1999). Across the OSPAR maritime area the assay has been used in several biological effects monitoring programmes using a range of indicator species including blue mussels, *Mytilus* sp, perch (*Perca fluviatilis*), dab (*Limanda limanda*), European flounder (*Platichthys flesus*), eelpout (*Zoarces viviparous*) and cod (*Gadus morhua*) (Ericson *et al*., 1998; Lyons *et al*., 1999; Lyons *et al*., 2000; Ericson *et al*., 2002; Aas *et al*., 2003; Akcha *et al*., 2004; Lyons *et al*., 2004a,b Balk *et al*., 2006). Studies from both North America and Europe have clearly demonstrated that when using non-migratory fish the levels of DNA adducts strongly correlate with the concentration of PAH sediment contamination (van der Oost *et al*., 1994; Ericson *et al*., 1999; Lyons *et al*., 1999). For example, studies using the eel (*Anguilla anguilla*) demonstrated a significant relationship between the level of DNA adducts and PAH contamination of the sediment (Van der Oost *et al*., 1994). Laboratory studies have demonstrated that fish exposed to PAHs accumulate hepatic DNA adducts in both a time- and a dose-dependent manner (French *et al*., 1996). It is known from experimental studies using both fish and shellfish that such DNA adducts may persist for many months once formed and are therefore particularly suited to monitoring chronic exposure to genotoxic contaminants (Stein *et al*., 1990; French *et al*., 1996; Harvey and Parry, 1998). Significantly, field based studies have investigated the relationship between DNA adduct formation and neoplastic liver disease and it has been shown that at certain contaminated sites the prevalence of DNA adducts are associated with the prevalence of toxicopathetic lesions including foci of cellular alteration and neoplasia (for review see Reichert *et al*., 1998).

Studies from North America and Europe suggest that DNA adduct levels in fish are not markedly influenced by factors such as age, sex, season or dietary status, which are known to confound the interpretation of other biomarkers (e.g. EROD). However,

validation of any biomarker, including DNA adducts in a species of interest is essential to ensure against any unforeseen species-specific responses (Reichert *et al*., 1999). While there is no evidence to suggest that environmental factors such as salinity and temperature significantly affect the formation of DNA adducts these factors should always be considered, as it is known that cellular detoxification systems (e.g. Cyp1A) are influenced by changes in environmental variables (Sleiderink *et al*., 1995).

Species selection and target tissue

The majority of hydrophobic genotoxins, such as PAHs, released into the marine environment quickly adhere or organic particular matter and settle into the sediment. Therefore, the majority of fish species used in PAH contaminant-monitoring programmes are benthic feeders, such as the marine flatfish. A particular advantage of the ^{32}P-postblabelling assay is that it is not species-specific and therefore can be utilized on any organism deemed fit for purpose. As such it has been used widely in a range of species (both vertebrate and invertebrate), ranging from filter-feeders to high-order predators. It should be noted that DNA adducts are known to accumulate and persist over time (Stein *et al*., 1990; French *et al*., 1996) and consequently should be considered a cumulative index integrating both past and present genotoxic exposure. Therefore, care needs to be taken when undertaking studies in migratory fish species as the detectable levels of DNA adducts may not be a true representation of the genotoxic contaminants at the site of capture. It has been suggested by Reichert *et al*, (1999) that in such situations biomarkers, such as bile metabolite analysis, should be employed in parallel as this would provide a relatively accurate index of recent PAH exposure and would therefore indicate whether the levels of DNA adducts were due to exposure at the site of capture.

Of the affected organs, liver is the most commonly studied when fish are used as sentinel organisms. Field data infers a chemical aetiology for many of the commonly observed hepatic lesions seen in wild fish collected from contaminated areas. Laboratory data supporting this association stems from biochemical and molecular studies which have shown the liver to be the major site for contaminant detoxification pathways (e.g. cytochrome P-450-mediated biotransformation enzyme systems). Furthermore, contaminant metabolism studies have shown fish liver microsomes are capable of producing the ultimate carcinogenic forms of common environmentally relevant PAHs, including benzo[a]pyrene, which bind to DNA to form adducts (Sikka *et al*., 1991). As mentioned previously, a major strength of the ^{32}P-postabelling assay is that it is not tissue specific and assuming sufficient DNA can be extracted it can be applied in a fit-for-purpose manner in any tissue of choice. To this end it has been used successfully in a range of tissues (both invertebrate and vertebrate), including liver, intestine, gill, brain, gonad and digestive gland (Ericson *et al*., 1999; French *et al*., 1996; Lyons *et al*., 1997; Harvey and Parry, 1998).

Methodology and technical considerations

^{32}P–postlabelling

In the ^{32}P-postlabeling method, DNA isolated from tissue is first hydrolysed enzymatically to 3′-monophosphates. The proportion of adducts in the enzyme hydrolysate are enriched by selective removal of unmodified nucleotides by enzymatic methods (Reddy and Randerath, 1986) or by extracting the adducts into n-butanol (Gupta, 1985) before labelling the mononucleotides with ^{32}P-ATP. For hydrophobic aromatic DNA adducts, such as PAH-DNA adducts, the enrichment steps can en-

hance the sensitivity of the assay to detect 1 adduct in 10^9–10^{10} bases (Reichert *et al*., 1999). Following the adduct enrichment step, the 3′-monophosphates are radiolabelled at the 5′-hydroxyl using ^{32}P-ATP and T4-polynucleotide kinase to form 3′, (^{32}P)5′-bisphosphates. Separation of the ^{32}P-labeled adducts is accomplished by multidimensional high-resolution anion exchange thin-layer chromatography. Autoradiography is then used to locate the radio-labelled adducts on the chromatogram and the radioactivity is measured by either liquid scintillation spectroscopy or storage phosphor imaging (IARC, 1993, Phillips and Castegnaro, 1999). Detailed methodologies have which have been through appropriate Quality Assurance (QA) programmes are now published by ICES and IARC (Phillips and Castegnaro, 1999; Reichert *et al*., 1999).

Radiation safety

The ^{32}P-postlabelling assay uses large amounts of ^{32}P, which is an energetic beta emitter (1.7 MeV) with a half-life of 14.3 days. Researchers using this isotope must receive detailed instruction before handling ^{32}P and must be frequently monitored for exposure to ^{32}P. In the UK the use of ^{32}P in scientific procedures is governed by Environment Agency. Institutes need to have an appointed Radiation Protection Supervisor (RPS) and follow designated licence consent criteria. Institutes wishing to conduct ^{32}P-postlabelling outside the UK must contact their own national licensing organization to clarify the legislative procedures required.

Main considerations to help minimize and monitor ^{32}P exposure:

- All researchers who handle ^{32}P must wear a whole body film badge and a finger dosimeter on the inside of each hand where there is the highest potential for radiation exposure. These badges should be monitored regularly.
- All laboratory operations are planned to minimize the time spent handling radioactivity, the use of tongs and forceps to minimize handling of tubes and vials is recommended.
- Double latex gloves are worn while handling ^{32}P and they should be regularly checked for radioactivity by passing them under a radiation monitor. Gloves should immediately be changed and discarded if contaminated.
- Laboratory working surfaces are checked frequently with the radiation monitor when handling ^{32}P. The monitor probe should be covered with a thin vinyl wrap to prevent contamination of the detector.
- After completion of work with radioactivity, the workers are to check themselves and their equipment with the radiation monitor. If any radioactivity is detected then they are to wash themselves and/or the equipment until free of radioactivity.

Equipment for handling and storage of ^{32}P

All ^{32}P is handled behind 1 cm Perspex/Plexiglas shielding. In addition, samples are kept in Perspex/Plexiglas containers that are at least 1 cm thick. Where possible all manipulations of eppendorfs and vials should be conducted using long armed tongs. It is recommended that radioactive waste is temporarily stored in a 1 cm thick Perspex/Plexiglas boxes. Such radiation specific safety equipment is available from most large scientific suppliers. Researchers should ensure that all safety procedures comply implicitly with their local radiation protection regulations. Detailed laboratory safety procedures are discussed in further in Castegnaro *et al*. (1993).

Status of quality control procedures and standardized assays

There are currently no active QA programmes running for the detection of DNA adducts using the ^{32}P-postlabeling method. Previous QA programmes have been conducted under the auspices of the EU funded Biological Effects Quality Assurance In Monitoring Programme (BEQUALM) and the International Agency for Research on Cancer (IARC). The IARC QA trial of the ^{32}P-postlabelling assay was conducted between 1994–1997 and involved 25 participants in Europe and the USA. The primary objectives of this project were to standardize the ^{32}P-postlabelling assay and improve inter-laboratory reproducibility. The IARC QA programme for ^{32}P-postlabelling led to a series of publications, which detailed a standardized protocol for the detection of bulky aromatic DNA adducts by the ^{32}P-postlabelling assay (IARC, 1993; Phillips and Castegnaro, 1999). The standardized protocol has now been adopted by the International Programme on Chemical Safety (IPCS) [1] and recommended for use in their guidelines for monitoring genotoxic carcinogens in humans (Richard *et al*., 2000). Essentially the same protocol is also published in an ICES Times technical document (Reichert *et al*., 1999).

Assessment criteria

It is recognized that setting baseline/background response levels have an important role in integrating biological effect parameters into environmental impact assessments of the marine environment. The general philosophy is that an elevated level of a particular biomarker, when compared with a background response, indicates that a hazardous substance has caused an unintended or unacceptable level of biological effect. Therefore, in order to understand and apply DNA adducts as a biomarker of genotoxic exposure it is of fundamental importance to gain information on the natural background levels in non-contaminated organisms. A number of studies have now examined fish collected from pristine areas (as supported by chemical and biomarker analyses) and the typical ^{32}P-postlaballing generated DNA adduct profiles either exhibited no detectable adducts or very faint diagonal radioactive zones (DRZs) (Figure 1A), suggesting minimal PAH exposure (Ericson *et al*., 1998; Reichert *et al*., 1998; Lyons *et al*., 2000; Aas, *et* al., 2003; Balk *et al*., 2006). In contrast, DNA adduct profiles in fish exposed to a complex mixture of PAHs will form DRZs on the chromatogram (Fig1B), which is a composite of multiple overlapping PAH-DNA adducts.

[1] International Programme on Chemical Safety (IPCS) was established in 1980 under the WHO, for more information visit: http://www.who.int/ipcs/en/

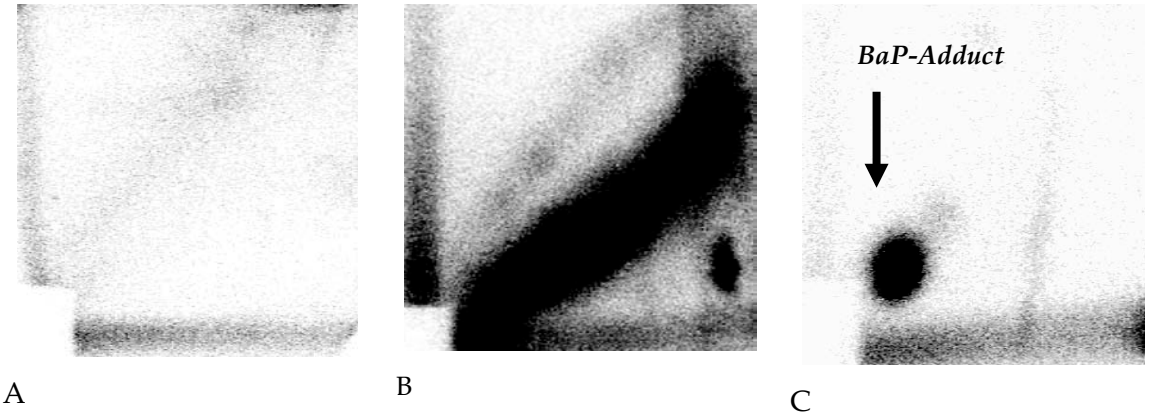


Figure 1. Representative hepatic DNA adducts profiles produced following ^{32}P-postlabelling. (A) DNA adduct profile obtained from a site with a low level of PAH contamination. A faint DRZ is visible, indicating a low level of DNA adducts representative of a clean reference location. (B) DNA adduct profile displaying a clear DRZ of ^{32}P-labelled DNA adducts indicating the fish has been exposed to a complex mixture of genotoxins. (C) Positive control consisting of BaP labelled DNA (115 nucleotides per 10^8 undamaged nucleotides) run with each batch (kindly provided by Professor David Phillips and Dr Alan Hewer, Cancer Research Institute, Sutton UK). Figure adapted from Lyons *et al.*, 2004b).

Determination of threshold level of significant effects (EAC) for DNA adducts in cod

The determined 90 percentile background level for DNA adducts in cod can be used to express the elevated-above-background level, however this level is not associated with significant effects on fitness in whole organisms. Therefore we have also defined a threshold value of significant effects. This is achieved by combining fitness effect data with DNA adduct data at corresponding oil concentrations.

Dose:reponse relationships between exposure concentrations of oil and DNA adducts in cod have been established in laboratory studies. We have used data from Skadsheim (2004), Skadsheim *et al.* (2009). Determination of significant whole organism effects on fitness is more uncertain. We have here assumed that this threshold level is found between 0,5 and 1,0 ppm of oil. We base this on reproduction effect data in model fish species *Cyprinodon variegatus* exposed to oil (Anderson *et al.*, 1977). This data has later been included in generic species sensitivity distribution for chronic whole organism effects (Scholten *et al.*, 1993; Smit *et al.*, 2009). This corresponds to mortality levels found in larval studies with the North East Atlantic relevant species cod, herring and halibut exposed to oil (Morton *et al.*, in submission; Ingvarsdottir *et al.*, in prep.).

Within the concentration range from 0,5 to 1,0 ppm oil, DNA adduct formation tends to increase strongly (Skadsheim, *op.cit*). The interpolated DNA adduct value at mid-range (0,75 ppm oil) was 6 nmol adducts pr. mol nucleotides. A similar value has also been found for flatfish (turbot) at this oil concentration (Jonsson *et al.*, in prep.). However, species differences exists, and in sprat (*Sprattus sprattus*), belonging to the herring family, the corresponding DNA adduct value is 1,4 nmol adducts pr. mol nucleotides (Bjørkblom *et al.*, in prep.).

The following issues are important and require consideration:

- ^{32}P-postlabelling studies should be conducted using internationally agreed protocols incorporating appropriate positive and negative control samples (Phillips and Castegnaro, 1999; Reichert *et al*., 1999).
- All studies need to include supporting environmental data to confirm the contaminant load at the reference location and where possible supporting biomarker and histopathological data to confirm health status of the individual.
- While the assay ^{32}P-postlabelling can be applied to any species deemed fit for purpose, it should only be applied to those species where there is sufficient background information available on life-history traits and behaviour (e.g. migration).

Derivation of assessment criteria

The UK has monitored DNA adducts in dab at offshore locations at 15 sites and for flounder in eight estuaries. Using these studies it has been possible to define reference locations and develop background response ranges. The approach used is similar to that adopted by the US EPA on Effect Range(ER) values. The ER-Low (ERL) value is defined as the lower tenth percentile of the effect. Data were available from Norway (IRIS and NIVA) for other species (IRIS database; BioSea project – Total E&P Norge & Eni Norge); data were reported as nmol adducts/mol DNA. The UK expressed results as adducted nucleotides per 10^8 normal nucleotides, which was converted to nmol adducts/mol DNA by dividing by 10.

Levels below 1 nmol adducts/mol normal nucleotides are considered too low to confirm genotoxic exposure. The derived ER-Low values for dab and flounder were 1.0 (background), and for Atlantic cod it was 1.6 (background) and for Haddock (Barents Sea) it was 3.0 (subtracting a species-specific spot). Threshold value assigned for significant effects (EAC) in Atlantic cod was 6 (see Assessment criteria (above) for method of estimation).This value is also indicative for flatfish (to be verified), but for herrings a considerably lower EAC threshold value of 1.4 has been determined.

Concluding remarks

- *DNA adducts as biomarkers of genotoxic exposure*. DNA adducts provide a measure of biologically active contaminant to have reached a critical cellular target (DNA). They are persistent and therefore considered a 'cumulative index' of exposure to genotoxins and a significant body of research demonstrates their importance in the initiation and progression of carcinogenesis induced by important environmental contaminants (e.g. PAHs).*Safety considerations when conducting the 32p-postlabelling assay.* The ^{32}P-postlabeling assay uses large amounts of ^{32}P, which is an energetic beta emitter. This requires specialist laboratories may limit the use of the assay to a few appropriately equipped research groups.*Applicability across OSPAR maritime area*. DNA adducts have been applied in a wide range of species across the whole OSPAR maritime area including blue mussels, *Mytilus* sp, perch (*Perca fluviatilis*), dab (*Limanda limanda*), European flounder (*Platichthys flesus*), eelpout (*Zoarces viviparous*) and cod (*Gadus morhua*). A particular advantage of the ^{32}P-postblabelling assay is that it is not species-specific and therefore can be utilized on any organism deemed fit for purpose.
- *Status of quality assurance*. There are currently no active QA programmes running for the detection of DNA adducts using the ^{32}P-postlabeling

method. However, inter laboratory QA programmes have previously been conducted under the auspices of BEQUALM and IARC and standardized protocols are available in the form of an ICES Times technical document and IARC publications.

- *Assessment criteria*. Provisional assessment criteria have been derived for flounder, dab and Atlantic cod. In addition, background criteria have been set for haddock and long rough dab. These have been derived from datasets from national monitoring programmes within the OSPAR maritime area. It is recommended that further work to refine these values is taken forward as and when new data becomes available through national monitoring programmes and through the activities of ICES WGBEC.

References

Aas, E., Liewenborg, B., Grøsvik, B.E., Campus, L., Jonsson, G., Børseth, J.F., Balk, L. 2003. DNA adduct levels in fish from pristine areas are not detectable or low when analysed using the nuclease P1 version of the ^{32}P-postlabelling technique. Biomarkers, 8 (6), 445-460.

Akcha, F., Leday, G., Pfohl-Leszkowicz, A. 2004. Measurement of DNA adducts and strand breaks in dab (*Limanda limanda*) collected in the field: effects of biotic (age, sex) and abiotic (sampling site and period) factors on the extent of DNA damage. Mutation Research 552(1–2):197–207.

Anderson, J.W., Dixit, D.B., Ward, G.S. and Foster, R.S. 1977. Effects of petroleum hydrocarbons on the rate of heart beat and hatching success of estuarine fish embryos. In Vernberg FJ, Calabrese A, Thurnberg FP, Vernberg WB, eds. *Physiological responses of marine biota to pollutant*. Acad. Press, New York, USA, pp 241–258.

Balk, L., Liewenborg, B., Larsen, B.K., Aas, E., Hylland, K., Sanni, S. 2006. Large hydrophobic adducts and strand breaks analysed in hepatic DNA from Atlantic cod (*Gadus morhua*) caged at the Statfjord oilfield. In Biological effects of contaminants in pelagic ecosystems. Editors Hylland K, Lang T, Thain J, Vethaak AD, Wosniok W. Brussels: Society of Environmental Toxicology and Chemistry (SETAC), 277–295.

Bauman, P.C. 1998. Epizootics of cancer in fish associated with genotoxins in sediment and water. Mutation Research 411, 227–233.

Beach, A.C., Gupta, P.C. 1992. Human biomonitoring and the ^{32}P-postlabelling assay. Carcinogenesis 13, 1053–1074.

Castegnaro, M. Bresil, H., Manin, J.P. 1993. Some safety procedures for handling ^{32}P during postlabelling assays, In Postlabelling Methods for Detection of DNA Adducts. Ed. D.H. Phillips, M. Castegnaro and H. Bartsch. Lyon, International Agency for Research on Cancer. IARC Scientific Publications 124, 87–92.

Ericson, G., Lindesjoo, E., Balk, L. 1998. DNA adducts and histopathological lesions in perch (*Perca fluviatilis*) and northern pike (*Esox lucius*) along a polycyclic aromatic hydrocarbon gradient on the Swedish coastline of the Baltic Sea. Canadian Journal of Fisheries and Aquatic Sciences, 55, 815–824.

Ericson, G., Noaksson, E., Balk, L. 1999. DNA adduct formation and persistence in liver and extrahepatic tissues of northern pike (*Esox lucius*) following oral exposure to benzo[a]pyrene, benzo[k]fluranthene and 7H-dibenzo[c,g]carbazole. Mutation Research, 427, 136–145.

Ericson, G., Skarphedinsdottir, H., Zuanna, L.D., Svavarsson, J. 2002. DNA adducts as indicators of genotoxic exposure in indigenous and transplanted mussels, *Mytilus edulis* L. from Icelandic coastal sites. Mutation Research 516(1–2):91–9.

French, B.L., Reichart, W.L., Hom, T., Nishimoto, M., Sanborn, H.R., Stein, J.E. 1996. Accumulation and dose-response of hepatic DNA adducts in English sole (*Pleuronectes vetulus*) exposed to a gradient of contaminated sediments. Aquatic Toxicology, 36 (1–2), 1–16.

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

Harvey, J.S., Parry, J.M. 1998. The analysis of DNA adduct formation, removal and persistence in the common mussel *Mytilus edulis* exposed to 4-nitroquinoline 1-oxide. Mutation Research, 399, 31–42.

IARC. 1993. Postlabelling Methods For Detection of DNA adducts. Edited by D, H. Phillips, M. Castegnaro, and H. Barysch. IARC Scientific Publications, Lyon. 1993. No. 124. 392 pages.

Lyons, B.P., Harvey, J.S., Parry, J.M. 1997. An initial assessment of the genotoxic impact of the Sea Empress oil spill by the measurement of DNA adduct levels in the intertidal teleost *Lipophrys pholis*. Mutation Research, 390, 263–268.

Lyons, B.P., Stewart, C., Kirby, M.K. 1999. The detection of biomarkers of genotoxin exposure in the European flounder (*Platichthys flesus*) collected from the River Tyne Estuary. Mutation Research, 446, 111–119, p116.

Lyons, B.P., Stewart, C., Kirby, M.K. 2000. 32P-postlabelling analysis of DNA adducts and EROD induction as biomarkers of genotoxin exposure in dab (*Limanda limanda*) from British coastal waters. Marine Environmental Research 50 (2000) 575–579.

Lyons, B.P., Bignell, J., Stentiford, G.D., Feist, S.W. 2004a. The viviparous blenny (*Zoarces viviparus*) as a bioindicator of contaminant exposure: application of biomarkers of apoptosis and DNA damage. Marine Environmental Research, 58, 757–762.

Lyons B.P., Stentiford, G.D., Green, M., Bignell,J., Bateman, K., Feist, S.W., Goodsir, F. Reynolds, W.J. Thain, J.E. 2004b. DNA adduct analysis and histopathological biomarkers in European flounder (*Platichthys flesus*) sampled from UK estuaries. Mutation Research, 552, 177–186.

Malins, D.C., Krahn, M.N., Myers, M.S., Rhodes, L.D., Brown, D.W., Krone, C.A., McCain, B.B., Cham, S. 1985. Toxic chemicals in sediments and biota from a creosote-polluted harbour. Relationships with hepatic neoplasms and other hepatic lesions in English sole (*Parophyrs vetulus*). Carcinogenesis, 6, 1463–1469.

Morton, C.H., G. Nyhammer, B.E. Grøsvik, V. Makhotin, A. Geffen, S. Boitsov, K.A. Kvestad, A.B. Kjersem, A. Goksøyr, A. Folkvord, J. Klungsøyr, A. Svardal, S. Meier. in submission. Development of Atlantic cod (*Gadus morhua*) exposed to produced water during early life stage. I. Effects on embryos, larvae, and juvenile fish.

Myers, M.S., Landahl, J.T., Krahn, M.M., McCain, B.B. 1991. Relationships between hepatic neoplasms and related lesions and exposure to toxic chemicals in marine fish from the US West Coast. Environmental Health Perspectives, 90, 17–26.

Phillips, D.H., Castegnaro, M. 1993. Results of an interlaboratory trail of 32P-postlabelling. In Postlabelling Methods for Decetion of DNA Adducts. Ed. D.H. Phillips, M. Castegnaro and H. Bartsch. Lyon, International Agency for Research on Cancer. IRAC Scientific Publications 124, 35–50.

Phillips, D.H., Castegnaro, M. 1999. Standardization and validation of DNA adduct postlabelling methods:report of interlaboratory trails and production of recommended protocols. Mutagenesis 14 (3), 301–-315.

Phillips, D.H. 2005. DNA adducts as markers of exposure and risk. Mutation Research 577, 284–292.

Phillips, D.H. 1997. Detection of DNA modifications by the 32P-postlabelling assay. Mutation Research, 378, 1–12.

Reddy, M.V., Randerath, K 1986. Nuclease-P1 mediated enhancement of sensitivity of the 32P-postlabelling test for structurally diverse DNA adducts. Carcinogenesis 9, 1543–1551.

Reichert, W.L., Myers, M.S., Peck-Miller, K., French, B., Anulacion, B.F., Collier, T.K., Stein, J.E., Varanasi, U. 1998. Molecular epizootiology of genotoxic events in marine fish: Linking contaminant exposure, DNA damage, and tissue-level alterations. Mutation Research, 411, 215–255.

Reichert, W.L., French, B.L,. Stein, J.E. 1999. Biological effects of contaminants: Measurements of DNA adducts in fish by 32P-postlabelling. ICES Techniques in Marine Environmental Sciences. No. 25.

Richard J. Albertini, R.J., Diana Anderson, D Douglas, D.R., Hagmar, L., Hemminki, K, Merlo, F., Natarajan, A.T., Norppa, H., Shuker, D.E.G., Tice, R., Waters, M.D., Aitio, A. 2000. IPCS guidelines for the monitoring of genotoxic effects of carcinogens in humans. Mutation Research-Reviews in Mutation Research. 463 (2), 111–172.

Sikka, H.C., Steward, A.R., Kandaswami, C., Rutkowski, J.P., Zaleski, J., Earley, K., Gupta, R.C. 1991. Metabolism of benzo[a]pyrene and persistence of DNA adducts in the brown bullhead (*Ictalurus nebulosus*). Comparative Biochemistry and Physiology, 100c, 25-28.

Scholten, M.C.Th, Schobben, H.P.M, Karman, C.C., Jak, R.G. and van het Groenewoud, H. 1993. De berekening van het maximal toetlaatbaar risico niveau van olie en oliecomonenten in water en sediment. Technical report R93/187, Dutch Organisation for Applied Scientific Research (TNO), Den Helder, The Netherlands.

Skadsheim, A. 2004. Cod stocks exposed to crude oils: uptake, metabolism and biomarker effects. Technical report AM-2004/005, Akvamiljo, Randaberg, Norway.

Skadsheim, A., Sanni, S., Pinturier, L., Moltu, U.-E., Buffagni, M. and Bracco, L. 2009. Assessing and monitoring long-range-transported hydrocarbons as potential stressors to fish stocks. Deep Sea Research Part II. Vol. 56, Issues 21–22, pp. 2037–2043.

Sleiderink, H.M., Beyer, J., Scholtens, E., Goksyr, A., Nieuwenhuize, J., Van Liere, J.M., Everaarts, J.M., Boon, J.P. 1995. Influence of temperature and polyaromatic contaminants on CYP1A levels in North Sea dab (*Limanda limanda*). Aquatic Toxicology, 32, 189–209.

Smit, M.G.D, Bechmann, R.K., Hendriks, A.J., Skadsheim, A., Larsen, B.K., Baussant, T., Bamber, S. and Sanni, S. 2009. Relating biomarkers to whole-organism effects using species sensitivity distributions: A pilot study for marine species exposed to oil. Environmental Toxicology and Chemistry, Vol. 28, No. 5, pp. 1104–1109.

Stein, J.E., Reichert, W.L., Nishimoto, M., Varanasi, U. 1990. Overview of studies on liver carcinogenesis in English sole from Puget Sound; evidence for a xenobiotic chemical etiology II: Biochemical studies. Science of the Total Environment, 94 51–69.

van der Oost, R., van Schooten, F.J., Ariese, F., Heida, H. 1994. Bioaccumulation, biotransformation and DNA binding of Pahs in feral eel (*Anguilla anguilla*) exposed to polluted sediments. Environmental Toxicology and Chemistry, 13: 859–870.

Annex 7: PAH metabolites in bile

Version date: 29 January 2010.

Background

Analyses of PAH metabolites in fish bile have been used as a biomarker of exposure to PAH contamination since the early 1980s. The presence of metabolites in bile (and in urine) is the final stage of the biotransformation process whereby lipophilic compounds are transformed to a more soluble form and then passed from the organism in bile or urine.

As a biomarker of exposure, measuring PAH metabolites in bile has many advantages over other techniques that require sophisticated tissue preparation protocols. The pretreatment of bile samples requires relatively simple dilution steps prior to analysis by direct fluorescence measurement. The bile is diluted in methanol:distilled water (1:1) and fluorescence is measured with a fluorometer. Fixed wavelength fluorescence is a suitable screening method for samples while HPLC/F or GC-MS SIM is utilized for qualitative and quantitative measures (Ariese *et al*., 2005; Jonsson *et al*., 2003; Lin *et al*., 2006; Aas *et al*., 2000a; 2000b).

Bile is generally stored in the gall bladder prior to episodic release into the esophagus where bile salts have a function to perform as part of the digestive process. This period of storage permits a degree of accumulation of metabolites and hence an increase in their concentration. The periodic release of bile does however introduce a variable into the technique, which must be accounted for. The feeding status of fish has been shown to influence both the volume and the density of the bile (Collier and Varanasi, 1991).

The ability of fish to biotransform PAHs into less lipophilic derivatives means that reliance on the detection of parent PAHs alone may lead to an underestimation of the *in vivo* exposure level of PAH in the fish. PAH metabolite detection, on the other hand, represents a quantification of the flux of PAHs streaming through the fish's body. From a toxicological point of view, flux information is more relevant to estimating the actual biotic stress due to PAH exposure, than the body burden data of the unmetabolised parent PAH compounds in tissues (most often liver). Despite this, body burden measurements are still more commonly used within monitoring studies than metabolite determination.

Dose–response (species–specific)

The PAH compounds are metabolised rapidly in the organisms and it is the endpoint of this metabolisms that is measured in the bile. The compounds are measured using chemical analysis. A consistent dose-response relationship has been demonstrated in laboratory studies between PAH exposure and the subsequent presence of metabolites in bile (Beyer *et al*., 1997; Aas *et al*., 2000). To establish a good dose-response relationship in field studies it is necessary to focus on aspects that influence the excretion of bile.

The method requires that bile is available in the gall bladder. Because the fish renew bile as part of normal metabolism and excrete it during digestion, it is important to know about the dietary status of the organism to establish a dose- response relationship. If the fish feed just before sampling, the gall bladder may become more or less

empty. After the gall bladder has been emptied, it will fill up and metabolites will be concentrated up to a plateau level corresponding to the exposure regime. Consequently the time since last digestion is important for the dose-response relationship. Fish generally have a very efficient metabolic excretion of most PAHs and it has been shown that most of the PAH will excreted after 2–8 days following exposure. This means that the PAH metabolites determined in bile will represent exposures on the scale of days and, at most, two weeks.

It has been shown in several field and laboratory studies that there is a good correlation between PAH exposure and bile metabolites. Because of the rapid metabolism and the correlation between bile content and digestive status it is difficult to make a dose-response relationship that can be used to quantify the exposure. Work has been done to try to correlate bile metabolite concentration to digestive status, by correlating it to the amount of protein or biliverdin in the bile. Absorbance at 380 nm is also used (similar to biliverdin) (Hylland, unpublished). This normalization is not standardized because it has been shown to only explain parts of the variability, but it is recommended to be part of the explaining factors in the interpretation of results. In laboratory studies it is normal to stop the feeding some days before sampling to ensure the bile quality. In field sampling this can be taken into account by letting the fish go some days in tanks before sampling, but this has some logistical challenges.

Species sensitivity

The background level differs between species so it is important to establish good baseline before using new species. It may be expected that species with fatty livers, i.e. most gadiids, may metabolise PAHs more slowly as more will partition into fat, but this has not been documented experimentally. Species differences have in general to be considered when calculating BACs and EACs, though in some cases the resulting ACs are so similar that a combined AC for several species is justified (Table 3 below gives an example).

Relevance of other factors

As mentioned above, food availability will affect the concentration of PAH metabolites in bile. In an assessment of data for more than 500 individual cod sampled through five years of national monitoring, variables such as size/age and sex explained some variability of multiple regression models (Ruus *et al*., 2003). This could be due to different feeding preferences, but also endogenous processes. In addition, the fat-content of the liver (measured as liver-somatic index, LSI) came out as significant, presumably because fat decreases the availability of PAH to the cellular compartments of liver cells. There are indications for seasonal differences between summer and winter values of PAH metabolites in dab (Kammann, 2007).

Background responses

Baseline levels of PAH metabolites have been established for many of the species relevant to monitoring in Norwegian coastal and offshore waters. From Ruus *et al*. (2003) values for the relevant species are: (all values standardized to absorbance at 380 nm) Atlantic cod: 0.6–4 μg/kg bile, flounder 27–89 μg/kg bile, dab 3.1–34 μg/kg bile, plaice 0.4–3 μg/kg bile (all quantified using HPLC separation and fluorescence detection and quantification). Standardisation at 380 nm is used to remove variability due to bile salts.

Assessment criteria

Assessment Criteria for PAH metabolites such as BAC (Background Assessment Criteria) have been derived from reference sites (Table 3). EAC (Environmental Assessment Criteria) can be derived from toxicological experiment data by linking oil exposure and PAH metabolites in fish with DNA-adducts and fitness data (Morton *et al*., subm; Skadsheim *et al*., 2004; Skadsheim *et al*., 2009), where the latter serves as the effect quantity for the calculation of the EAC presented in Table 3. Some variation in PAH metabolites in bile appear to be related to sex and size/age (Ruus *et al*., 2003), knowledge of which should be included in the sampling design.

Quality assurance

A general protocol outlining analytical strategies and their strengths as well as weaknesses has recently become available (Ariese *et al*., 2005). There have been international intercalibration exercises for the determination of PAH-metabolites in fish bile, arranged in collaboration between an EU-project and QUASIMEME [2]. Reference bile samples were generated as part of the aforementioned EU project and are now available through IRMM, JRC, Geel, Belgium (http://www.irmm.jrc.be/html/homepage.html). An intercalibration for PAH metabolites is taking place in the framework of the EU funded BONUS project "BEAST" in 2010.

Acknowledgement

The current review has been derived from an overview prepared for the Norwegian offshore companies through OLF (Hylland *et al*., 2006).

[2] QUASIMEME; organisation that offers quality assurance for chemical endpoints; http://www.quasimeme.org

References

Aas, E., T. Baussant, *et al*. 2000a. PAH metabolites in bile, cytochrome P4501A and DNA adducts as environmental risk parameters for chronic oil exposure: a laboratory experiment with Atlantic cod. Aquatic Toxicology 51, 241–258.

Aas, E., Beyer, J., Goksøyr, A. 2000b. 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.

Ariese, F., Beyer, J., Jonsson, G., Visa, C.P., Krahn, M.M. 2005. Review of analytical methods for determining metabolites of polycyclic aromatic compounds (PACs) in fish bile. ICES Techniques in Marine Environmental Sciences, 39: 1–41.

Beyer J, Sandvik M, Skåre JU, Egaas E, Hylland K, Waagbø R, Goksøyr A. 1997. Time- and dose-dependent biomarker responses in flounder (*Platichthys flesus* L.) exposed to benzo[a]pyrene, 2,3,3',4,4',5-hexachlorobiphenyl (PCB-156) and cadmium. Biomarkers 2: 35–44.

Collier, T.K. and 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.

Hylland, K., Ruus, A., Børseth, J.F., Bechmann, R., Barsiene, J., Grung, M., Tollefsen, K., Myhre, L.P. 2006. Biomarkers in monitoring - a review. NIVA-report 5205. 106 p.

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, 2003. 5: p. 513–520.

Kammann U. 2007. PAH metabolites in bile fluids of dab (*Limanda limanda*) and flounder (*Platichthys flesus*) – spatial distribution and seasonal changes. Environ Sci Pollut Res 14(2) 102–108.

Morton, C.H., G. Nyhammer, B.E. Grøsvik, V. Makhotin, A. Geffen, S. Boitsov, K.A. Kvestad, A.B. Kjersem, A. Goksøyr, A. Folkvord, J. Klungsøyr, A. Svardal, and S. Meier (in subm.): Development of Atlantic cod (*Gadus morhua*) exposed to produced water during early life stage. I. Effects on embryos, larvae, and juvenile fish.

Lin, E.L.C., Cormier, S.M.Torsella, J.A. 1996. Fish biliary polycyclic aromatic hydrocarbon metabolites estimated by fixed-wavelength fluorescence: Comparison with HPLC-fluorescent detection. Ecotoxicology and environmental safety 35, 16–23.

Ruus, A, Hylland, K, Green, N. 2003. Joint Assessment and Monitoring Programme (JAMP).Biological Effects Methods, Norwegian Monitoring 1997–2001. SFT Report no. 869/03. TA- 1948/2003, 139 p.

Skadsheim, A. 2004. Cod stocks exposed to crude oils: uptake, metabolism and biomarker effects. Technical report AM-2004/005, Akvamiljo, Randaberg, Norway.

Skadsheim, A., Sanni, S., Pinturier, L., Moltu, U.-E., Buffagni, M. and Bracco, L. 2009. Assessing and monitoring long-range-transported hydrocarbons as potential stressors to fish stocks. Deep Sea Research Part II. Vol. 56, Issues 21–22, pp. 2037–2043.

Table 1. Overview of field and laboratory studies – PAH metabolites measured by fixed fluorescence.

Species	Substance (lab/field)	Test concentrations/area	Exposure time	Metabolite	Baseline	control or reference	exposed /control	reference/comments
Cod (*Gadus morhua*)	Feral fish	Barents Sea	Baseline					Aas *et al*., 2003
Cod (*Gadus morhua*)	Feral fish	Egersund	Baseline non polluted area	Naph type Pyren type BaP type	5.3 ug/ml 0.8 ug/ml 0.4 ug/ml			Klungsøyr *et al*., 2003
Cod (*Gadus morhua*)	Feral fish	Sleipner	Baseline polluted area?	Naph type Pyren type BaP type	6.1 ug/ml 1.0 ug/ml 0.5 ug/ml			Klungsøyr et al. 2003
Cod (*Gadus morhua*)	Feral fish	Statfjord	Baseline polluted area?	Naph type Pyren type BaP type	5.9 ug/ml 0.9 ug/ml 0.3 ug/ml			Klungsøyr *et al*., 2003

Species	Substance (lab/field)	Test concentrations/area	Exposure time	Metabolite	Baseline	control or reference	exposed /control	reference/comments
Cod (*Gadus morhua*)	Feral fish	Frøy, ceased installation 10 000 m (ref) 2000 m–200 m	Baseline polluted area?	Naph type		3.9 ug/ml	1.1–1.1	Beyer *et al*., 2003
				Pyren type		0.6 ug/ml	1.1–0.9	
				BaP type		0.3 ug/ml	0.9–0.9	
Cod (*Gadus morhua*)	Feral fish	Barents sea	Baseline	Naph type	2,15 ug/g			Sundt, 2002
				Pyren type	1,63 ug/g			
				BaP type	0,69 ug/g			
Cod (*Gadus morhua*)	Feral fish	Barents sea	Baseline	Naph type	5,8 ug/g			Aas and Børseth, 2002
				Pyren type	1,7 ug/g			
				BaP type	0,8 ug/g			
Cod (*Gadus morhua*)	Laboratory	1 ppm crude oil Statfjord B	14 days					Aas *et al*., 2002
Cod (*Gadus morhua*)	Laboratory	0.06–0.25–1 ppm Oil	average 3, 7, 14, 24 days	Naph type		3,9 ug/g	7,5–23, 7–31,4	Skadsheim *et al*., 2004
				Pyren type		2,6 ug/g	3,6–10, 6–13	
				BaP type		1,0 ug/g	1,7–2, 4–2,2	

Species	Substance (lab/field)	Test concentrations/area	Exposure time	Metabolite	Baseline	control or reference	exposed /control	reference/comments
Cod (*Gadus morhua*)	Laboratory	0.06–0.25–1 ppm Oil	average 3, 17, 31 day	Naph type		53.1 ug/g	0.7–2.3–2.9	Skadsheim *et al*., 2004
				Pyren type		7.0 ug/g	1–2.9–3.3	
				BaP type		1.0 ug/g	1.1–1.5–1.5	
Cod (*Gadus morhua*)	Laboratory	Oil 0.06–0.25–1 ppm	30 days	Naph type		7.1 fi	5.1–9.5–227.5	Aas *et al*., 2000
				Pyren type		2 fi	6.4–12.7–43.3	
				BaP type		0.8 fi	2.3–3.6–9.6	
Cod (*Gadus morhua*)	Laboratory	PW Oseberg, 1:1000–1:200–0.2 ppm oil–0.2 ppm oil + PAHmix	15 days	Naph type		12.6 ug/ml	1.3–2.5–3.6–5.4	Sundt, 2004
				Pyren type,		4 ug/ml	1.7–3.7–4.1–17.8	
				BaP type,		1.8 ug/ml	1.3–1.8–1.5–2.4	
Cod (*Gadus morhua*)	Field, Caged	North Sea - Statfjord, 10 000 m–2000 m–500 m German bight G	5.5 weeks	Naph type	7.5 ug/ml	0,7	1.7–1.9–2.1	Aas *et al*., in press
				Pyren type	3.1 ug/ml	0,7	1.2–1.5–1.6	
				BaP type	1.2 ug/ml	0,8	1.2–1.1–1.2	

Species	Substance (lab/field)	Test concentrations/area	Exposure time	Metabolite	Baseline	control or reference	exposed /control	reference/comments
Cod (*Gadus morhua*)	Field, Caged	German bight G4 (Ref) G1 - G2 - G3	5.5 weeks	Naph type	7.5 ug/ml	0,4	0.9–0.9–1.6	Aas *et al*., in press
				Pyren type	3.1 ug/ml	0,5	0.8–0.9–1.7	
				BaP type	1.2 ug/ml	0,7	0.8–1–1.3	
Cod (*Gadus morhua*)	Field, Caged	North Sea - Troll, 1000 m–500 m	6 weeks	Naph type	4.6 ug/ml	1,4	1.7–2.5	Børseth *et al*., 2004
				Pyren type	2.4 ug/ml	0,9	1.1–1.3	
				BaP type	0.9 ug/ml	1,1	1.1–1.3	
Cod (*Gadus morhua*)	Field, Caged	North Sea - Tampen, 10 000–2500–1000–500	6 weeks	Naph type		8.8 ug/ml	1.0–1.5–1.2–1.2	Hylland *et al*., 2005
				Pyren type				
				BaP type		1.4 ug/ml	0.9–0.7–0.8–0.9	
Haddock (*Melanogrammus aeglefinus*)	Feral fish	Egersund	Baseline non polluted area	Naph type	5.1 ug/ml			Klungsøyr *et al*., 2003
				Pyren type	1.4 ug/ml			
				BaP type	0.7 ug/ml			
Haddock (*Melanogrammus aeglefinus*)	Feral fish	Sleipner	Baseline polluted area?	Naph type	6.8 ug/ml			Klungsøyr et al. 2003
				Pyren type	1.9 ug/ml			
				BaP type	0.8 ug/ml			
Haddock (*Melanogrammus aeglefinus*)	Feral fish	Statfjord	Baseline polluted area?	Naph type	11.2 ug/ml			Klungsøyr *et al*., 2003
				Pyren type	2.5 ug/ml			

Species	Substance (lab/field)	Test concentrations/area	Exposure time	Metabolite	Baseline	control or reference	exposed /control	reference/comments
				BaP type	0.7 ug/ml			
Haddock (*Melanogrammus aeglefinus*)	Feral fish	Barents sea		Naph type	2.52 ug/g			Sundt, 2004
				Pyren type	1.69 ug/g			
				BaP type	0.77 ug/g			
Haddock (*Melanogrammus aeglefinus*)	Feral fish	Barents sea		Naph type	2.0 ug/g			Aas and Børseth, 2004
				Pyren type	1.3 ug/g			
				BaP type	0.6 ug/g			
Haddock (*Melanogrammus aeglefinus*)	Feral fish	Frøy, ceased installation 10 000 m (ref) 2000 m–200 m	Baseline polluted area?	Naph type		5.6 ug/ml	1.3 - 2.2	Beyer *et al*., 2003
				Pyren type		1.4 ug/ml	1.4 - 0.7	
				BaP type		0.75 ug/ml	1.8 - 0.6	
Sheepshead minnow	Laboratory	North sea oil A 0.1–0.4–0.7 ppm	5 weeks	Naph type		6916	2.3–6.2–9.3	Bechmann *et al*., 2004
				Pyren type		569	2.5–5–6.3	
				BaP type		107	4–13.1–19.2	
Sheepshead minnow	Laboratory	North sea oil B 0.1–0.9–5.6 ppm	6 weeks	Naph type		18 164	1.8–4.3–12.5	Bechmann *et al*., 2004
				Pyren type		438	5.6–12.6–30.8	
				BaP type		110	12.6–42.7–123.9	
Sheepshead minnow	Laboratory	2–14–214 ppb	5 weeks	Naph type		267 280	0.9–2.2–18.6	Bechmann *et al*., 2004
				Pyren type		9926	0.9–1.5–9.6	
				BaP type		5152.7	3–17.4–207	

Species	Substance (lab/field)	Test concentrations/area	Exposure time	Metabolite	Baseline	control or reference	exposed /control	reference/comments
Polar cod (*Boreogadus saida*)	Laboratory, feral fish 2001, 2002	1.5 ppm StatfjA oil , baseline, control	14 days	Naph type	16.0 ug/g	2	16,9	Sundt and Bechmann, 2004
				Pyren type	0.9 ug/g	5,5	74,4	
				BaP type	0 ug/g	0	1,8	

Table 2. PAH-metabolites in marine fish; measured by GC-MS.

Species	Substance (lab/field)	Test concentrations	Exposure time	Metabolite	Baseline	control or reference	exposed/control	reference
Cod (*Gadus morhua*)	Feral fish	Barents sea	baseline	Naph sum	150,6 ng/g			Aas and Børseth, 2002
				Phen sum	61,2 ng/g			
				Pyren	4,6 ng/g			
Cod (*Gadus morhua*)	Feral fish	Barents sea	baseline	Naph sum	1285 ng/g			Sundt, 2004
				Phen sum	220 ng/g			
				Pyren	3.5 ng/g			
Cod (*Gadus morhua*)	Feral fish	Egersund	Baseline non polluted area	Naph sum	2005.1 ng/g			Klungsøyr *et al.*, 2003
				Phen sum	230.2 ng/g			
				Pyren	3.9 ng/g			
Cod (*Gadus morhua*)	Feral fish	Sleipner	Baseline polluted area?	Naph sum	1296.1 ng/g			Klungsøyr *et al.*, 2003
				Phen sum	197.8 ng/g			
				Pyren	0 ng/g			
Cod (*Gadus mor-*	Feral fish	Statfjord	Baseline polluted	Naph sum	1361.7 ng/g			Klungsøyr *et al.*,

Species	Substance (lab/field)	Test concentrations	Exposure time	Metabolite	Baseline	control or reference	exposed/control	reference
hua)			area?	Phen sum	351.1 ng/g			2003
				Pyren	4.0 ng/g			
Cod (*Gadus mor-*	Laboratory	0.06–0.25–1 ppm Oil	average 3, 7, 14, 24 days	Naph sum		2549 ng/g	4.6–13.4–23.6	Skadsheim *et al.*, 2004
hua)				Phen sum		691 ng/g	7.7–22.9–34.9	
				Pyren		27 ng/g	7.3–16.2–25.1	
Cod (*Gadus mor-*	Laboratory	0.06–0.25–1 ppm Oil	average 3, 17, 31 day	Naph sum		5702 ng/g	4–13.3–12,7	Skadsheim *et al.*, 2004
hua)				Phen sum		377 ng/g	10,5–40,3–48,7	
				Pyren		5 ng/g	8,6–63–88,4	
Cod (*Gadus mor-*	Field, Caged	North Sea - Statfjord, 500–2000–10 000 m		Naph sum		1150 ng/g	3.0–2.0–1.3	Aas *et al.*, in press
hua)				Phen sum		340 ng/g	3.5–2.7–2.5	
				Pyren		-	-	
Cod (*Gadus mor-*	Field, Caged	North Sea - Troll, 1000 m–500 m	6 weeks	Naph sum	1515.1 ng/g	1,1	1.1–1.2	Børseth *et al.*, 2004
hua)				Phen sum	327.2 ng/g	1,6	2.1–2.0	
				Pyren	173.2 ng/g	1,2	0.9–1.2	
Cod (*Gadus mor-*	Field, Caged	North Sea - Tampen, 10 000–2500–1000–500	6 weeks	Naph sum		965.3 ng/g	0.9–1.7–0.9–1	Hylland *et al.*, 2005
hua)				Phen sum		934.5 ng/g	1.4–3–1.8–1.5	
				Pyren		3.7 ng/g	0–0–0.5–0.0	
Cod (*Gadus mor-*	Field, Caged	North Sea - Statfjord, 10 000 m–	5.5 weeks	Naph sum	228 ng/g	0,2	0.9–1.1–0.9	Aas *et al.*, in press
hua)				Phen sum	482 ng/g	2,0	3–4.5–6.7	

Species	Substance (lab/field)	Test concentrations	Exposure time	Metabolite	Baseline	control or reference	exposed/control	reference
		2000 m–500 m		Pyren	28 ng/g	10,2	29.5–31.1–41.5	
Cod (*Gadus morhua*)	Field, Caged	German bight G4 (Ref) G1 - G2 - G3	5.5 weeks	Naph sum	228 ng/g	0,8	1–1–1.9	Aas *et al*., in press
				Phen sum	482 ng/g	1,0	0.7–0.8–0.8	
				Pyren	28 ng/g	0,0	0–0–0	
Haddock (*Melanogrammus aeglefinus*)	Feral fish	Egersund	Baseline non polluted area	Naph sum	1346.9 ng/g			Klungsøyr *et al*., 2003
				Phen sum	526.8 ng/g			
				Pyren	5.7 ng/g			
Haddock (*Melanogrammus aeglefinus*)	Feral fish	Sleipner	Baseline polluted area?	Naph sum	1111.5 ng/g			Klungsøyr *et al*., 2003
				Phen sum	331.5 ng/g			
				Pyren	10.4 ng/g			
Haddock (*Melanogrammus aeglefinus*)	Feral fish	Statfjord	Baseline polluted area?	Naph sum	1279.7 ng/g			Klungsøyr *et al*., 2003
				Phen sum	331.9 ng/g			
				Pyren	3.1 ng/g			
Haddock (*Melanogrammus aeglefinus*)	Feral fish	Barents sea		Naph sum	1474 ng/g			Sundt, 2004
				Phen sum	165 ng/g			
				Pyren	0			
Polar cod (*Boreogadus saida*)	Laboratory, feral fish 2001, 2002	1.5 ppm StatfjA oil , baseline, control	14 days	Naph sum	1330 ng/g	1,3	114	Sundt and Bechmann, 2004
				Phen sum	538 ng/g	0,9	90	
				Pyren	52 ng/g	14,6	60	

Table 3. Biological assessment criteria (BAC) and Environmental assessment criteria (EAC) for two PAH metabolites, different fish species and methods. Data partly taken from WKIMC 2009.

Biological Effect	Fish species	BAC [ng/ml] HPLC-F	EAC [ng/g] GC/MS
Bile metabolit 1-hydroxypyrene	dab	16	
	cod	21	483
	flounder	16 4)	
	haddock	13	
	dab, cod, haddock	17	
	turbot		909
	halibut		745
Bile metabolit 1-hydroxyphenanthrene	dab	3,7	
	cod	2.7	518
	flounder	3.7 4)	
	haddock	0.8	
	dab, cod, haddock	2.4	
	turbot		1832
	halibut		262

Biological Effect	Fish species	BAC [µg/ml] syncronuos Fluor. 341/383	EAC [µg/ml] Fixed Fluor. 341/383
Bile metabolites of pyrene-type	dab	0.15	22 1)
	cod	1.1	35
	flounder	1.3	29 2)
	haddock	1.9	35 3)
	turbot		29
	halibut		22
	herring/sprat		16

AC based on 1)halibut, 2)turbot, 3)cod and 4)dab

Annex 8: Externally visible fish diseases, macroscopic liver neoplasms and liver histopathology

Version date : 29 January 2010.

Compiled by S W Feist[3] and T Lang[4]

Summary

Applicability across OSPAR maritime area. Externally visible fish diseases have been used internationally for many years as an integrative response for general biological effects monitoring, measuring the general health status at the individual and population level. The method is used for a variety of fish species, including dab (*Limanda limanda*), flounder (*Platichthys flesus*) and cod (*Gadus morhua*) and is easily adaptable for other species such as whiting (*Merlangius merlangus*) and haddock (*Melanogrammus aeglefinus*). Methodologies and diagnostic criteria involved in the monitoring of contaminant-specific macroscopic liver neoplasms (= liver nodules) and liver histopathology have largely been developed based on experiences with flatfish species (in Europe mainly dab and flounder) but can also be adapted to other flatfish species and also to bottom-dwelling roundfish species.

Status of quality assurance. Quality assurance procedures for externally visible fish diseases, macroscopic liver neoplasms and liver histopathology are in place and operational through ICES activities and under BEQUALM. Largely through activities of the International Council for the Exploration of the Sea (ICES), standardized methodologies for surveys on the occurrence of diseases of flatfish species from the North Sea and adjacent areas have been developed and intercalibrated repeatedly. Practical guidelines have been established for all methodologies involved, including sampling of fish, diagnosis of diseases, reporting of data to ICES and statistical data analysis. As part of the work carried out in BEQUALM, these guidelines were reviewed and, where necessary, additional details and methodologies for the collection, diagnosis and reporting of fish disease data are provided. Under BEQUALM, a number of ring-tests and intercalibration workshops were held. ICES TIMES series publications have been published (nos. 19 and 38).

Influence of environmental variables. Justification is provided that externally visible diseases provide an appropriate indicator of the general health of individuals and populations. The factors that affect disease are multifactorial and include endogenous and exogenous effects on the immune response of the fish as well as specific and non-specific contaminant-related effects at differing biological levels of organization. Certain types of non-neoplastic and neoplastic liver lesions (as specified in the guidelines for the JAMP/CEMP) are known to be associated with prior exposure to carcinogenic contaminants such as PAHs.

Assessment of thresholds. For externally visible diseases and non-specific liver histopathology, absolute threshold or background levels have not been defined due to the

[3] Centre for Environment, Fisheries and Aquaculture Science, Weymouth Laboratory, Barrack Road, The Nothe, Weymouth, Dorset. DT4 8UB. UK.

[4] Federal Research Centre for Fisheries, Institute of Fishery Ecology, Deichstrasse 12, D-27472 Cuxhaven, Germany.

natural variability of disease prevalences between regions. However, significant changes in disease prevalence levels and trends are a basis for threshold assessments. For macroscopic liver neoplasms and contaminant-specific liver histopathology, assessment criteria have been proposed by the 2009 ICES/OSPAR Workshop on Assessment Criteria for Biological Effects Measurements (WKIMC).

Proposals for assessment tools. The WGPDMO developed a Fish Disease Index (FDI) to be used for the analysis and assessment of fish disease data. At the 2009 ICES/OSPAR Workshop on Assessment Criteria for Biological Effects Measurements (WKIMC), assessment criteria for macroscopic liver neoplasms and for contaminant-specific liver histopathology were proposed.

Final remarks. Some amendments should be made to the JAMPGuidelines for PAH-specific biological effects monitoring related to liver histopathology.

Assessment of the applicability of fish disease and liver pathology techniques across the OSPAR maritime area

Diseases of wild marine fish have been studied on a regular basis by many ICES Member Countries for more than two decades. Disease surveys are often integrated with other types of biological and chemical investigations as part of national monitoring programmes aiming at an assessment of the health of the marine environment, in particular in relation to the impact of human activities (Lang, 2002).

On an international level, fish disease data have been used for environmental assessments in the framework of the North Sea Task Force and its Quality Status Report (North Sea Task Force, 1993), the OSPAR Quality Status Report 2000 (OSPAR Commission, 2000) and in the 3rd and 4th HELCOM assessments (HELCOM, 1996, 2002). Studies on externally visible diseases, macroscopic liver neoplasms (= liver nodules) and liver histopathology are on the list of techniques for general and contaminant-specific biological effects monitoring as part of the OSPAR pre-CEMP (see Tables 1 and 4).

At present, annual or biannual fish disease surveys in the North Sea are carried out by Germany (vTI, Inst. of Fishery Ecology, Cuxhaven), The Netherlands (RIKZ) and the UK (Cefas, Weymouth; Marine Scotland, Aberdeen). However, more data are available from monitoring programmes that were terminated in the 1990s or early 2000s (e.g. carried out by Belgium, Denmark and Sweden).

The following environmental monitoring programmes incorporating pathology and diseases of marine organisms are routinely performed in the OSPAR area:

Germany: Surveys are carried out twice a year in offshore areas of the North Sea and the southwestern Baltic Sea. The major target fish species in the North Sea is dab (*Limanda limanda*), in the Baltic Sea flounder (*Platichthys* flesus) and cod (*Gadus morhua*). Externally visible diseases/parasites and liver anomalies (macroscopic and histopathological) are recorded according to ICES guidelines. The data are submitted to the ICES Data Centre.

The Netherlands: Diseases surveys are done annually in three North Sea offshore areas, sites in the western Wadden Sea and in coastal zone of the Eastern Scheldt with dab and flounder as target species. Externally visible diseases/parasites and liver anomalies (macroscopic and histopathological) are recorded according to ICES guidelines. The data are submitted to the ICES Data Centre.

UK: The UK National Marine Monitoring Programme (NMMP) was established to detect long-term trends in physical, biological and chemical variables at selected estuarine and coastal sites in the North Sea, Irish Sea and the English Channel. 10–15 offshore areas are included. The biological effect component of this programme includes assessment of the disease status of target flatfish species (dab and flounder). In addition, data on diseases and parasites in commercial species are also collected. Estuarine monitoring activities have been undertaken more recently using flounder and viviparous blenny (*Zoarces viviparus*) as the target species. In Scotland, externally visible diseases/parasites and liver anomalies of dab, cod and haddock (*Melanogrammus aeglefinus*) are monitored at sampling sites in the Firth of Forth, east of Orkney and in the Moray Firth. Diseases are recorded according to ICES guidelines and the data are submitted to the ICES Data Centre.

Many of these national programmes have increasingly evolved into integrated monitoring programmes, including studies on chemical contamination and on biological effects of contaminants.

Externally visible disease studies are being conducted in a variety of fish species, including dab (*Limanda limanda*), flounder (*Platichthys flesus*) and cod (*Gadus morhua*) and methodologies are easily adaptable for other species such as whiting (*Merlangius merlangus*) and haddock (*Melanogrammus aeglefinus*). Methodologies and diagnostic criteria involved in the monitoring of contaminant-specific liver neoplasms and liver histopathology have largely been developed based on studies with flatfish species, in Europe mainly dab and flounder, but can also be adapted to other flatfish species (e.g. plaice (*Pleuronectes platessa*) or long rough dab (*Hippoglossoides platessoides*)) and possibly also to bottom-dwelling roundfish species, such as dragonet species (*Callionymus* spp.) or viviparous blenny (*Zoarces viviparus*).

In conclusion it can be stated that fish disease and liver histopathology techniques are applicable across the OSPAR maritime area. The application of the Fish Disease Index (FDI) facilitates a comparison of disease data over larger geographical areas and between species (see Chapter 'Proposals for assessment tools').

Status of quality assurance techniques for fish diseases and liver pathology

Since the early 1980s, ICES has played a leading role in the initiation and coordination of fish disease surveys and has contributed considerably to the development of standardized methodologies. Through the work of the ICES Working Group on Pathology and Diseases of Marine Organisms (WGPDMO), its offspring, the Sub-Group/Study Group on Statistical Analysis of Fish Disease Data in Marine Stocks (SGFDDS) (1992–1994) and the ICES Secretariat, quality assurance procedures have been implemented at all stages, from sampling of fish to submission of data to the ICES Data Centre and to data assessment.

A number of practical ICES sea-going workshops on board research vessels were organized by WGPDMO in 1984 (southern North Sea), 1988 (Kattegat), 1994 (Baltic Sea, co-sponsored by the Baltic Marine Biologists, BMB) and 2005 (Baltic Sea) in order to intercalibrate and standardize methodologies for fish disease surveys (Dethlefsen *et al.*, 1986; ICES, 1989, 2006a; Lang and Mellergaard, 1999) and to prepare guidelines. Whilst first guidelines were focused on externally visible diseases and parasites, WGPDMO developed guidelines for macroscopic and microscopic inspection of flat-

fish livers for the occurrence of neoplastic lesions at a later stage. Further intercalibration and standardization of methodologies used for studies on liver pathology of flatfish were a major issue of the 1996 ICES Special Meeting on the Use of Liver Pathology of Flatfish for Monitoring Biological Effects of Contaminants (ICES, 1997). This formed the basis from which the BEQUALM programme developed for the application of liver pathology in biological effects monitoring (Feist *et al*., 2004) (Table 2).

A fish disease database has been established within the ICES Data Centre, consisting of disease prevalence data of key fish species and accompanying information, submitted by ICES Member Countries. Submission of fish disease data to the ICES Marine Data Centre has been formalized by the introduction of the ICES Environmental Reporting Format designed specifically for the purpose. This is used for fish disease, contaminant and biological effects data. The programme includes internal screening procedures for the validation of the data submitted providing further quality assurance.

The ICES fish disease database is extended on an annual basis to include data from other species and areas within the OPSPAR maritime area as well as data on studies into other types of diseases, e.g. macroscopic liver neoplasms and liver histopathology. To date, the data comprise mainly information from studies on the occurrence of externally visible diseases and macroscopic liver lesions in the common dab (*Limanda limanda*) and the European flounder (*Platichthys flesus*) from the North Sea and adjacent areas, including the Baltic Sea, Irish Sea, and the English Channel. In addition, reference data are available from pristine areas, such as waters around Iceland. In total, data on length, sex, and health status of more than 700 000 individual specimens, some from as early as 1981, have been submitted to ICES, as well as information on sampling characteristics (Wosniok *et al*., 1999, Lang and Wosniok, 2008).

Current ICES WGPDMO activities have focused on the development and application of statistical techniques for an assessment of disease data with regard to the presence of spatial and temporal trends in the North Sea and western Baltic Sea (Wosniok *et al*., 1999; Lang and Wosniok, 2008). An output of WGPDMO's activities is the ICES web-based report on wild fish diseases, consisting of trend maps and associated information. In a more holistic approach, pilot analyses have been carried out combining the disease data with oceanographic, nutrient, contaminant and fishery data extracted from the ICES DataCentre in order to improve the knowledge of the complex cause–effect relationships between environmental factors and fish diseases (Lang and Wosniok, 2000; Wosniok *et al*., 2000). These analyses constituted one of the first attempts to combine and analyses ICES data from various sources and can, therefore, be considered as a step towards a more comprehensive integrated assessment.

Quality assurance is in place for externally visible diseases, macroscopic liver neoplasms and liver histopathology via the ongoing BEQUALM programme (additional information under 'Assessment of thresholds' below). Regular intercalibration and ring-test exercises are conducted. The basis for QA procedures are provided in two key publications in the ICES TIMES series (Bucke *et al*., 1996, Feist *et al*., 2004) and a BEQUALM CD ROM of protocols and diagnostic criteria and reporting requirements for submission of data to ICES.

Review of the environmental variables that influence fish diseases and liver pathology

The multifactorial aetiology of diseases, in this context in particular of externally visible diseases, is generally accepted. Therefore, externally visible diseases have correctly been placed into the General biological effect component of the OSPAR CEMP. Most wild fish diseases monitored in past decades are caused by pathogens (viruses, bacteria). However, other endogenous or exogenous factors may be required before the disease develops. One of these factors can be environmental pollution, which may either affect the immune system of the fish in a way that increases its susceptibility to disease, or may alter the number and virulence of pathogens. In addition, contaminants may also cause specific and/or non-specific changes at various levels of biological organization (molecule, subcellular units, cells, tissues, organs) leading to disease without involving pathogens.

The occurrence of significant changes in the prevalence of externally visible fish diseases can be considered a non-specific and more general indicator of chronic rather than acute (environmental) stress, and it has been speculated that they might, therefore, be an integrative indicator of the complex changes typically occurring under field conditions rather than a specific marker of effects of single factors. Because of the multifactorial causes of externally visible diseases, the identification of single factors responsible for observed changes in disease prevalence is difficult, and scientific proof of a link between contaminants and externally visible fish diseases is hard to achieve. Nevertheless, there is a consensus that fish disease surveys should continue to be part of national and international environmental monitoring programmes because they can provide valuable information on changes in ecosystem health and may act as an "alarm bell" potentially initiating further more specific studies on cause and effect relationships.

In the statistical analysis of ICES data on externally visible diseases (lymphocystis, epidermal hyperplasia/papilloma, acute/healing skin ulceration) of dab from different North Sea regions, it could be demonstrated that there were significant spatial differences, both in terms of absolute levels and the temporal changes in disease prevalence in the North Sea. While data from the 1990s revealed stable or decreasing disease prevalences in the majority of sampling sites, some areas in the North Sea showed increasing trends for some of the diseases, indicating a change in environmental conditions adversely affecting the health status of dab (Wosniok *et al*., 1999). The results from the subsequent multivariate analysis on the relationship between the prevalence of the diseases with potentially explanatory environmental and host-specific factors (also extracted from the ICES fishery, oceanography and environmental databases) clearly highlighted the multifactorial aetiology of the diseases under study. A number of natural and anthropogenic factors (stock composition, water temperature, salinity, nutrients, contaminants in water, sediments and biota) were significantly related to the temporal changes in disease prevalence. However, depending on area, time range and data availability, different sets of factors were identified. This reflects the multifactorial aetiology of the diseases covered, but was also attributed to some high correlations between the explaining quantities (Lang and Wosniok, 2000; Wosniok *et al*., 2000).

The presence of macroscopic liver neoplasms and of certain types of histopathological liver lesions is a more direct indicator of contaminant effect and has been used for many years in environmental monitoring programmes around the world. Liver neoplasms (either detected macroscopically or by histopathological analysis) are likely to

be associated to exposure to carcinogenic contaminants, including PAHs, and are therefore considered appropriate indicators for General and for PAH-specific biological effects monitoring. Therefore, monitoring of macroscopic liver neoplasms in the CEMP should not only be part of the CEMP general biological effects monitoring but also of the CEMP PAH-specific biological effects monitoring. The study of liver histopathology (comprises the detection of more lesion categories (non-specific, neoplastic and non-neoplastic toxicopathic lesions), reflecting responses to a wider range of contaminants (including PAHs) but also to other environmental stressors and is, therefore, considered an appropriate indicator for both General and PAH-specific biological effects monitoring.

The liver is the main organ involved in the detoxification of xenobiotics and several categories of hepatocellular pathology are now regarded as reliable biomarkers of toxic injury and representative of biological endpoints of contaminant exposure (Myers *et al*., 1987; 1992; 1998; Stein *et al*., 1990; Vethaak and Wester, 1996; Stentiford *et al*., 2003; Feist *et al*., 2004). The majority of lesions observed in field collected animals have also been induced experimentally in a variety of fish species exposed to carcinogenic compounds, PAHs in particular, providing strong supporting evidence that wild fish exhibiting these lesions could have been exposed to such environmental contaminants.

Assessment of the thresholds when the response (prevalence and incidence of fish disease) can be considered to be of concern and/or require a response

As indicated above, ICES has developed requirements for the international reporting of fish diseases over many years in order to minimize variation between laboratories regarding the accuracy and reproducibility of data generated. These have been reviewed by BEQUALM and produced in CD-ROM format. Each grossly visible disease (lymphocystis, acute and healing skin ulcerations, epidermal hyperplasia/papilloma and liver nodules etc.) has a minimum requirement for reporting and severity is assessed according to criteria allocated to three stages (lymphocystis, ulcerations and epidermal hyperplasia/papilloma only). Macroscopic liver neoplasms are only recorded if the minimum diameter exceeds two mm. Each case has to be verified histologically to exclude the possibility that the macroscopic lesion is the response to parasites, cysts, necrotic or inflammatory foci. As such the acceptable limits of variation for disease recording are well established.

With regard to the application of liver histopathology as a tool in biological effects monitoring, the activities undertaken in ICES and within BEQUALM have been successful in the establishment of the methodology and diagnostic criteria. The diagnostic key (see below) provides clear criteria to discriminate between the lesion types, thus minimizing the possibility of mis-diagnosis. Ring tests and other intercalibration exercises are regularly undertaken in order to minimize inter-observer variation and to establish acceptable limits of variation. These are carried out as an ongoing process in order to ensure continuous quality assurance of data obtained.

These quality assurance procedures implemented are a crucial prerequisite for the establishment of assessment criteria (see below) and reference or threshold values applied by all institutions involved in fish disease monitoring in order to take decisions on further actions. The ICES WGPDMO and the 2009 ICES/OSPAR Workshop on Assessment Criteria for Biological Effects Measurements (WKIMC) addressed the

question of establishing background/reference levels of disease and criteria for their assessment (see Chapter 'Proposals for assessment tools').

Proposals for assessment tools

The development of assessment tools for externally visible diseases, macroscopic neoplasms and liver histopathology has been carried out by the ICES Working Group on Pathology and Diseases of Marine Organisms (WGPDMO) (ICES 2006b, 2007, 2008, 2009) and additions were proposed at the 2009 ICES/OSPAR Workshop on Assessment Criteria for Biological Effects Measurements (WKIMC).

The ICES WGPDMO developed a Fish Disease Index (FDI) using data on diseases of the common dab (*Limanda limanda*) as a model, the aim of which is to summarize information on the disease status of individual fish into one robust and easy-to-understand and easy-to-communicate numeric figure. By applying defined assessment criteria and appropriate statistics, the FDI can be used to assess temporal changes in the health status of fish populations and can, thus, serve as a tool for the assessment of the ecosystem health of the marine environment, e.g. related to the effects of anthropogenic and natural stressors. Its design principle allows the FDI to be applied to other species with other sets of diseases. Therefore, the FDI approach is applicable for wider geographical areas, e.g. as part of the convention-wide OSPAR monitoring and assessment programme.

For the calculation of the FDI, the following components are required:

- Data on diseases of the common dab (*Limanda limanda*) (can be adapted to other fish species, provided that sufficient appropriate data are available);
- Information on the presence or absence of a range of diseases monitored on a regular basis, categorized as externally visible diseases (EVD: 9 key diseases, incl. 3 parasites), macroscopic liver neoplasms (MLN: 2 key diseases) and liver histopathology (LH: 5 key diseases) (see Table 1);
- For most diseases, data on three severity grades (reflecting a light, medium or severe disease status) are included;
- Disease-specific weighting factors, reflecting the impact of the diseases on the host (assigned based on expert judgements);
- Adjustment factors for effects of size and sex of the fish as well as for season effects;

Table 1. Disease categories and key diseases to be used for calculating the Fish Disease Index for dab (*Limanda limanda*) (ICES 2009).

Externally visible diseases	Macroscopic liver neoplasms	Liver histopathology
Lymphocystis	Benign neoplasms	Non-specific lesions
Epidermal hyperplasia/papilloma	Malignant neoplasms	Early non-neoplastic toxicopathic lesions
Acute/healing skin ulceration		Pre-neoplastic lesions (FCA)
X-cell gill disease		Benign neoplasms
Hyperpigmentation		Malignant neoplasms
Acute/healing fin rot/erosion		
Stephanostomum baccatum		
Acanthochondria cornuta		
Lepeophtheirus pectoralis		

The result of the calculation is a FDI value for individual fish which is scaled in a way that values can range from 0 to 100, with low values representing healthy and high values representing diseased fish. The maximum value of 100 can only be reached in

the (purely theoretical and unrealistic) case that a fish is affected by all diseases at their highest severity grades. From the individual FDIs, mean FDIs for a sample from a fish population in a given sampling area can be calculated. All assessment is based on mean FDI values calculated from samples. Depending on the data available, FDIs can be calculated either for single disease categories or for combinations thereof.

The assessment of the FDI data generated is done on a region-wide basis, considering (a) FDI levels and (b) FDI trends in geographical units, e.g. ICES statistical rectangles. The assessment approach developed does not apply any global background or reference values or assessment criteria as is often done for chemical contaminants or for biochemical biomarkers. Instead, the assessment of the FDI values is based on the development of the mean FDI within the geographical units over a given period of time, based on which region-specific assessment criteria are defined. The reason for choosing this approach is the known natural regional variability of the disease prevalence (even in areas considered to be pristine), making it impossible to define generally applicable background/reference values that can uniformly be used for all geographical units to be assessed. This approach is based on the availability of disease data over a longer period of time (ideally ten observations, e.g. in the case of biannual monitoring over a period of five years) for every geographical area to be assessed. However, the FDI can also be used for exploratory monitoring in areas not studied before or for newly installed fish disease monitoring programmes after some modification.

The final products of the assessment procedure are:

- graphs showing the temporal changes in mean FDI values in geographical units over the entire observation period and
- maps in which the geographical units assessed are marked with green, yellow or red smiley faces, indicating current changes (e.g. within the past five years) in health status of the fish population (green: improvement of the health status; yellow, indifferent change; red: worsening of the health status, reason for concern and motivation for further research on causes) (see Figures 3 and 4).

The ICES WGPDMO applied the FDI approach and the assessment for the common dab from the North Sea using ICES fish disease data extracted from the ICES Environmental Data Centre twice in 2008 and, using an extended dataset, in 2009 (ICES 2008, 2009). The results will be included in the OSPAR QSR 2010 as a case study.

At the 2009 ICES/OSPAR Workshop on Assessment Criteria for Biological Effects Measurements (WKIMC), additional assessment criteria for macroscopic liver neoplasms and for the contaminant-specific components of liver histopathology were proposed. These are provided in Table 3 together with information on the use of assessment criteria for externally visible diseases and for the non-specific component of liver histopathology. The latter two are identical with the assessment strategy of the original FDI approach developed by the WGPDMO.

Final remarks

Some amendments still need to be made by OSPAR in the JAMP Guidelines for General and for PAH-specific biological effects monitoring and the terminology used therein:

- In the JAMP Guidelines for PAH-specific biological effects monitoring, Chapter 4.1 and 5, the term ‘Liver pathology’ should be changed to ‘Liver

histopathology' and the term 'external diseases' should be changed to 'externally visible diseases' because these terms more correctly describe the technique to be applied.

- In the table of contents of the JAMP Guidelines for PAH-specific biological effects monitoring, the terms 'histopathology' and 'liver pathology' should be replaced by 'liver histopathology' because this term more correctly describes the technique to be applied.

Literature cited

Bucke, D., Vethaak, A. D., Lang, T., and Mellergaard, S. 1996. Common diseases and parasites of fish in the North Atlantic: Training guide for identification. ICES Techniques in Marine Environmental Sciences, 19. 27 pp.

Dethlefsen, V., Egidius, E., and McVicar, A. H. 1986. Methodology of fish disease surveys. Report of an ICES Sea-going Workshop held on RV 'Anton Dohrn' 3-12 January 1984. ICES Cooperative Research Report, 140. 33 pp.

Feist, S. W., Lang, T., Stentiford, G. D. and Köhler, A. 2004. The use of liver pathology of the European flatfish, dab (Limanda limanda L.) and flounder (Platichthys flesus L.) for monitoring biological effects of contaminants. ICES Techniques in Marine Environmental Science, 28: 47pp.

HELCOM. 1996. Third Periodic Assessment of the State of the Marine Environment of the Baltic Sea, 1989-1993; Background Document. Baltic Sea Environment Proceedings N0 64 B. 252 pp.

HELCOM. 2002. Environment of the Baltic Sea area 1994-1998. Baltic Sea Environment Proceedings N0 82 B. 215 pp.

ICES. 1989. Methodology of fish disease surveys. Report of an ICES Sea-going Workshop held on RV U/F 'Argos' 16–23 April 1988. ICES Cooperative Research Report, 166. 33 pp.

ICES. 1997. Report of the Special Meeting on the Use of Liver Pathology of Flatfish for Monitoring Biological Effects of Contaminants. ICES CM 1997/F:2. 75 pp.

ICES. 2005. ICES Advice. http://www.ices.dk/committe/acfm/comwork/report/2005/may/Technical%20Annex%20to%20the%20review%20of%20WKIMON.pdf

ICES. 2006. Report of the ICES/BSRP Sea-going Workshop on Fish Disease Monitoring in the Baltic Sea (WKFDM). ICES CM 2006, BCC:02, 85 pp.

ICES. 2006. Report of the Working Group on Pathology and Diseases of Marine Organisms (WGPDMO), ICES CM 2006/MCC:01, 104 pp.

ICES. 2007. Report of the Working Group on Pathology and Diseases of Marine Organisms (WGPDMO), ICES CM 2007/MCC:04, 92 pp.

ICES. 2008. Report of the Working Group on Pathology and Diseases of Marine Organisms (WGPDMO), ICES CM 2008/MCC:01, 128 pp.

ICES. 2009. Report of the Working Group on Pathology and Diseases of Marine Organisms (WGPDMO), ICES CM 2009/MCC:01, 111 pp.

Lang, T. 2002. Fish disease surveys in environmental monitoring: the role of ICES. ICES Marine Science Symposia, 215: 202–212.

Lang, T., and Mellergaard, S. 1999. The BMB/ICES Sea-Going Workshop 'Fish Diseases and Parasites in the Baltic Sea' – introduction and conclusions. ICES Journal of Marine Science, 56: 129–133.

Lang, T., and Wosniok, W. 2000. ICES data for an holistic analysis of fish disease prevalence. In ICES 2000, Report of the ICES Advisory Committee on the Marine Environment, 1999, pp. 193-209. ICES Cooperative Research Report, 239. 277 pp.

Lang, T. and Wosniok, W. 2008. The Fish Disease Index: a method to assess wild fish disease data in the context of marine environmental monitoring. ICES CM 2008/D:01, 13 pp.

Myers, M. S., Rhodes, L. D. and McCain, B.B. 1987. Pathologic anatomy and patterns of occurrence of hepatic neoplasms, putative preneoplastic lesions, and other idiopathic conditions in English sole (*Parophyrs vetulus*) from Puget Sound, Washington. Journal of the National Cancer Institute 78, 333–363.

Myers, M.S., Olson, O.P., Johnson, L.L., Stehr, C.M., Hom, T., and Varanasi, U. 1992. Hepatic lesions other than neoplasms in subadult flatfish from Puget Sound, WA: Relationships with indices of contaminant exposure. Mar. Environ. Res. 34, 45–51.

Myers, M.S., Johnson, L.L., Hom, T., Collier, T.K., Stein, J.E., and Varanasi, U. 1998. Toxicopathic lesions in subadult English sole (*Pleuronectes vetulus*) from Puget Sound, Washington, USA: Relationships with other biomarkers of contaminant exposure. Marine Environmental Perspectives, 45, 47–67.

North Sea Task Force. 1993. North Sea Quality Status Report 1993. Oslo and Paris Commissions, London. Olsen and Olsen, Fredensborg, Denmark. 132 pp.

OSPAR Commission. 2000. Quality Status Report 2000, Region II-Greater North Sea. OSPAR Commission, London. 136+xiii pp.

Stein, J.E., Reichert, W.L., Nishimito, M., and Varanasi, U. 1990. Overview of studies on liver carcinogenesis in English sole from Puget Sound; evidence of a xenobiotic chemical etiology, 2, Biochemical studies. Sci. Total Environ. 94, 51–69.

Stentiford, G. D., Longshaw, M., Lyons, B. P., Jones, G., Green, M. and Feist, S. W. 2003 Histopathological biomarkers in estuarine fish species for the assessment of biological effects of contaminants. Marine Environmental Research 55, 137–159.

Vethaak, A. D. and Wester, P. W. 1996. Diseases of flounder *Platichthys flesus* in Dutch coastal and estuarine waters, with particular reference to environmental stress factors. II. Liver histopathology. Dis. Aquat. Orgs., 26, 81–97.

Wosniok, W., Lang, T., Dethlefsen, V., Feist, S. W., McVicar, A. H., Mellergaard, S., and Vethaak, A. D. 2000. Analysis of ICES long-term data on diseases of North Sea dab (*Limanda limanda*) in relation to contaminants and other environmental factors. ICES CM 2000/S:12. 15 pp.

Wosniok, W., Lang, T., Vethaak, A. D., des Clers, S., and Mellergaard, S. 1999. Statistical analysis of fish disease prevalence data from the ICES Environmental Data Centre. In Report of the ICES Advisory Committee on the Marine Environment, 1998, pp. 297–327. ICES Cooperative Research Report, 233. 375 pp.

Table 2. BEQUALM categories of histopathological liver lesions in fish that should be used for the CEMP General and PAH-specific Biological Effects Monitoring.

Histopathology Categories	Histopathological Lesions
Non-specific lesions	Coagulative necrosis
	Apoptosis
	Lipoidosis
	Haemosiderosis
	Variable glycogen content
	Increased numbers and size of macrophage aggregates
	Lymphocytic/monocytic infiltration
	Granuloma
	Fibrosis
	Regeneration

Early toxicopathic non-neoplastic lesions	Phospholipidosis Fibrillar inclusion Hepatocellular and nuclear polymorphism Hydropic degeneration Spongiosis hepatis
Foci of cellular alteration	Clear cell foci Vacuolated foci Eosinophilic foci Basophilic foci Mixed cell foci
Benign neoplasms	Hepatocellular adenoma Cholangioma Haemangioma Pancreatic acinar cell adenoma
Malignant neoplasms	Hepatocellular carcinoma Cholangiocarcinoma Pancreatic acinar cell carcinoma Mixed hepatobiliary carcinoma Haemangiosarcoma Haemangiopericytic sarcoma

Table 3. Assessment criteria proposed for the assessment of contaminant-specific effects on fish health (Note: the colour 'red' should be used for graphical representations of the categories 'elevated response/above background' as well as for 'significant response/unacceptable effects' in maps or similar illustrations).

Disease category	Background	Elevated response/ above background	Significant response/ unacceptable effects
Externally visible diseases (to be used as additional information for the assessment)	Not applicable	Statistically significant increase in mean FDI level in the assessment period compared with a prior observation period or Statistically significant upward trend in mean FDI level in the assessment period	Statistically significant increase in mean FDI level in the assessment period compared with a prior observation period or Statistically significant upward trend in mean FDI level in the assessment period
Liver histopathology: non-specific (to be used as additional information for the assessment)	Not applicable	Statistically significant increase in mean FDI level in the assessment period compared with a prior observation period or Statistically significant upward trend in mean FDI level in the assessment period	Statistically significant increase in mean FDI level in the assessment period compared with a prior observation period or Statistically significant upward trend in mean FDI level in the assessment period
Liver histopathology: contaminant-specific	Mean FDI <2	Mean FDI ≥ 2 A value of FDI = 2 is, e. g., reached if the prevalence of liver tumours is 2 % (e. g., one specimen out of a sample of 50 specimens is affected by a liver tumour). Levels of FDI ≥ 2 can be reached if more fish are affected or if combinations of other toxicopathic lesions occur.	Mean FDI ≥2 A value of FDI = 2 is, e. g., reached if the prevalence of liver tumours is 2 % (e. g., one specimen out of a sample of 50 specimens is affected by a liver tumour). Levels of FDI ≥ 2 can be reached if more fish are affected or if combinations of other toxicopathic lesions occur.
Macroscopic liver neoplasms	Mean FDI <2	Mean FDI ≥ 2 A value of FDI = 2 is reached if the prevalence of liver tumours (benign or malignant) is 2 % (e. g., one specimen out of a sample of 50 specimens is affected by a liver tumour). If more fish are affected, the value is FDI > 2.	Mean FDI ≥2 A value of FDI = 2 is reached if the prevalence of liver tumours (benign or malignant) is 2 % (e. g., one specimen out of a sample of 50 specimens is affected by a liver tumour). If more fish are affected, the value is FDI > 2.

Table 4. Fish disease monitoring in the OSPAR Coordinated Environmental Monitoring Programme (CEMP) reflecting ICES advice (ICES 2005).

TABLE A: PAH-SPECIFIC BIOLOGICAL EFFECTS MONITORING				
	SPECIES	DISEASES	NUMBERS	GUIDELINES
Macroscopic liver neoplasms	Dab (1st priority) (*Limanda limanda*)	Macroscopic liver nodules > 2 mm in diameter, subsequent quantification of histologically identified liver neoplasms	Size group ≥ 25 cm: 50 (if not available in sufficient numbers, include size group 20–24 cm)	JAMP Guidelines based on: Bucke et al. 1996. Common diseases and parasites of fish in the North Atlantic: Training guide for identification. ICES TIMES No. 19. Relevant in addition: ICES 1989. Methodology of fish disease surveys. ICES Coop. Res. Rep. 166. Feist et al. 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 TIMES 38, 42 pp. BEQUALM
	Flounder (*Platichthys flesus*)		Size group ≥ 30 cm: 50 (if not available in sufficient numbers, include size group 25–29 cm)	
Liver histopathology	Dab (1st priority) (*Limanda limanda*)	Non-specific lesions Early toxicopathic non-neoplastic lesions Foci of cellular alteration Benign neoplasms Malignant neoplasms	Size group 20–24 cm : 50	JAMP Guidelines based on: ICES 1997. Report of the Special Meeting on the Use of Liver Pathology of Flatfish for Monitoring Biological Effects of Contaminants. ICES CM 1997/F:2. Relevant in addition: Feist et al. 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 TIMES 38, 42 pp. BEQUALM
	Flounder (*Platichthys flesus*)		Size group 25–29 cm : 50	
	Dragonet (*Callionymus* spp.)		Size group 10–15 cm : 50	No JAMP guidelines so far Relevant: Feist *et al.*, 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 TIMES 38, 42 pp.

Table B: General biological effects monitoring

	Species	Diseases	Numbers	Guidelines
Externally visible fish diseases	Dab (1st priority) (*Limanda limanda*)	Lymphocystis Epidermal hyperplasia/papilloma Acute/healing skin ulcers X-cell gill disease Hyperpigmentation	Size group 15–19 cm: 100 Size group 20–24 cm: 100 Size group ≥ 25 cm : 50	JAMP Guidelines based on: Bucke et al. 1996. Common diseases and parasites of fish in the North Atlantic: Training guide for identification. ICES TIMES No. 19. Relevant in addition: ICES 1989. Methodology of fish disease surveys. ICES Coop. Res. Rep. 166. BEQUALM
	Flounder (*Platichthys flesus*)	Lymphocystis Acute/healing skin ulcers	Size group 20–24 cm: 100 Size group 25–29 cm: 100 Size group ≥ 30 cm: 50	
	Cod (*Gadus morhua*)	Acute/healing skin ulcers Skeletal deformities Pseudobranchial swelling *Cryptocotyle* sp.	Size group < 29 cm: 100 Size group 30–44 cm: 100 Size group ≥ 45 cm: 50	
	Whiting (*Merlangius merlangus*)	Epidermal hyperplasia/papilloma *Lernaeocera branchialis* *Diclidophora merlangi* *Clavella adunca*	Size group 15–19: 100 Size group 20–29: 100 Size group ≥ 30: 50	No JAMP guidelines so far Relevant: Bucke *et al.*, 1996. Common diseases and parasites of fish in the North Atlantic: Training guide for identification. ICES TIMES No. 19.
Macroscopic liver neoplasms	Dab (1st priority) (*Limanda limanda*)	Macroscopic liver nodules > 2 mm in diameter, subsequent quantification of histologically identified liver neoplasms	Size group ≥ 25 cm: 50 (if not available in sufficient numbers, include size group 20–24 cm)	JAMP Guidelines based on: Bucke *et al.*, 1996. Common diseases and parasites of fish in the North Atlantic: Training guide for identification. ICES TIMES No. 19. Relevant in addition: ICES 1989. Methodology of fish disease surveys. ICES Coop. Res. Rep. 166. Feist *et al.*, 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 TIMES 38, 42 pp. BEQUALM
	Flounder (*Platichthys flesus*)		Size group ≥ 30 cm: 50 (if not available in sufficient numbers, include size group 25–29 cm)	
Liver histopathology	Dab (1st priority) (*Limanda limanda*)	Non-specific lesions Early toxicopathic non-neoplastic lesions	Size group 20–24 cm: 50	JAMP Guidelines based on:

Table B: General biological effects monitoring

Species	Diseases	Numbers	Guidelines
Flounder (*Platichthys flesus*)	Foci of cellular alteration Benign neoplasms Malignant neoplasms	Size group 25–29 cm: 50	Bucke *et al.*, 1996. Common diseases and parasites of fish in the North Atlantic: Training guide for identification. ICES TIMES No. 19. Relevant in addition: ICES 1989. Methodology of fish disease surveys. ICES Coop. Res. Rep. 166. Feist *et al.*, 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 TIMES 38, 42 pp. BEQUALM
Dragonet (*Callionymus* spp.)		Size group 10–15 cm : 50	No JAMP guidelines so far for Dragonet Relevant: Feist *et al.*, 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 TIMES 38, 42 pp.

Annex 9: Water *in vivo* bioassays

Executive summary

Applicability across the OSPAR maritime area. Water *in vivo* bioassays are available for immediate deployment within the OSPAR JAMP CEMP. These bioassays have been recommended by ICES and are of sufficient standing in terms of methodological development, ease of use and application for uptake across the whole OSPAR area. The preferred method is short-term tests on concentrates of water. This includes both broad-spectrum (acute and short-term chronic) bioassays, (and can be combined with specific *in vitro* bioassays), which can be applied to salt water, brackish water and fresh water, allowing all types of water to be assessed in the same way, and thereby giving a comprehensive picture of an entire area. If the focus is also on specific groups of substances or a specific toxicity, such as hormone-disrupting effects or neurotoxicity, *in vitro* bioassays can be used, on concentrates or otherwise. Chronic (long-term) *in vivo* bioassays would appear to be most suited to site-specific assessment and comparison with the field situation (e.g. to provide sufficient evidence to support the conclusion that a problem no longer occurs). The long-term exposure without concentration of the sample means these tests give the most realistic estimate of the possible effects in the field. Relevant acute bioassays can be a quick and cheap alternative, as can *in vitro* tests.

Water bioassays should be deployed as a "battery of tests" and should include a minimum basic set, possibly of three or more. However, the composition of what the set needs to comprise of requires further work. The range of bioassays needs to be expanded to include all trophic levels and phyla such as echinoderms.

Quality assurance. QA procedures are in place for most of the (water) bioassays and is provided for by BEQUALM (www.bequalm.org), therefore bioassay data can be submitted to the ICES database for subsequent assessment as appropriate by ICES/OSPAR. A standardized protocol for bioassay extractions is required to ensure consistency of application between laboratories and member states and comparability of reported data for assessment purposes. A protocol for extraction methods for bioassays will become available as ICES TIMES series document in 2010.

Influence of environmental variables. Abiotic testing conditions, such as temperature, salinity, solids and especially dissolved oxygen and pH, can dramatically influence test variability. The same is true for the condition and age of test organisms and storage conditions of test samples. In general, these factors are standardized in the test procedures and controlled during the test period by the use of positive and negative controls. The use of extracts/concentrates will further reduce any disturbing factors.

Thresholds and assessment tools. Three assessment classes were derived for water bioassays; a background response, a warning level and a level of serious concern. For the water bioassays (*Tisbe* sp., *Acartia* sp., sea urchin and bivalve larvae) the background responses were 10%, 10%, 10% and 20% mortality (or deformity as appropriate) respectively; the level of serious concern was 100% mortality, and the warning level between these values. These figures however need to be defined and further established.

In this document we describe and propose an ecotoxicological metric for acute and chronic in vivo bioassays. An acute/chronic ratio of 10 is used to convert the acute data to chronic data. If data are available from three bioassays, a preliminary effect assessment can be performed. If at least four chronic values are available for different

taxonomic groups, a refined effect assessment can be carried out whereby the potentially affected fraction (PAF) approach is used to calculate the percentage of affected species in the ecosystem in question. With its 'negligible effect', 'maximum permissible effect' and 'serious effect' classification, this method assessment is consistent with the current Dutch standard framework and terminology (environmental risk limits). It is however equally suited to the current OSPAR and EU-WFD assessment frameworks.

Synergism between CEMP and WFD. There are clear opportunities for synergism between the CEMP and WFD for water bioassay applications in coastal and estuarine areas, but further work and agreement is needed.

Recommendations. The sampling strategy and design of water quality monitoring for spatial and temporal monitoring purposes needs to be clearly defined and in particular the role of water concentrates. In this respect there is an important need to develop and validate appropriate protocols for extraction methods and subsequent *in vivo* (and *in vitro*) testing. More research is also needed to link bioassay responses to actual impacts on the aquatic system. The application of passive samplers for bioassay assessment of water also warrants special attention. It is recommended that a pilot study be carried out to test the practical application of the proposed metric (or any other available tool) for water in the OSPAR maritime area.

Assessment of the applicability of water *in vivo* bioassays across the OSPAR maritime area

Most existing bioassays have been used for reporting to regulatory commissions on individual hazardous substances and the determination of environmental quality standards (den Besten and Munawar, 2005). Over the past few decades, bioassays have also been used for the risk assessment and management of saline and fresh water whole effluents (e.g. Oris and Klaine, 2000; Power, 2004), and for dredged material (e.g. Stronkhorst *et al*., 2003).

To date, there are numerous studies illustrating the application of bioassays to assess the toxicity of environmental samples from marine and inland surface water (e.g. Karbe, 1992; Hill *et al*., 1993; Matthiesen *et al*., 1993; Hendriks *et al*., 1994; Thomas *et* al., 1999; Kirby *et al*., 1998; Peters *et al*., 2002; Akerman and Smit, 2003; Derksen *et al*., 2004). For example, bioassay assessment of fresh surface water has been used successfully for many years in the Netherlands in the context of the surveillance monitoring of the Meuse, Scheldt and Rhine river basins (Maas *et al*., 2003). This assessment used acute bioassays (or *in vitro* bioassays) (including CALUX systems, Microtox®, Daphnia and whole sediment, pore water) on XAD concentrates of the water (e.g. Hendriks *et al*., 1998; Maas *et al*. 2003). The ICES/IOC Bremerhaven Workshop on biological effects of contaminants in the North Sea and the ICES BECPELAG Workshop on biological effects in pelagic ecosystems have clearly demonstrated the potential applicability of a variety of *in vivo* bioassays to coastal and offshore water column and micro surface layer monitoring (Stebbing *et al*., 1992; Hylland *et al*., 2002, 2006).

Water bioassays recommended for use in different monitoring strategies are well described in OECD, ASTM, ISO, SETAC and ICES test protocols (see also USEPA, 1995; Tonkes *et al*., 2005). Bioassays are widely recognized within Europe to be an efficient way to assess water quality. Bioassays are also applied on national level by several countries (ICES, 2004). The uptake of water bioassays, such as the oyster embryo assay (Thain *et al*., 1991), in monitoring programmes across the OSPAR maritime area is however still poor (so far, only UK; see ICES, 2004). *In vivo* bioassays and *in vitro* tests

with micro-organisms are now also frequently used as tools in estimating the potential risk of contaminants of estuarine and marine waters (e.g. Thomas *et al.*, 2002; Murk *et al.*, 2002; Klamer *et al.*, 2003; Akerman *et al.*, 2004).

Introduction of water *in vivo* bioassays to the CEMP and status of quality assurance

ICES agreed on the following revised criteria for recommended monitoring methods:

a) A recommended method needs to be an established technique that is available as a published method in the TIMES series or elsewhere. This applies to both the bioassay itself and the preparation phase (such as the sampling and extraction methods).

b) A recommended method (or combination of methods) must have been shown to respond to contaminant exposure in the field.

c) A recommended method (or combination of methods) must be able to differentiate the effects of contaminants from natural background variability.

The OSPAR JAMP CEMP lists water bioassays as Category-II-rated. The corresponding Technical Annex 2, 3 and 4 of the JAMP Guidelines for General Biological Effects Monitoring relate to the following bioassay methods: *Tisbe battagliai*, oyster embryo, *Nitocra* and *Dinophilus.* However, other species are now also appropriate and have been recommended by ICES and include the methods; turbot juvenile acute, *Daphnia* acute and chronic, *Acartia* acute, *Skeletonema* 72-h growth).

Quality assurance through BEQUALM is in place or currently running (JAMP, 1998; ASMO, 2003; ICES, 2005). So far, uptake of water bioassays in BEQUALM has been slow but is increasing. Protocols exist for water extracts, but they have not been agreed, standardized and "transcribed" into OSPAR guidelines. A standardized protocol for bioassay extractions is required to ensure consistency of application between laboratories and member states and comparability of reported data for assessment purposes. Also these protocols are used as standard procedures for BEQUALM intercalibrations. The Protocol for Extraction Methods for Bioassays will be published in the *ICES Techniques in Marine Environmental Sciences* series on Biological Effects of Contaminants.

Synergism between CEMP, MSFD and WFD

Though bioassays are not included as ecological quality elements in the monitoring for the Water Framework Directive (WFD) (CIS, 2003), it is generally accepted that they will be able to contribute to the Pressures and Impacts/Risk Assessment process (this is especially true of chronic water and sediment bioassays). This process, being carried out by national authorities, is designed to identify water bodies at risk of failing to achieve good ecological status during the later classification exercise. Further chemical analysis can be combined with water bioassays at smaller interval time points for the purposes of trend monitoring. In this way bioassays can be used as a partial replacement for chemical analysis of priority and/or other relevant substances and prioritizing locations for further chemical analysis. This "bioanalysis approach" can lead to more cost-efficient and cost-effective monitoring and would put the precautionary principle called for in the WFD into practice. Pilot studies carried out in the Netherlands to explore these possibilities have had promising results (van de Heuvel *et al.*, 2005; Maas *et al.*, 2005). It can be concluded that clear opportunities exist

for synergism between the CEMP or the MSFD and WFD for bioassay applications in coastal and estuarine areas, but that further work and agreement are needed.

Thresholds and assessment tools

General

Thresholds for water bioassays are available. Effects measured include acute (e.g. mortality) or chronic endpoints (sub lethal endpoint such as growth, development and reproduction) and hence are generic indicators of toxicity of the water. Values of EC_{xx}, LC_{xx}, NOEC and LOEC are usually used where appropriate to evaluate the test responses and to estimate toxicity. Results of bioassays from a contaminated area can be compared with a reference area, in a dose-response relationship between sites or by using time-series analysis, multivariate analysis such as principal component analysis (PCA), and toxicological risk ranking methods (e.g. Hartwell, 1998; Péry *et al*., 2002). Ecotoxicological assessment criteria for in vivo bioassays ((water and sediment (whole sediment and pore water)) and *in vitro* bioassays are described in the next section. Assessment tools will also need to be developed for data derived from bioassay directed water extract testing.

Water *in vivo* bioassays include techniques that use specific testing regimes and species. Therefore for the purposes of developing background responses and assessment values each technique will require separate review.

Methods for water in vivo bioassays currently in JAMP

Water *in vivo* bioassays

The species recommended for water *in vivo* bioassays are:

- Copepod (*Tisbe battaglii* and *Acartia* sp); 48 h exposure using mortality as the endpoint.
- Bivalves (*Crassostrea gigas, Mytilus* spp) embryos: 24 h exposure using Percent Net Response as the endpoint.
- Sea urchin (*Paracentrotus lividus*): 24 h embryo exposure using percent normal development and larval length as the endpoints.

The methodology for water bioassays is well developed and available through ICES TIMES and/or OECD. Quality Assurance is provided via BEQUALM for the bivalve tests and *Tisbe* assay.

In all water bioassays a control and positive control is used. The control is a "pristine water" of known water quality and characteristic i.e. no contamination, full salinity, appropriate pH and dissolved oxygen e.g. natural seawater from the Atlantic from ICES reference station or Cape Wrath. The *control water* is used in all tests and test animal response in all field and test samples are compared with the test animal response in the control water. A positive control is always used in each experimental design to assess the performance of the testing procedures, including the sensitivity of the test organism. The *positive control* consists of the control water spiked with a reference compound (usually Zn). A *reference water* may also be included for site-specific programmes and may be considered as the control water for the sampling area or region under investigation and ideally should give the same response as the control water.

The methodology for the extraction or concentration generally requires sample manipulation and/or concentration techniques, and clean-up using extraction procedures analogous to those used in chemical analysis. These procedures and QA are currently being developed and documents for ICES and OSPAR are being prepared by the UK and NL. When they are fully in place it will be appropriate to develop the background responses and assessment criteria for these techniques. This may be progressed through the current ICES OSPAR framework.

Assessing the data

The data for water bioassays can be considered in much the same way as for sediment bioassays and the background response is defined as the upper level of natural variation and can be determined as a percentile (for instance 90%) of the individual responses (mortality or malformation) of the control water.

From experience in the UK, Netherlands and Spain the max background level response is of the order of 10% for *Tisbe* sp and *Acartia* sp bioassays, 10% for sea urchin and 15% for the bivalve embryo bioassay. These figures however need to be defined and further established when further data becomes available (see also Table 3 below). Above two times these values and up to 100% the bioassay response is categorized at a level of serious concern (i.e. malformation and mortality is regarded as a serious high level individual population response). Data in this response range should trigger immediate follow up investigations. Responses between background and two times background should be categorized as a cause for concern and prompting further sampling in terms geographical spread and frequency of sampling (possibly time-integrated water sampling). Responses at the serious concern level would initiate further assay of the water test samples using a dilution series in order to quantify the toxicity using a ECx (percent dilution causing a x% reduction in the endpoint) or toxic units (TU=100/ECx) approach. A phased Toxicity Identification Evaluation (TIE) can be conducted to further describe the nature of the toxicity or potential toxicants present.

Preliminary assessment of background response level of available data for water bioassays

A preliminary derivation of background response levels was attempted at the meeting for the water bioassays using *Tisbe bataglii*, bivalve embryo and echinoderm embryo. However, it should be noted that the raw data available at the meeting was limited and only tentative background responses could be calculated. The data were entered into a template (Table 2) and the following calculations made. Data from controls were collected for several tests from different sources. When individual datasets were obtained these were averaged per sample and listed in a databases with standard deviation. From resulting samples the averaged per lab/country was calculated together with the 0.1, 0.5 (median) and 0.9 percentile. In case more datasets were available the same was done with lab/countries datasets (Table 3).

Table 2. Template of data available during the meeting used for calculations of background responses for water and whole sediment bioassays (Median, Min and max are optional).

Test	Name of the test
reference	Reference to the origin of the data
year	Year of production
Country	
lab	Laboratory that performed the analyses
type	Is it a control or other type of sample
Endpoint	Type of measurement
unit	
idnr	Sample number within a dataset
Replicates	Number of replicates
Result	Average value of the control
Median	Median of the individual data
Min	Minimum of the individual data
Max	Maximum of the individual data
Stdv	Standard deviation f the individual exposures
Sed-ino	Information about sediment properties

Table 3. Preliminary results of background response levels for water bioassays.

Test	lab		Average	0.1 perc	0.5 perc	0.9 perc	n
Mussel embryoIEOV		Control	14.1	12.0	13.5	16.5	3

Test	lab		Average	0.1 perc	0.5 perc	0.9 perc	n
Copepods							
Tisbe	Cefas	Control	1.3	0.0	0.0	5.0	28

Ecotoxicological assessment criteria for bioassays

Introduction and definitions

The standard for bioassays described and proposed is based on a recent internal report produced by the Dutch Ministry of Transport, Public Works and Water Management/RWS (Maas *et al*., 2003) and is primarily intended as a step towards the incorporation of biological effect assessment (bioassays in this case) into the CEMP, as desired within OSPAR.

The following definitions and terminology are used.

Bioassays can be divided into *in vivo* and *in vitro* bioassays. A distinction can also be drawn between broad-spectrum bioassays and bioassays based on a specific action mechanism.

In *in vivo* bioassays, whole living organisms (including bacteria) are exposed to environmental samples, or extracts of samples. The tests may be of short duration (lasting several hours to several days), and designed to identify acute effects, or of longer duration (days or months), to determine chronic effects. They can be carried out in a laboratory or in the field (*in situ*). The effects noted, known as 'endpoints', are compared with the endpoints of a control test. *In vivo* bioassays have been developed so as to provide broad-spectrum analysis.

In vitro bioassays are laboratory tests using prepared cells or sub cellular fractions isolated from organisms or modified bacteria. These tests are mechanism-based. They are of short duration (lasting from several minutes to several days), quick to perform and small-scale.

Acute tests provide an initial screening, are of short duration and identify 'crude' effects, such as the death of the test organism. They simulate a 'realistic worst-case' scenario: a one-off, short-term exposure to relatively high concentrations of pollutants.

Chronic tests are designed to emulate the actual situation more closely: longer exposure (i.e. for a substantial proportion of the lifetime of the test organism) to lower concentrations. Endpoints include reduced reproduction or growth in the test organism. Chronic tests are generally more sensitive, but they are also more expensive and more complex in practice than acute tests.

The decision as to whether to perform an acute or chronic test will depend on the degree of pollution in the compartment. In surface waters, for instance, acute effects can be observed near point sources and after incidental adverse events; however, in salt water and fresh water it is usually only possible to observe chronic effects. In cases where neither chronic nor acute effects have been measured, but there is a need to identify trends in toxicity or show the current level of toxicity, acute tests can be performed on concentrates of surface water. However, it must be remembered that not all substances can be concentrated to the same degree using the techniques available (see also 5.3).

The advantages of acute tests are that several tests can be performed simultaneously, that they produce rapid results, that a smaller sample volume is needed and that they are generally cheaper. Water samples are also more constant in acute tests than in chronic tests.

In vivo and *in vitro* bioassays each have their own specific strengths and weaknesses. *In vivo* assays use the entire organism. The exposure situation in such tests is more consistent with the actual situation than in tests where only parts of organisms are used. Processes that play a role in toxicity, such as biological availability, metabolism and bioaccumulation, can therefore be included.

The advantage of chronic *in vivo* bioassays is that they indicate potential longer term effects. However, some chronic tests take a great deal of time, space, manpower and, therefore, money. This applies particularly for larger, longer-lived organisms such as fish. However, some chronic tests can be completed within a fairly short time and cost little more than acute tests. They include growth inhibition tests on bacteria.

Preconditions and criteria for bioassays

To ensure their application and acceptance it is important that bioassays conform to certain criteria and include factors such as *relevance* and *reliability*, for example.

The requirements for recommending a bioassay for JAMP purposes have been proposed by ICES and are described above (Section 2) and must include inter and intra laboratory Quality Assurance procedures. These are provided using agreed international procedures and through BEQUALM and intercalibration exercises. Several further requirements are listed and discussed below. The basic principle is that these tools should allow the ecosystem to be protected as much as possible. The ideal set of bioassays would be representative of all organisms and trophic levels in the ecosystem in question and that the most sensitive species are used. The idea being that the

ecosystem as a whole will be protected if a number of 'trigger species' from several taxonomic groups are protected. Furthermore, in such an ideal situation, the response from the set of bioassays should enable all possible substances to be covered, at both the acute and the chronic level. The set should therefore also have the following qualities:

Ecologically and/or toxicologically relevant

Relevance refers to the guarantee that the bioassay will measure the toxic and ecological effect one is actually interested in. Relevance is determined, among other things, by the test's sensitivity, specificity and discriminatory capacity. Ideally the measured effect should be ecologically relevant and if it is a species that is of ecological/commercial importance then this would be an additional advantage. Bioassays are 'merely' a model of reality. The ecological relevance, in particular, of *in vitro* assays is the subject of debate. We also know too little about how to link the effects at bioassay level with real impacts on the aquatic system. Results from a combined set of bioassays (both *in vivo* and *in vitro*) might, however, provide a weight of evidence as to the ecological relevance of the observed effects.

Representative of all organisms and trophic levels in the ecosystem in question

There is currently no bioassay that is representative of all organisms and trophic levels. This means that a *set* of bioassays is always needed, to cover the ecosystem as fully as possible. Ideally, this set would consist of bioassays for every class of organism: algae, bacteria, crustacea, mollusca, pisces, aves, etc. In line with the guidelines used in chemical standard-setting – at least three or four different taxonomic groups, at least one of which must be vertebrate – a set of at least three or four *in vivo* bioassays would be needed, one of which used fish.

Covering all effects of all possible substances and action mechanisms, both acute and chronic

In vivo bioassays are whole organism tests and therefore by definition respond in an integrated manner to all the contaminants that are present in a test sample (i.e. tests lack specificity but have high relevance). At the moment, there is no one *in vivo* bioassay that could be used to detect all possible mechanisms of toxicity and indeed no in *vitro* bioassay that is capable of detecting all substances or possible action mechanisms. The best way to address this issue is to use a set of *in vivo* and *in vitro* bioassays that cover as many different action mechanisms as possible (see also de Zwart and Sterkenburg, 2002). However, some action mechanisms are not covered fully by *in vivo* bioassays, either because the tests are less sensitive, or because the effect occurs only after long-term exposure. This applies particularly to genotoxicity, immunotoxicity, hormone-disrupting effects and dioxin-like toxicity, as well as the initial signs of neurotoxicity. Effects via these mechanisms are more likely to be detected with *in vitro* bioassays.

Sufficiently sensitive, specific and discriminatory to predict effects

Some bioassays are very sensitive to very small quantities of contaminants in the tested material. This is particularly true of *in vitro* tests, which can respond specifically to a particular contaminant or have specific modes of action. Sometimes, an effect found in an *in vitro* test cannot be replicated in an *in vivo* bioassay. In such cases, the *in vitro* assay is probably too unspecific, so that it also responds to non-active substances present either naturally or otherwise in the matrix. The reverse also occurs: no response *in vitro*, response *in vivo*. In this case, it might be that the *in vitro* bioassay is

too insensitive, or that there has been a loss of compounds during the exposure or processing of the environmental sample. In conclusion, all scenarios can be obviated by using a battery of test methods, or, targeted bioassay use when prior knowledge of the presence of a contaminant is suspected. The bioassay methods described above (see 2) are well tried and intercalibrated and as such the inherent variability of the endpoints of each assay is well documented. Therefore, it is possible to design sampling and test strategies with adequate replication to provide good discriminatory power between test samples.

Reliable and reproducible

The reliability or precision of a bioassay relies on its reproducibility within the same laboratory, or in other laboratories (intra- and inter-laboratory reproducibility). Reproducibility is determined by the stability of the bioassay. A standardized method laid down in a protocol with validity criteria and control for modifying factors is essential to a stable bioassay. All bioassay tests now use positive controls; this consists of a standardized reference material, which is run alongside the test samples and ensures that the response of the assay organism and the conditions are valid for the test.

Availability of test species

For the widespread use and acceptance of a bioassay it is essential that the test organism is widely available geographically and that the species can either be collected easily and cheaply from the wild or is easily cultured in the laboratory. Care also needs to be taken to ensure that too much inbreeding in cultured organisms or seasonality in wild collected organisms does not affect the response of the assay, but this should be taken account of if positive controls are employed.

Clearly, when compiling a set of bioassays for assessing the quality of water one must also take into account other financial and practical considerations. Further conditions therefore include:

Financial

In general bioassays are not expensive (relative to other methodologies) and their incorporation into the CEMP should not entail excessive cost. However it is not possible to specify any particular sum, but it is realized that expensive bioassay packages that could include long-term exposure with chronic endpoints will have little chance of successful introduction and should be confined to targeted and site-specific problems.

Laboratory availability

The introduction of bioassays into the CEMP will place major demands on the available laboratory capacity. This capacity should therefore ideally be expanded. There should preferably be more contract laboratories that can routinely perform bioassays. The bioassays recommended in the JAMP CEMP have well documented protocols and the procedures are easy to learn and in most cases do not require expensive or sophisticated equipment or capital expenditure. Current methods tend to be microscale in operation, which by definition require less space and are more cost-effective.

Use of test animals

Society across Europe wishes to reduce the use of test animals, particularly vertebrates like fish. This trend is only likely to strengthen in future. This automatically

means that *in vivo* bioassays with invertebrate organisms are preferable, and that more effort must be focused on the development of *in vitro* bioassays.

Availability of test and incorporation into metric

By no means all of the promising tests have been worked out to the extent that they can be included in a set of biological effect instruments. The results of the CEMP bioassays in the set must of course be consistent with the proposed metrics.

Taking account of these extra conditions will allow a pragmatic set of bioassays to be selected from the ideal, scientifically sound set of bioassays. Ideally this set should include a minimum of three acute or chronic *in vivo* bioassays on at least three different taxonomic groups, preferably not using vertebrates, and one or more *in vitro* bioassays.

Towards a normative framework for bioassays

The proposed framework for bioassays should preferably be generic, tying in readily with existing policy frameworks and with national and international criteria. An entirely new and unknown system would not be desirable. On the other hand, however, it must be possible to estimate location-specific risks.

It is usually necessary, when conducting *in vitro* tests and rapid, acute *in vivo* tests on surface waters, to produce a concentrate of the surface water. This is necessary because the concentration of contaminants in the bulk water is not acutely toxic, exceptions may be samples taken in estuaries or close to discharge points. Typically, a seawater concentrate is a method whereby contaminants are selectively extracted from a surface water sample (e.g. 100 litres) onto a medium; the medium is eluted with an appropriate solvent, evaporated to a small volume which is subsequently taken back up in seawater (e.g. 100 ml). In this example, a 1000 fold concentration of extractable contaminants and dilutions of this concentrate are bioassayed. Working with concentrates has a number of important advantages:

All kinds of disturbing factors are automatically removed from the test sample during the extraction procedure. They include a high ammonium content, salinity, a high or low pH value, any ion imbalance and hardness. The great advantage is that all water types – fresh water, salt water or brackish water – can be tested using the same (fresh-water or salt-water) methods. This allows one to obtain a picture of the entire OSPAR Convention area, for example, and to compare all locations. Concentrates can be diluted again, so it is almost always possible to obtain a quantitative measure of the toxicity. Using a selective extraction method allows one to determine the cumulative effect of an entire group of substances with the same action mechanism, such as substances with an estrogenic effect.

Bioassays conducted on surface water samples generally use a small sample volume, typically 20–100ml taken from a discrete water sample of say two litres. Water extraction procedures require a larger sample volume (e.g. 100 litres) which can be regarded as a more representative and integrated sample. Furthermore, a greater integration can be achieved by taking samples over time, and subsequently bulking the water samples prior to extraction.

A major advantage of water extraction techniques is that a positive bioassay response can be followed up by bioassay led TIE (Toxicity Identification Evaluation; USEPA 1991 and 1993) procedures. This is a procedure whereby a targeted bioassay response and targeted analytical chemistry can be used to identify the type or, in some cases the specific compound causing the reduced water quality.

There are also drawbacks, however. Usually only a proportion of the substances are extracted and the efficiency of the extraction process will depend on the medium and solvent used. Metals, in particular, tend to get left behind in the current procedures. This restricts our view of the total toxicity of the surface water, forcing us to overlook the combined effects of several substance groups with different action mechanisms, such as metals and organic micro pollutants. The current extraction methods would appear to be broad enough for organic micro pollutants. If not, two extracts can be mixed together, broadening the range of extracted substances. Passive samplers should be considered for the assessment of contaminant concentrations in water (replacing water samples); extracts from passive samplers could then be used for acute *in vivo* bioassays and *in vitro* bioassays. This approach could be used to detect the presence of new chemicals in areas selected for such monitoring. For more discussion of extraction methods, see ICES 2005.

Chronic *in vivo* bioassays would seem to be most suited to site-specific assessment and comparison with the field situation. Long-term exposure without concentration gives the most ecological realistic estimate of possible effects in the field. Appropriate acute bioassays, such as fertilization and embryo development tests, can be a quick, cheap alternative, as can *in vitro* tests.

Availability of test and incorporation into metrics

Not all of the promising bioassays have been established to the extent that they can be included in a set of biological effect instruments. The results from the CEMP bioassays in the set must of course be consistent with the proposed metrics. Taking account of these extra conditions will allow a pragmatic set of bioassays to be selected from the ideal, scientifically sound set of bioassays. Ideally, this set should include a minimum of three acute or chronic *in vivo* bioassays on at least three different taxonomic groups, preferably not using vertebrates, and one or more *in vitro* bioassays.

Assessment framework: metric and criteria

Experience in the Netherlands

The premise of the effects-oriented track for water and sediments is that exposure to substances should not result in "adverse" effects on humans and ecosystems. The metric should therefore be consistent with the environmental risk limits (ERLs) for individual substances. Initially, the ERLs applying in the Netherlands were selected: serious risk (SR), maximum permissible risk (MPR) and negligible risk (NR). However, the term 'risk' is too strongly associated with the derivation of risk limits for single substances based on simple toxicity tests. The following new terms are therefore proposed:

- negligible effect (NE)
- maximum permissible effect (MPE)
- serious effect (SE)

The criteria for water and sediment (i.e. the details of the metric) are set out below, for both *in vivo* and *in vitro b*ioassays. A schematic representation of the metrics is shown in Figure 1.

Proposed metric and criteria for use of *in vivo* bioassays

For the scaling of the results of these bioassays, a metric consistent with the NR-MPR-SR concept has been chosen: the NE-MPE-SE metric. Two points should however be noted regarding consistency with standards for individual substances:

a) Concerning the method: the same methods have been used for the metric as for substance standards, as described in the RIVM report 'Guidance Document on Deriving Environmental Risk Limits (Traas, 2001):

   i) if NOEC values are present for four or more taxonomic groups, refined effect assessment is used. This uses species sensitivity distributions (SSDs) based on the method according to Aldenberg and Jaworska (2000). The criterion for the MPR (or MPE in this case) is the 95% protection level, or PAF_5 (PAF = potentially affected fraction);

   ii) if this condition is not met, preliminary effect assessment is performed, using 'assessment factors'. These factors range from 10 to 1000, depending on the nature of the study-acute or chronic-and the number of ecotoxicity data.

   The same methods are thus used in the metric for bioassays proposed here, the actual choice of method depending on the number of *chronic* data available. It should be noted that the assessment factors for the preliminary effect assessment are applied differently in the metric, though the principle is the same.

b) As regards the factor for MPE/SE: a factor 100 is used to derive the SR for individual substances from the MPR. This factor was chosen because many substances are often found together in the environment, and it takes account of the possible effects of combined toxicity (INS Steering Group, 1999). In bioassays, where samples from the field are used, this effect has already been taken into account, and a factor 10 can be used for converting MPE to SE.

There are also a number of essential differences between *in vivo* bioassays with aquatic organisms and with sediment dwellers, which have implications for the metric:

- in sediment, unlike in fresh water, it is virtually only possible to use *chronic* tests;
- it is possible to use dilutions for both surface water and sediment, based on the undiluted or untreated sample (the 'as is' sample). However, unlike sediment, a water sample can be concentrated, for example with a 1:1 mix of XAD-4 and XAD-8 (de Zwart and Sterkenburg, 2002). Using this technique on water samples makes it easier to scale up the results of *in vivo* bioassays using aquatic organisms to the 'full' metric NE-MPE-SE (so including SE).

Standard for in vivo bioassays for surface water

Method 1. Standard with 'preliminary effect assessment' (Cf = concentration factor compared with the untreated sample (original water sample); this can be seen as the 'assessment factor' applied in the case of three acute or chronic tests from different taxonomic groups).

The table below details the metrics for surface water.

ACUTE TESTS	
NE (negligible effect)	in 3 acute tests effect = 0 (in practice < EC50), Cf = 100
MPE (maximum permissible effect)	in 3 acute tests effect = 0 (in practice < EC50), Cf = 10
SE (serious effect)	in 1 acute test effect ≥ EC50, Cf = 10 or in 2 acute tests EC20< effect < EC50, Cf = 10
CHRONIC TESTS	
NE (negligible effect)	in 3 chronic tests effect = 0, Cf = 10
MPE (maximum permissible effect)	in 3 chronic tests effect = 0, Cf = 1
SE (serious effect)	in 1 chronic test effect ≥ EC50, Cf = 1 or in 2 chronic tests NOEC < effect < EC50, Cf = 1
EC50 = Mean effective concentration, produces a 50% effect in the bioassay	
NOEC = no-observed-effect concentration	

Method 2. Standard with 'refined effect assessment' (PAF approach; see Figure 2)

The method works as follows:

- At least 4 chronic values for different taxonomic groups must be available.
- Both acute and chronic bioassays can be used.
- Results of acute tests are expressed as the concentration factor necessary to reach a 50% effect in the bioassay. These results are transformed into a chronic value by applying an acute-chronic ratio (ACR) of 10. (de Zwart (2002)).
- For chronic values a species sensitivity distribution is assessed following a log-logistic distribution (Traas (2000).
- The extent to which the PAF5 (for the MPE) and PAF50 (for the SE) are exceeded in the undiluted Cf=1 sample is determined.

In order to determine the NE, the Cf (associated with the MPE (PAF5)) is defined and divided by 10. This gives the concentration factor at which the NE acts. This result is compared with the results of the undiluted sample in order to determine whether this conforms to the MPE or the NE.

Standard for *in vivo* bioassays for sediment (whole sediment/pore water)

The standard is as follows:

- Only MPE and SE levels are inferred. This is due to practical issues associated with concentrating sediments.
- To determine the MPE at least three chronic tests must be available, including at least two 'whole sediment' tests.
- As for surface water, the MPE is: in three chronic tests effect = 0. If a negative effect is measured in at least one of three chronic tests, the MPE is exceeded
- The SE level is reached when an effect ≥ EC_{50} is measured in one chronic test (on the 'as is' sample), or an effect between NOEC and EC_{50} in two chronic tests.

The MPE on the metric for both surface water and sediment thus corresponds to the level at which no effect is measured in three *chronic* tests with different taxonomic groups on the 'as is' sample (Cf = 1). On the basis of three *acute* tests the MPE corresponds to the level at which no effect (in practice<EC_{50}) is measured when the sample is concentrated by a factor 10 (Cf = 10) relative to the 'as is' sample. This factor 10 is based on the ACR of 10 (see above). The SE has been derived only for surface water and not for sediment, as it is not possible to concentrate the sediment sample.

The above presentation of a metric for *in vivo* bioassays in surface water states no preference for the use of acute or chronic bioassays. A metric has been developed for both types. The choice of chronic or acute will depend partly on the specific circumstances at the locations studied: the compartment to be assessed, knowledge of the degree of pollution, etc. A choice will therefore have to be made for each type of study and compartment. In this choice, the advantages of acute tests will often outweigh the drawbacks. For instance, chronic effects are sometimes difficult to observe even in concentrates. It is easier to conduct several acute tests simultaneously. Furthermore, the shorter duration of acute tests means the composition of the matrix (water) is more constant, an issue that has proven problematic in chronic tests. If the choice of more acute tests or more chronic tests depends on cost, in our experience the first option is generally preferred (more acute tests, with other organisms or other taxonomic groups).

It is possible to illustrate how the metric for surface waters works in practice on the basis of a 1996 study of the toxicity of surface water in Dutch waters at 15 locations (de Zwart and Sterkenburg, 2002). Acute toxicity tests were performed with five *in vivo* bioassays: the Microtox assay, an algal photosynthesis test using *Selenastrum capricornutum*, the Rotox test, the Thanmotox test and the Daphnia IQ test. A PAF curve was fitted after the acute EC50 values were extrapolated to chronic NOEC values with a factor 10. Although de Zwart and Sterkenburg (2002) estimated the toxicity of the original water sample using the pT method (pT: toxic potency, or the PAF of the undiluted water sample), it is also possible to deduce from their results whether the MPE or SE was exceeded.

Another example of toxicity-based assessment is illustrated in Table 1. Water samples from the surface water monitoring programme of the Western Scheldt estuary (NL) in the period 2000–2005 were extracted using XAD extraction method (de Zwart and Sterkenburg, 2002). This is necessary to achieve an extract in which acute toxicity can be measured. The matrix of the samples is displaced by a standardized medium. Noise effects from for instance nutrients or salt concentrations are removed in order to decrease the number of false positive effects. The extracts were assayed with three different bioassays. To interpret the test results it is important to set criteria for acceptable effects in the undisturbed sample, as explained in §5.4. Table 1 shows the results of a preliminary effect assessment using the test results of the three bioassays.

Table 1. Indication of toxicity in surface water of the Western Scheldt estuary on basis of three different bioassay responses allowing a preliminary effect assessment as proposed in Maas *et al.*, 2003).

LOCATION	DATE	CF (ECF50)*			CF (MTE)
		Daphnia	Algae	Microtox	(from PAF5)
SvOD-1	12-2-00	42	20	19	
SvOD-2	9-4-00	28	16	24	
SvOD-3	11-6-00	54	2.4	23	

SvOD-4	2-8-00	56	3.5	35	
SvOD-5	17-10-00	96	4.5	62	
SvOD-6	15-12-00	87	9	31	
SvOD-1	13-01-05	95	20	27	
SvOD-2	9-03-05	87	30	29	
SvOD-3	2-05-05	127	17	43	
SvOD-4	27-6-05	197	14	44	
SvOD-5	23-8-05	251	10	38	
SvOD-6	19-10-05	94	12	70	
W.Scheldt Vlissingen	4-6-03	416	52	15	2.0
W Scheldt Honte	4-6-03	180	56	38	3.2
W Scheldt Terneuzen	4-6-03	403	28	57	4.0
W Scheldt Hansweert	2-6-03	243	16	84	17.2
W Scheldt Boei s.v WO3	2-6-03	271	15	97	3.2
Scheldt Bath	3-6-03	271	9	52	1.8
Schaar vo Doel (SvoD)	3-6-03	92	9	50	1.6
Scheldt Antwerpen	18-6-03	144	2	23	0.4

corrected for recovery

Expected chronic effect in surface water:

green = negligible effect (NE)

yellow = NE<effect< maximum permissible effect (MPE)

red = serious effect (SE)

Proposed metric and criteria for use of in vitro bioassays

There are several ways of devising a metric for *in vitro* bioassays. Two approaches are regarded as promising:

- Divide Dutch surface waters or sediment into 'clean' and polluted locations on the basis of existing measurements taken in *in vitro* bioassays. The SE is then the average maximum found at locations assumed to be 'clean'. This by analogy with the CTT value for DR-CALUX for the distribution of saline dredged material (Stronkhorst *et al.*, 2001). This value (50 ng TEQ/kg) has been derived from criteria for PCBs and dioxins applying in other countries and observed effects in reference sediments in Dutch coastal waters.
- Regard *in vitro* bioassays as exposure assays, with a response caused by a substance (or group of substances) that triggers the assay. An MPE (or TEQ-MPE) can be calculated via a link to substances on the basis of MPR and SR values for this substance.

A combination of these two approaches would probably be the most realistic, with the first serving to test the derived TEQ-MPE and TEQ-SE.

Another, more complex, approach to producing a metric would be to base it on levels of substances in biological tissue. This is a particularly interesting approach for *in vitro* bioassays that respond to compounds that accumulate in the food chain, such as the dioxin-like compounds measured in the DR-CALUX test, as it takes into account transfer in the food chain. Furthermore, the shifts that occur in the relative content and significance of such substances in the water→sediment→organism→predator chain are also taken into account. The advantage of this approach is that it is good at

predicting the risks to groups of predators or organisms (it has high ecological relevance). Building on this, a metric might also be developed for extrapolation to the ecosystem, taking account of the shift in accumulation patterns at higher trophic levels. Conclusions as to local risks at ecosystem level could be drawn on the basis of the *in vitro* bioassay response in sediment. Such a metric could be established after long-term research, but it would remain limited to a particular food chain in a single area. Although we probably do not yet have the knowledge required to develop a metric in this way, and it would probably not be feasible to apply such an approach to all the different areas and food chains in the OSPAR maritime area, it should not be abandoned entirely. Given the ease of extrapolation to ecosystem level, it is important that we retain this option for the future.

Experience in the UK

The oyster embryo bioassay has been used widely for the measurement of water quality. Surveys in the early 1990s showed no adverse water quality offshore and occasional instances of poor water quality in some UK estuaries. Recent surveys have only been conducted in estuaries. The range of response measured is Percent Net Response (PNR); values range from 0 to 100, where 100 indicates that no oyster embryos developed. A value of 20 or more PNR is regarded as an adverse but negligible effect, a value of between 50 and 80 cause for concern (maximum permissible effect) and in excess of 80 a serious effect. PNR values of between 20–50 have been measured in some UK estuaries but repeated sampling has shown the poor water quality to be transitory.

Over the past six years trials have been conducted using water extraction techniques. Initially these were conducted using a hexane liquid-liquid extraction technique (Thain *et al*., 1996). More recently SPMD extraction procedures have been used successfully (Thomas *et al*., 1999; 2000) and we have developed a battery of bioassay tools to use which include; bivalve embryo development, *Tisbe* bioassay, echinoderm larval development, fish embryo survival, phytoplankton growth and a number of in vitro bioassays, YES and YAS oestrogen screen and the Ahr receptor-based assay. The data has not yet been published but assessment of the water quality results show that Contaminant Concentration Factors (CCF i.e. the concentration of the contaminants in a water sample required to elicit an EC50) are generally;

- >1000 at distant offshore station such as the ICES Reference Stations;
- 500–1000 offshore stations such as the western English Channel;
- 200–500 intermediate stations;
- 50–200 inshore stations;
- 10–50 coastal stations and estuaries;
- >10 only observed in estuaries.

The use of these bioassays and water concentration techniques is in development and therefore no assessment framework has been established. However, it is clear that the procedures permit water quality to be assessed and mapped but that this has to be interpreted within the limitations and restrictions of the chemical process (see 5.3 above).

Conclusions

- Water *in vivo* bioassays are available for immediate deployment within the OSPAR JAMP CEMP. These bioassays have been recommended by ICES

and are of sufficient standing in terms of methodological development, ease of use and application for uptake across the whole OSPAR area. Quality assurance procedures are in place for most of the bioassays and are provided for by BEQUALM. Therefore bioassay data can be submitted to the ICES database for subsequent assessment as appropriate by ICES/OSPAR.

- Bioassays should be deployed as a "battery of tests" and should include a minimum basic set, possibly of three or more. However, the composition of what the set needs to comprise of requires further work. The range of bioassays needs to be expanded to include all trophic levels and phyla such as echinoderms.
- The sampling strategy and design of water quality monitoring for spatial and temporal monitoring purposes needs to be clearly defined and in particular the role of water concentrates. In this respect there is an important need to develop and validate appropriate protocols for extraction methods and subsequent *in vivo* and *in vitro* testing.
- Background response levels and assessment criteria for water bioassays currently in JAMP are available.

References

Åkerman, J and T. Smit. 2003. Effect directed analyses in surface water of the Wadden Sea and Eems-Dollard (in Dutch. RIKZ working document RIKZ/AB/2003.607X.

Åkerman, J *et al*. 2004. Stoffen in de Noordzee in de Nederlandse kustzone in 2003. Internal report nr.2004.040. RWS-RIKZ, The Hague, The Netherlands.

Aldenberg, T. and J. Jaworska. 2000. Uncertainty of the hazardous concentration and fraction affected for normal species sensitivity distributions. Ecotox. Environ. Saf. 25, 48–63.

CIS. 2003. Guidance on Monitoring for the Water Framework Directive. WFD Common Implementation Strategy Working Group 2.7 Monitoring. 23 January 2003.

Den Besten, PJ and M Munawar (eds). 2005. Ecotoxicological testing of marine and freshwater ecosystems: emerging techniques, trends and strategies. CRC Press. Boca Raton, FL USA: 271 pp.

Derksen J.G.M., J.L. Spier, H. Maas and A.D. Vethaak. 2004. Bioassays onder de loep. Selectie van bioassays voor zoet, zout en brak oppervlaktewater. RIZA 2004.012; RIKZ 2004 029; ISBN 9036956366.

Hartwell, S. Ian. 1998. Demonstration of a toxicological risk ranking method to correlate measures of ambient toxicity and fish community diversity. Environmental Toxicology and Chemistry 16: 361–371.

Hendriks, A.J., J.L. Maas-Diepeveen, A. Noordsij and M.A. van der Gaag. 1994. Monitoring response of XAD concentrated water in the Rhine Delta: A major part of the toxic compounds remains unidentified. Water Res. 28: 581–598.

Hill, I.R, P. Matthiessen and F. Heimbach. 1993. Guidance document on sediment toxicity tests and bioassays for freshwater and marine environments. Workshop on sediment toxicity assessment held at Slot Moermond Congress Centre, Renesse, The Netherlands, 8–10 November 1993. SETAC publication.

Hylland, K., Becker, G., Klungsøyr, J., Lang, T., McIntosh, A., Serigstad, B., Thain, T., Thomas, K., Utvik, T.I.R., Vethaak, D., Wosniok, W. 2002. An ICES workshop on biological effects in pelagic ecosystems (BECPELAG): summary of results and recommendations. ICES ASC 2002 CM 2002/X:13.

Hylland Ketil *et al*. (eds). 2006. Biological effects of contaminants in marine pelagic ecosystems. SETAC Special Book Publications (in press).

ICES. 2004. Report of the Working Group on biological effects of contaminants. Report 2004. ICES Marine Habitat Committee ICES CM 2004/E:04 Ref. ACME.

ICES. 2005. Report of the Working Group on biological effects of contaminants. Report 2005.

ICES. CM/2005./E:08. Ref. ACME.

JAMP. (Joint Assessment and Monitoring Programme). 1998. JAMP guidelines for general biological effects monitoring. Oslo and Paris Commissions, 15 p.

Karbe, L. 1992. Toxicity of surface microlayer, subsurface water and sediment-elutriates from the German Bight: summary and conclusions. Marine Ecology Progress Series 91:197–201.

Kirby M. F., M. A. Blackburn, J. E. Thain and M. J. Waldock. 1998. Assessment of water quality in estuarine and coastal waters of England and Wales using a contaminant concentration technique. Marine Pollution Bulletin 36: 631–642.

Maas, J.L., E.J. van de Plassche, A. Straetman, A.D.Vethaak and A. Belfroid. 2003. Normstelling voor bioassays. Uitwerking voor oppervlaktewater en waterbodems. RWS-RIZA/RIKZ report 2003.007.

Maas, J.L., M. van de Heuvel-Greve. 2005. Opportunities for bio-analysis in WFD chemical monitoring using bioassays. RWS-RIZA/RIKZ. Internal report.

Matthiessen Peter, John E. Thain, Robin J. Law and Tim W. Fileman. 1993. Attempts to assess the environmental hazard posed by complex mixtures of organic chemicals in UK estuaries. Marine Pollution Bulletin 26: 90–95.

Oris, J. and S. Klaine. 2000. Whole effluent toxicity testing. Environmental Toxicology and Chemistry 1–256.

Péry, Alexandre R.R , Patrick Flammarion, Bernard Vollat, Jacques J.M. Bedaux, Sebastiaan A.L.M. Kooijman, and Jeanne Garric. 2002. Using a biologic-based model (DEBTOX) to analyze bioassay in ecotoxicology. Opportunities and recommendations. Environmental Toxicology and Chemistry 21: 459–465.

Peters C., S. Becker, U. Noack, S. Pfitzner, W. Bülow, K. Barz, W. Ahlf, and R. Berghahn. 2002. A Marine Bioassay Test Set to Assess Marine Water and Sediment Quality – its Need, the Approach and First Results. Ecotoxicology 11 (5): 379–383.

Power, Elizabeth A. 2004. International Trends in Bioassay Use for Effluent Management. Ecotoxicology 13 (5): 377–398.

Stebbing, A.R.D., Dethlefsen, V., M. Carr. 1992. Results of the Bremerhaven workshop on biological effects of contaminants. Mar. Ecol. Prog. Ser., 91, 361 pp.

Thain, J.E. 1991. Biological effects of contaminants: Oyster (*Crassostrea gigas*) embryo bioassay. Techniques in Marine Environmental Sciences, No. 11. 12 pp.

Thain, J.E. and Kirby, M. 1996. Improving the sensitivity of biological water quality measurements in marine waters. In Proceedings of the Scientific Symposium on the North Sea Quality Status Report 1993. 18–21 April 1994, Ebeltoft, Denmark, pp 151–156. Danish Environmental Protection Agency.

Thomas, Kevin V; Hurst, Mark R; Matthiessen, Peter; McHugh, Mathew; Smith, Andy; Waldock, Michael J. 1992. An assessment of *in vitro* androgenic activity and the identification of environmental androgens in United Kingdom estuaries. Environmental Toxicology and Chemistry 7: 1456–1461.

Thomas, Kevin V., Rachel E. Benstead, John E. Thain and Michael J. Waldock. 1999. Toxicity Characterization of Organic Contaminants in Industrialized UK Estuaries and Coastal Waters. Marine Pollution Bulletin38: 925–932.

Thomas, K.V., Hurst, M.R., Lavender, J., Matthiessen, P., Thain, J.E., Waldock, M.J. 2000. Characterising hazardous substances in the UK marine environment. ICES ASC CM 2000/S:11.

Tonkes, M, PJ den besten, and D Leverett. 2005. Bioassays and tiered approaches for monitoring surface quality and effluents. In: Den Besten, PJ and M Munawar (eds). Ecotoxicological testing of marine and fresh-water ecosystems: emerging techniques, trends and strategies. CRC Press. Boca Raton, FL USA: 43–86.

Traas, T.P. (ed.). 2001. Guidance Document on deriving Environmental Risk Limits. RIVM rapport nr. 601501 012.

US EPA. 1991. Methods for Aquatic Toxicity Identification Evaluation: Phase I Toxicity Characterisation Procedures. EPA Report EPA-600/6-91/003. USEPA, Duluth, MN, USA.

US EPA. 1991. Methods for Aquatic Toxicity Identification Evaluation: Phase II Toxicity Characterisation Procedures. EPA Report EPA-600/R-92/08. USEPA, Duluth, MN, USA.

USEPA. 1995. Short-Term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to West Coast Marine and Estuarine Organisms. National Exposure Research Laboratory, Office of Research and Development. United States Environmental Protection Agency. Cincinnati. 648 pp.

van den Heuvel-Greve, J.L. Maas and A.D. Vethaak. 2005. Ecoassays: an exploratory study on the possible application of ecoassays in the context of the Water Framework Directive. RWS-RIKZ/RIZA. Report nr RIKZ/2005.019.

Zwart, D. de and A. Sterkenburg. 2002. Toxicity-based assessment of water-quality. In: Posthuma, L., G.W. Suter II and T.P. Traas (Eds). Species Sensitivity Distributions in Ecotoxicology. Lewis Publishers. Thain, J.E. 1991. Biological effects of contaminants: oyster (*Crassostrea gigas*) embryo bioassay. ICES TIMES series No. 11. 12 p.

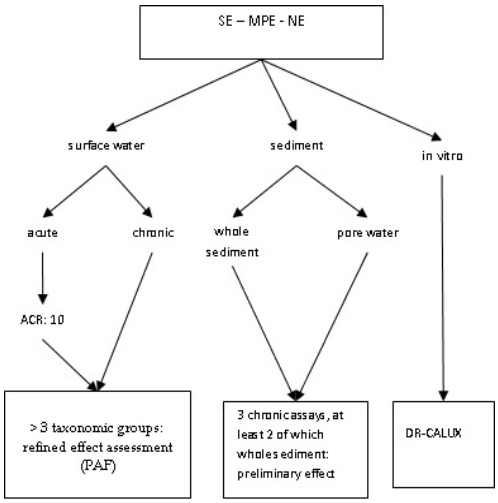


Figure 1. Summary of the metrics based on *in vivo* bioassays for surface water and sediment, and on *in vitro* bioassays (ACR: acute-chronic ratio; PAF: potentially affected fraction).

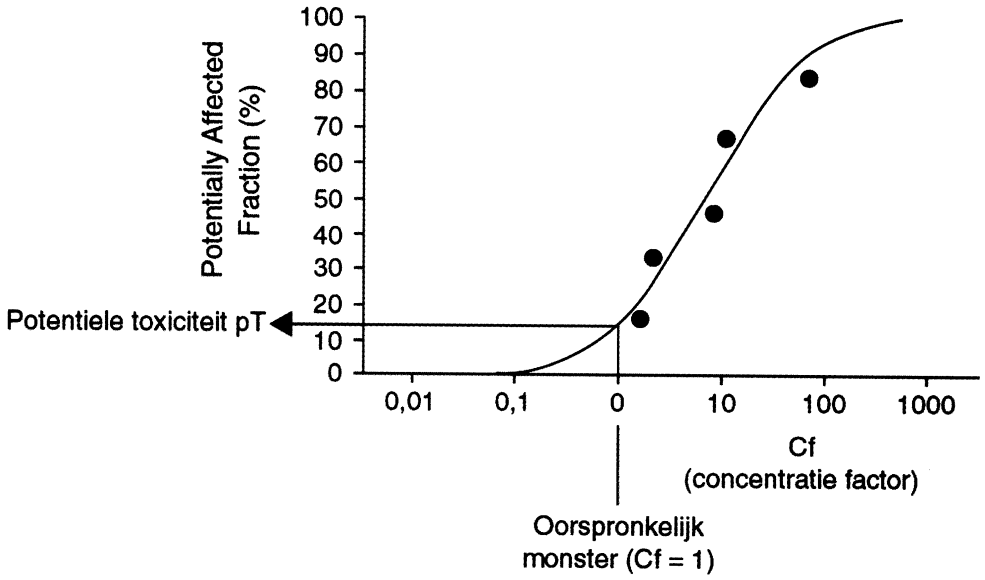


Figure 2. Oorspronkelijk monster=orginal sample (Cf=1).

Annex 10: Acetylcholinesterase as a method for assessing neurotoxic effects in aquatic organisms

Version date: 29 January 2010.

Compiled by: Thierry Burgeot[1], Joelle Forget-Leray[2] and Gilles Bocquené[1]

[1]**Ifremer laboratory of ecotoxicology, BP 21105, 44311 Nantes, France**

[2]**LEMA, Le Havre University, Rue P. Lebon, 76, Le Havre, France**

Background

The analysis of acetylcholinesterase (AChE; EC 3.1.1.7) activity in marine organisms has been shown to be a highly suitable method for assessing exposure to neurotoxic contaminants in aquatic environments. In general, the methods developed are sensitive to detect neurotoxic effects of contaminant concentrations occurring in marine waters. AChE activity method is applicable to a wide range of species and has the advantage of detecting and quantifying exposure to neurotoxic substances without a detailed knowledge of the contaminants present. As applied in human medicine, AChE activity is a typical biomarker that can be used in *in vitro* bioassays and field applications.

AChE is present in most animals and is responsible for the rapid hydrolytic degradation of the neurotransmitter acetylcholine (ACh) into the inactive products choline and acetic acid. AChE has highest specificity for ACh of any other choline ester, while butyrylcholinesterase has the highest specificity for butyrylcholine. The inhibition of AChE leads to an accumulation of ACh which, in turn, over-stimulates sensitive neurons at the neuromuscular junction which results in tonic spasm and tremors. The presence of AChE has been demonstrated in a variety of tissues of marine organisms including muscle and brain tissue of fish, adductor muscle, foot tissue, haemocytes and gills of shellfish, and abdominal muscle of crustaceans (Bocquené and Galgani, 1998). The highest activities have been found in the brain and muscle of fish and in the muscle of prawn. Molluscs in general show low activity (Bocquené *et al*., 1998). In vertebrates, neurotoxic poisoning with hyperactivity, tremors, convulsions and paralysis may finally lead to death.

Being an indicator of neurotoxic effects, AChE has traditionally been used as a specific biomarker of exposure to organophosphate and carbamate pesticides (e.g. Coppage and Braidech, 1976; Day and Scott, 1990; Bocquené and Galgani, 1998; Printes and Callaghan, 2004; Hoguet and Key, 2007). More recently, its responsiveness has been demonstrated to various other groups of chemicals present in the marine environment including heavy metals, detergents and hydrocarbons (Zinkl *et al*., 1991; Payne *et al*., 1996; Guilhermino *et al*., 1998; Forget *et al*., 1999; Brown *et al*., 2004). Its usefulness as a general indicator of pollution stress in mussels from the Baltic Sea has recently been suggested and it has been used for this purpose (Schiedek *et al*., 2006; Kopecka *et al*., 2006, Barsiene *et al*., 2006).

The existence of extremely low thresholds for induction of inhibitory effects on AChE suggests that detection is possible after exposure to low concentrations of insecticides (0.1 to 1 $\mu g l^{-1}$; Habig *et al*., 1986).

Confounding factors

It is important to know the natural limits of variability of AChE activity in the species of interest to assess the significance of the observed depression in activity. A number

of factors are known to affect the level of AChE activity, knowledge of possible variations related to sex, size, state of gonadal maturation and the influence of seawater temperature should be systematically determined. AChE activity of juveniles of *Callionymus lyra* in the Atlantic sea and in *Serranus cabrilla* in the Mediterranean Sea is higher than that of adults but no differences were determined between males and females in *Limanda limanda* in the Atlantic Ocean (Galgani and Bocquené, 1992).

Different biotic and abiotic factors are known to modulate AChE activity, including trace metals (cadmium copper, mercury, zinc) and variation of natural factors, i.e. seawater temperature and salinity (Pfeifer *et al*., 2005; Leiniö and Lehtonen, 2005; Rank *et al*., 2007). In *Mytilus edulis* and *Macoma balthica* from the northern Baltic Sea, mean values of AChE values vary ca. twofold depending on season, following closely changes in temperature (Leiniö and Lehtonen, 2005). Seasonal variability has also been shown as different responses to natural factors in coastal areas compared with offshore sites (Dizer *et al*., 2001; Burgeot *et al*., 2006; Bodin *et al*., 2003). The presence of and exposure to biotoxins or cyanobacteria/cyanobacterial extracts (*Nodularia spumigena*) in mussels has been demonstrated to affect AChE activity (Dailianis *et al*., 2003; Lehtonen *et al*. ,2003; Frasco *et al*., 2005; Kankaanpää *et al*., 2006). Anatoxins produced by algal blooms of cyanobacteria such as *Microcystiis aeruginosa*, *Anabaena flos-aquae* and *Aphanizomenon flos-aquae* have been reported to be strong AChE inhibitors. To avoid confounding factors, it is recommended that the presence of any algal blooms and their identity should be noted when the samples are collected.

The hormone 20-hydroxyecdysone is the primary mechanism controlling molting in crustaceans and has been positively correlated with neurological activity (i.e. AChE) e.g. in *Artemia franciscana* (Gagne and Blaise, 2004). Molting rate increases with the development, specifically peaking at the juvenile stage. The subsequent decline in AChE may also be explained by reduced molting frequencies in adults.

The process and mechanisms of biological response in each organism require further investigation in specific habitats with specific chemical contamination. The mussel *Mytilus galloprovincialis* shows a great heterogeneity of esterases and a particular sensitivity to specific components such as paraoxon (Ozretic and Krajnovic-Ozretic, 1992; Brown, 2004).

Enzymatic polymorphism has also been demonstrated in the oyster *Crassostrea gigas*, and two forms of AChE with different sensitivity to paraoxon have been described (Bocquené *et al*., 1997). Thus, extraction of the sensitive form now identified in some organisms would provide greater precision for determination of AChE enzymatic activity than would an overall measurement of acetylcholinesterases.

Exploration of genetic variability and the influence of environmental factors on specific habitats should lead to a better distinction between natural and pollutant effects.

Ecological relevance

AChE inhibition results in continuous and excessive stimulation of nerve and muscle fibre, producing tetany, paralysis and death. Sublethal exposure affecting AChE can alter the animal's behaviour and locomotive abilities, potentially affecting reproduction, fitness and survival. Evidence of AChE activity modulation by organic chemicals, including fuel oil, has been described in marine organisms including crustaceans (Signa *et al*., 2008). The evaluation of the variations of AChE activity in different species allows characterization of neurotoxic effects of a wide spectrum of organic and inorganic contaminants in the marine environment.

Quality assurance

The large experience acquired in conducting AChE measurements in the field makes it possible today to evaluate the effects of diffuse contamination in some marine organisms sampled in the Atlantic Ocean, the Baltic Sea and the Mediterranean Sea.

A microplate assay technique established for *in vitro* detection of AChE inhibition (Bocquené and Galgani, 1998) has been applied in the monitoring of coastal and offshore waters. This technique has a specific sensitivity comparable with chemical analyses with a detection limit of 100 $ng.L^{-1}$ for carbamates and 10 $ng.L^{-1}$ for organophosphates (Kirby *et al*., 2000).

Standardisation of the sampling strategy and regular intercalibration exercises on specific organisms sampled in the Atlantic Ocean, Mediterranean and the Baltic Sea are still necessary before using AChE in routine pollution monitoring.

No formal quality assurance programmes are currently run within the BEQUALM programme but one major intercalibration exercise was carried out during the BEEP project (Biological Effects of Environmental Pollution in marine coastal ecosystems, EU project EVK3-2000-00543) in 2002.

Background responses and assessment criteria

Baseline levels of AChE in different marine species have been estimated from results derived from French studies in the Atlantic Ocean, the Mediterranean and the Baltic Sea (Table 1).

Therefore, in order to understand and apply the AChE enzymatic activity as a biomarker of neurotoxic exposure it is of fundamental importance to gain information on the natural background levels in non-contaminated organisms during at least two seasonal cycles. The baseline level (35 $nmol.min^{-1}$ mg $prot^{-1}$) of the seasonal cycle of the mussel *Mytilus edulis* studied during three years along the Atlantic coast demonstrated a maximum of amplitude of 30% (Bocquené *et al*., 2004).

Generally it has been accepted that 20% reduction in AChE activity in fish and invertebrates indicates **exposure** to neurotoxic compounds. Depression in AChE activity more than 20% up to 50% indicates **sublethal impact** (Dizer *et al*., 2001). In the field, several species have baseline AChE activities within the same order of magnitude among different studies/measurements (Table 1). However, differences between sea areas and seasons are obvious, with values activity values in *Mytilus* spp. varying from 25 to 54 nmol min mg protein.

Table 1. Assessment of acetylcholinesterase after *in vivo* exposure to commonly used biomonitoring organisms. The table summaries a number of field studies that have utilized commonly deployed species collected from reference locations.

Organisms	Tissue	Reference location or control conditions	Exposure time	Temperature °C or month	Background response AChE activity nmol.min^{-1} mg prot^{-1}	Ref.
Hexaplex trunculus	Digestive gland	Field Mediterranean sea			60	Romeo *et al.* (2006)
Nereis diversicolor	Whole organism	Field			20	Scaps and Borot (2000)
Macoma balthica	Whole organism	Field Baltic Sea			40	Lehtonen and Leiniö (2003)
M. edulis	Gills	Caging in field North Sea	3 weeks	5	25	Burgeot *et al.*, 2006
Mytilus galloprovincialis	Gills	Caging in field Mediterranean sea		September	28	Bodin *et al.*, 2004
Mytilus gallorpovincialis	Gills	Caging in field Mediterranean sea		April	54	Bodin *et al.*, 2004
Mytilus edulis	Gills	Wild mussels Atlantic ocean			35	Bocquené *et al.*, 2004
Maia verrucosa	Whole organism				86	Talesa *et al.* (1992)
Palinurus vulgaris	Whole organism				21	Talesa *et al.* (1992)
Carcinu maenas	Hemolymph				200	Lundebye *et al.* (1997)
Chasmagnathus granulata	Thoracic ganglia				1750	Monserrat and Bianchini (1998)
Palaemonetes pugio	Whole organism				57	Key and Fulton (2002)
Litopenaeus vannamei	Muscle Eyes				6 23	Garcia-de la Parra *et al.* (2006)
Crangon crangon	Cephalothorax				50	Menezes *et al.* (2006)
Squilla mantis	Whole organism				34	Talesa *et al.* (1990)
Artemia salina	Whole organism				3	Varo *et al.* (2002)
Artemia parthenogenetica	Whole organism				4	Varo *et al.* (2002) Nunes *et al.* (2006)
Paracentrotus lividus	Ambulacra				47	Cunha *et al.* (2005)
Polybius henslowii		Field Spain	*In situ*		45	Signa *et al.*, 2008
Trigriopus brevicornis	Whole organism	Field France	In situ		300	Forget *et al.*, 2001
Eurytemora affinis	Whole organism	Field France estuary	In situ		750	Cailleaud *et al.* (2007)
Vertebrates						

Organisms	Tissue	Reference location or control conditions	Exposure time	Temperature °C or month	Background response AChE activity nmol.min^{-1} mg prot^{-1}	Ref.
Plathichthys flesus	Muscle	Field Baltic sea	*In situ*	6.6	50	Kopecka J. and Pemkowiak J., 2008
Plathichthys flesus	Muscle	Field Baltic sea	*In situ*	22	103	Kopecka J. and Pemkowiak J., 2008
Serranus Cabrilla		Field Mediterranean sea	*In situ*	20	89	Diezer *et al.*, 2001
Acanthurus Bahianus		Field Caribbean and Atlantic ocean	*In situ*	26	220	Galgani *et al.*, 1996

Future work

Consensus is required on a standardized AChE measurement protocol for the main species currently used in marine biomonitoring programmes (OSPAR, HELCOM, MEDPOL and MSFD).

Further information should be gathered to confirm baseline activity levels in specific habitats.

Assessment criteria (background activities need to take account of seasonal (temperature) patterns of Ache activity (See ICES/OSPAR SGIMC 2010 report).

References

Baršiene, J., Lehtonen, K. K., Köhler, A., Broeg, K., Vuorinen, P.J., Lang, T., Pempkowiak, J., Šyvokiene, J., Dedonyte, V., Rybakovas, A., Repecka, R., Vuontisjärvi, H., Kopecka, J. 2006. Biomarker responses in flounder *(Platichthys flesus)* and mussel (*Mytilus edulis)* in the Klaipeda-Butinge area (Baltic Sea). Marine Pollution Bulletin 53: 422–436.

Bocquené G., Roig A., Fournier D. 1997. Acetylcholinesterases from the common oyster *Crassotrea gigas*. Evidence for the presence of a soluble acetylcholinesterase insensitive to organophosphate and carbamate inhibitors. *FEBS Lett* 3:261–266.

Bocquené, G., and Galgani F. 1998. Biological effects of contaminants: Cholinesterase inhibition by organophosphate and carbamate compounds. *Ices Techniques in Marine Environmental Sciences*. No 22.12p. ICES Copenhaguen DK.

Bocquené G., Chantereau S., Raffin B., Minier C., Clérendeau C., Leskowicz A., Beausir E., Burgeot T., Ménard D. 2004. The monitoring of the biological effects of the ERIKA oil spill on the blue mussel (*Mytilus edulis*). *Aquatic Living Resources*. 17. 309–316.

Bodin N., T.Burgeot, J.Y. Stanisière, G. Bocquené, D.Menard, C. Minier, I. Boutet, A. Amat, Y.Cherel , H.Budzinski. 2004. Seasonal variations of a battery of biomarkers and physiological indices for the mussel *Mytilus galloprovincialis* transplanted into the Northwest Mediterranean Sea. *Comparative Biochemistry and Physiology.* 138:411–427.

Brown R.J., Galloway T.S.? Lowe D., Browne M.A., Dissanayake A., Jones M.B. and Depledge M.H. 2004. Differential sensitivity of three marine invertebrates to copper assessed using multiple biomarkers. *Aquat. Toxicol*, 66.267–278.

Burgeot T., Faucet J, Ménard D, Grosjean P., and Bocquené G. 2006. Variations of benzo(a)pyrene hydroxylase and cholinesterase activities in mussels caged in the North Sea (German Bight and Statfjord). Special Issue. *Environmental Toxicology and chemistry*. 171–185.

Coppage D.L., Braidech T.E. 1976. River pollution by antilcholinesterase agents. *Water Research*, 10: 19–24.

Dailianis, S., Domouhtsidou, G.P., Raftopoulou, E., Kaloyianni, M., Dimitriadis, V.K. 2003. Evaluation of neutral red retention assay, micronucleus test, acetylcholinesterase activity and a signal molecule (cAMP) in tissues of *Mytilus galloprovincialis* (L.), in pollution monitoring. *Marine Environmental Research* 56, 443–470.

Day K. E. and Scott I. M. 1990. Use of acetylcholinesterase activity to detect sublethal toxicity in stream invertebrates exposed to low concentrations of organophosphate insecticides. *Aquatic Toxicology,* 18: 101–114.

Dizer H., da Silva de Assis H.C. and P.-D. Hansen. 2001. Cholinesterase activity as a bioindicator for monitoring marine pollution in the Baltic Sea and the Mediterranean sea.pp331–342. In: Biomarkers in marine organisms; a practical approach. Edited by Ph. Garrigues, H. Barth, C.H. Walker and J.F. Narbonne. Eds Elsevier.491p.

Frasco, M.F., Fournier, D., Carvalho, F. and Guilhermon, L. 2005. Do metal inhibit acetylcholinesterase (AChE)? Implementation of assay conditions for the use of AChE activity as a biomarker of metal toxicity. *Biomarkers*, 10, 360–375.

Gagne F. and Blaise C. 2004. Shell protein characteristics and vitellogenin-like proteins in brine shrimp *Artemia fransciana* exposed to municipal effluent and 20-hydroxyecdsone. *Comp. Biochem. Physiol. Part C Pharmacol.Toxicol* 138. 515–522.

Galgani F., Bocquené G. and Cadiou Y. 1992. Evidence of variation in cholinesterase activity in fish along a pollution gradient in the North Sea. Mar. ecol. Prog. Ser. 91: 1–3).

Galgani F. , Bocquené G and T. Burgeot. 1996. Acetylcholinesterase and ethoxyresorufin-o-deethylase in the surgeonfish *Acanthurus bahianus* around Martibique Island (French West Indies). *Biomarkers* 1. 1–3.

Habig C., Di Guilio R.T., Nomeir A.A. and Abou Donia A.M.B. 1986. Comparative toxicity, cholinergic effects and tissue levels of S,S,S-tri-nButyl Phosphoritrithioate 'DEF) to Channel catfish (*Ictarus punctatus*) and bleu crab (*Callinectes sapidus*). *Aquatic Toxicology*, 9:193–206.

Hansson, T., Schiedek, D., Lehtonen, K.K., Vuorinen, P.J., Liewenborg, B., Noaksson, E., Tjärnlund, U., Hanson, M. and Balk, L. 2006. Biochemical biomarkers in adult female perch (*Perca fluviatilis*) in a chronically polluted gradient in the Stockholm recipient (Sweden). Marine Pollution Bulletin 53: 451–468.

Hoguet J. and P.B.Key. 2007. Activities of biomarkers in multiple life stages of the model crustacean, *Palaemonetes pugio. Journal of experimental marine biology and ecology*. 235–244.

Kankaanpää, H., Leiniö, S., Olin, M., Sjövall, O., Meriluoto, J., Lehtonen, K. K. 2007. Accumulation and depuration of cyanobacterial toxin nodularin and biomarker responses in the mussel *Mytilus edulis*. Chemosphere 68: 1210–1217.

Kirby M.F., Morris S., Hurst M. Kurby S.J., Neall P., Tylor T. and Fagg A. 2000. The use of cholinesterase activity in flounder (*Platichthys flesus*) muscle tissue as a biomarker of neurotoxic contamination in UK estuaries. *Mar. Pollut. Bull*. 9:780–791.

Kopecka, J., Lehtonen, K. K., Baršienė, J., Broeg, K., Vuorinen, P.J., Gercken, J., Balk, L., Pempkowiak, J. 2006. Measurements of biomarker levels in flounder (*Platichthys flesus*) and blue mussel (*Mytilus trossulus*) from the Gulf of Gdańsk (southern Baltic). Marine Pollution Bulletin 53: 406–421.

Lehtonen, K. K., Kankaanpää, H., Leiniö, S., Sipiä, V.O., Pflugmacher, S, Sandberg-Kilpi, E. 2003. Accumulation of nodularin-like compounds from the cyanobacterium *Nodularia spumigena* and changes in acetylcholinesterase activity in the clam *Macoma balthica* during short-term laboratory exposure. Aquatic Toxicology 64: 461–476.

Lehtonen, K. K., Leiniö, S. 2003. Effects of exposure to copper and malathion on metallothionein levels and acetylcholinesterase activity of the mussel *Mytilus edulis* and the clam *Macoma balthica* from the northern Baltic Sea. Bulletin of Environmental Contamination and Toxicology 71: 489–496.

Leiniö, S., Lehtonen, K. K. 2005. Seasonal variability in biomarkers in the bivalves *Mytilus edulis* and *Macoma balthica* from the northern Baltic Sea. Comparative Biochemistry and Physiology C 140: 408–421.

Ozretið and Krajnovið-Ozretið. 1992. Esterase heterogeneity in mussel *Mytilus galloprovincialis*: effects of organophosphate and carbamate pesticides in vitro. *Comp Biochem Physiol* 103:221–225.

Payne J.F., Mathieu A., Melvin W., Fancey L.L. 1996. Acetylcholinesterase, an old biomarker with a new future? Field trials in association with two urban rivers and a paper mill in Newfoundland. *Mar. Pollut. Bull*. 32:225–231.

Pfeifer, S., D. Schiedek and J. W. Dippner. 2005. Effect of temperature and salinity on acetylcholinesterase activity, a common pollution biomarker, in Mytilus sp. from the south-western Baltic Sea. J. exp. mar. biol. ecol. 320: 93-103.

Printes L.B. and Callaghan A. 2004. A comparative study on the relationship between acetylcholinesterase activity and acute toxicity in *Daphnia magna* exposed to anticholinesterase insecticides. *Environ. Toxicol.Chem.*23, 1241–1247.

Rank J., Lethonen K.K.? Strand J., Laursen M. 2007. DANN damage, acetylcholinesterase activity and lysosomal stability in native and transplanted mussels (*Mytilus edulis*) in areas close to coastal chemical dumping sites in Denmark. *Aquatic Toxicology* 84, 50–61.

Rank, J., Lehtonen, K. K., Strand, J., Laursen, M. 2007. DNA damage, acetylcholinesterase activity and lysosomal stability in native and transplanted mussels (*Mytilus edulis*) in areas close to coastal chemical dumping sites in Denmark. Aquatic Toxicology 84: 50–61.

Signa, G., Cartes J.E., M. Sole., Serrano, A. and Sanchez, F. 2008. Trophic ecology of the swimming crab *Polybius henslowii* Leach, 1820 in Galician and Cantabrian Seas: Influences of natural variability and the Prestige oil spill. *Continental Shelf Research*. 2659–2667.

Schiedek, D., Broeg, K., Baršiene, J., Lehtonen, K. K., Gercken, J., Pfeifer, S., Schneider, R., Vuontisjärvi, H., Vuorinen, P. J., Koehler, A., Balk, L. 2006. Biomarker responses and indication of contaminant effects in blue mussel (*Mytilus edulis*) and female eelpout (*Zoarces viviparus*) from the western Baltic Sea. Marine Pollution Bulletin 53: 387–405.

Zinkl, J.G., Lockhart, W.L., Kenny, S.A., Ward, F.J. 1991. The effects of cholinesterase inhibiting insecticides on fish. In: Mineau P (ed) Cholinesterase Inhibiting Insectisides. Elsevier, Amsterdam, pp. 233-254.

Annex 11: Micronucleus assay as a method for assessing DNA damage in marine organisms

Version date: 29 January 2010.

Author: Brett Lyons

Background

Micronuclei are acentric fragments of chromosomes or whole chromosomes which are not incorporated into daughter nuclei at anaphase (Al-Sabti and Metcalfe, 1995). A micronucleus (MN) arises due to either spindle apparatus malfunction or chromosomal fragmentation. The lagging or whole/fragmented chromosome then form a secondary small nucleus (MN) during the telophase stage of the cell cycle. Scoring MN during interphase then provides a measure of genotoxicity. Due to the fact that MN may arise from fragmented or whole chromosomes, their measurement can indicate exposure to either clastogenic or aneupoidy inducing contaminants (Figure 1).

The MN assay was first developed as a routine *in vivo* mutagenicity assay for detecting chromosomal mutations in mammalian studies (Heddle, 1973). Hooftman and de Raat (1982) were the first to successfully apply the assay to aquatic species when they demonstrated the induction of micronuclei in erythrocytes of the eastern mudminnow (*Umbra pygmaea*), following waterborne exposure to the known mutagen ethyl methanosulphate (EMS). Since these initial experiments other studies have validated the detection of micronuclei as a suitable biomarker of genotoxicity in a wide range of both vertebrate and invertebrate species (for review see Chaudhary *et al*., 2006; Udroiu *et al*., 2006). While most studies have utilized circulating erythrocytes, cells can be sampled from a number of tissues, such as gill epithelium, liver (Barŝiené *et al*., 2006), digestive gland and kidney (ref).

The application of the assay to blood samples of fish is particularly attractive due to the ease of sampling and the number of assessable cells present.

However recent studies have demonstrated that the frequencies of MN may differ between tissues and species (Barŝiené *et al*., 2006). The frequencies of micronucleated erythrocytes sampled from peripheral blood and cephalic kidney differed and the authors proposed the hypothesis that species such as cod and turbot remove micronuclei from the peripheral circulation via their spleen. This phenomenon has also been recognized in mammals and has led to the Organization for Economic Cooperation & Economic Development (OECD) statement "any appropriate mammalian species may be used provided it is a species in which the spleen does not remove micronucleated erythrocytes" (OECD, 1997).

Scoring of micronuclei

When selecting a test species, consideration must be given to its karyotype as many teleosts are characterized by an elevated number of small chromosomes (Udroiu *et al*., 2006). Thus, in certain cases micronuclei formed after exposure to clastogenic contaminants will be very small and hard to detect by light microscopy. This can be addressed to a certain extent by using fluorescent staining. After selecting suitable species, researchers should also ensure that other factors including age, sex, temperature and diet are similar between the sample groups. If conducting transplantation studies, consideration needs to be given to the cellular turnover rate of the tissue be-

ing examined to ensure sufficient cells have gone through cell division. For example, if using blood the times of erythropoiesis should be known prior to sampling. After sampling, slides should be coded and "blindly" scored. The area to be scored should first be examined under low magnification to select the part of the slide showing the highest quality (good staining, non overlapping cells). Scoring of micronucleated cells should then be undertaken at 1000x magnification. Previous reviews have suggested that when using fish erythrocytes at least 2000 cells should be scored per animal (Udroiu *et al*., 2006).

Diagnostic criteria for micronuclei identification were developed for micronucleus scoring according to several authors:

- The diameter of fish micronuclei varies between 1/10th and 1/30th of the mean diameter of the main nuclei.
- Micronuclei are non-refractile and can therefore be distinguished from artefacts such as staining particles.
- Micronuclei are not connected to the main nuclei.
- Micronuclei may touch but not overlap the main nuclei and the micronuclear boundary should be distinguishable from the nuclear boundary.
- Micronuclei usually have the same staining intensity as the main nuclei but occasionally staining may be more or less intense.

Background responses and assessment criteria

MN frequency lower than 0.05‰ (the Baltic Sea) and lower than 0.1‰ (the North Sea) could be suspected as a reference level in the peripheral blood erythrocytes of flatfish (Rybakovas *et al*., 2009).

Background levels of micronuclei in marine species derived from field reference sites or control laboratory conditions are summarized in Table 1.

Table 1. Reported background (field reference sites or control laboratory conditions) levels of MN in marine species.

Species	Tissue	Endpoint	Exposure	Response	Ref.
M. edulis	Haemolymph	MN/1000 cells	Lab control	<5	Canty *et al.*, 2009
M. galloprovincialis	Haemolymph	MN/1000 cells	Field sites	<5	Burgeot *et al.*, 1996
M. galloprovincialis	Haemolymph	MN/1000 cells	Field sites	<2	Dailianis *et al.*, 2003
M. galloprovincialis	Haemolymph	MN/1000 cells	Field sites	1.43 – 1.5	Klobucar *et al.*, 2008
M. galloprovincialis	Gill	MN/1000 cells	Field sites	<3	Dailianis *et al.*, 2003
M. galloprovincialis	Gill	MN/1000 cells	Lab control	2.90± 1.24 6.75 ± 1.97	Scarpato *et al.*, 1990
M. galloprovincialis	Gill	MN/1000 cells	Field control	2.94–4.25	Magni *et al.*, 2006
M. edulis	Gill	MN/1000 cells	Field reference	1.2–2.72	Baršiené *et al.*, 2006
M. edulis	Gill	MN/1000 cells	Field reference	1.42–5.75	Izquierdo *et al.*, 2003
M. galloprovincialis	Gill	MN/1000 cells	Field reference	≈ 5	Nigro *et al.*, 2006
Dicentrarchus labrax	Erythrocytes	MN/1000 cells	Field reference	1·25 ± 1·97	Strunjak-Perovic *et al.*, 2009
Mugil spp.	Erythrocytes	MN/1000 cells	Field reference	2·04 ± 4·41	
P. flesus	Erythrocytes	MN/1000 cells	Field reference (Swedish West coast)	0.04–0.05	Baršiené *et al.*, 2006
P. flesus	Erythrocytes	MN/1000 cells	Field reference (Baltic)	0.19–0.32	Rybakovas *et al.*, 2009
Limanda limanda	Erythrocytes	MN/1000 cells	Field reference (Baltic)	0.13–0.15	
L. limanda	Erythrocytes	MN/1000 cells	Field reference (N.Sea)	0.02–0.44	
L. limanda	Erythrocytes	MN/1000 cells	Field reference (R.Elbe)	0.12–0.21	
P. flesus	Erythrocytes	MN/1000 cells	Lab Control	0.2–0.56	Lyons, unpublished
			Field reference (UK estuaries)	0.27–0.66	

Assessment criteria for MN in *M. edulis* have been proposed by Hagger *et al*. (2008; 2009).

They divided the data (a combination of previous laboratory and field studies) into four categories: those that had slight alteration from baseline responses were assigned a numerical rank of four, moderate alterations were assigned a three, major alterations a two and, finally, biological responses that were severely altered from previously recorded baseline values were assigned a one.

Table 2. Proposed assessment criteria for *Mytilus edulis* and *Carcinus maenas* (Hagger *et al.*, 2008; Hagger *et al.*, 2009).

Species	Tissue	Endpoint	Slight	Moderate	Major	Severe
M. edulis	Haemolymph	MN/1000 cells	0–1.9	2.0–3.9	4.0–5.9	6.0+
C. maenas	Haemolymph	MN/1000 cells	0–0.99	1.0–1.99	2.0–2.99	3.0+

Assessment criteria for MN fish require to be developed.

References

Baršiene, J., Rybakovas, A., Förlin, L., Šyvokiene, J. 2008. Environmental genotoxicity studies in mussels and fish from the Göteborg area of the North Sea. Acta Zoological Lithuania 18 (4), pp. 240–247.

Barŝiené J., Lehtonen K.K., Koehler. A., Broeg. K., Vuorinen, P.J., Lang T., Pempkowiak J., Syvokiene J., Dedonyte V., Rybakovas A., Repecka A., Vuontisja¨rvi H., Kopecka J. 2006. Biomarker responses in flounder (*Platichthys flesus*) and mussel (*Mytilus edulis*) in the Klaipeda-Butinge area (Baltic Sea). Marine Pollution Bulletin 53 (2006) 422–436.

Burgeot, T., Wollt, S., Galgani, T. 1996. Evaluation of the Micronucleus Test on *Mytilus galloprovincialis* for Monitoring Applications along French Coasts Marine Pollution Bulletin, Vol. 32, No. 1, pp. 39–46.

Chaudhary, R., Pandy, S., Kushwaha, B., Gaur, K.K., Nagpure, N.S. 2006. Fish micronucleus assay: A sensitive tool for ecogenotoxicity studies. Journal of Ecophysiology and Occupational Health 6 (3-4), pp. 143–147.

Dailianis, S., Domouhtsidou, G.P., Raftopoulou, E., Kaloyianni, M., Dimitriadis, V.K. 2003. Evaluation of neutral red retention assay, micronucleus test, acetylcholinesterase activity and a signal transduction molecule (cAMP) in tissues of *Mytilus galloprovincialis* (L.), in pollution monitoring. Marine Environmental Research 56, 443–470.

Hagger *et al*. 2008. Application of biomarkers for improving risk assessments of chemicals under the Water Framework Directive: A case study, Marine Pollution Bulletin, vol. 56, p. 1111–1118.

Hagger *et al*. 2009. Application of biomarkers to assess the condition of European Marine Sites. Environmental Pollution 157:2003-10.

Izquierdo, J.I., Machado, G., Ayllon, F., d'Amico, V.L., Bala, L.O., Vallarino, E., Elias, R. and Garcia-Vazquez, E. 2003. Assessing pollution in coastal ecosystems: a preliminary survey using the micronucleus test in the mussel *Mytilus edulis*, Ecotoxicology and Environmental Safety 55 24–29.

Klobučar, G.I.V., Štambuk, A., Hylland, K., Pavlica, M. 2008. Detection of DNA damage in haemocytes of *Mytilus galloprovincialis* in the coastal ecosystems of Kaštela and Trogir Bays, Croatia. Science of the Total Environment 405 (1–3), pp. 330–337.

Magni, P., De Falco, G., Falugi, C., Franzoni, M., Monteverde, M., Perrone, E., Sgro, M., Bolognesi C. 2006. Genotoxicity biomarkers and acetylcholinesterase activity in natural populations of *Mytilus galloprovincialis* along a pollution gradient in the Gulf of Oristano (Sardinia, western Mediterranean). Environmental Pollution 142, 65–72.

Nigro, M., Falleni, A., Del Barga, I., Scarcelli, V., Lucchesi, P., Regoli, F., Frenzilli, G. 2006. Cellular biomarkers for monitoring estuarine environments: Transplanted versus native mussels. Aquatic Toxicology 77, 339–347.

Rybakovas, A., Baršiene, J., Lang, T. 2009. Environmental genotoxicity and cytotoxicity in the offshore zones of the Baltic and the North Seas. Marine Environmental Research 68 (5), pp. 246–256.

Scarpato, R., Migliore, L., Alfinito-Cognetti, G. and Barale, R. 1990. Induction of micronuclei in gill tissue of *Mytilus galloprovincialis* exposed to polluted marine waters. Marine Pollution Bulletin, 2(2) 74–80.

Strunjak-Perovic, I., Topic Popovic, N., Coz-Rakovac, R., Jadan, M. 2009. Nuclear abnormalities of marine fish erythrocytes. Journal of Fish Biology 74 (10), pp. 2239–2249.

Udroiu, I. 2006. The micronucleus test in piscine erythrocytes. Aquatic Toxicology 79 (2), pp. 201–204.

Annex 12: Comet assay as a method for assessing DNA damage in aquatic organisms

Version date: 29 January 2010.

Author: Brett Lyons

Background

The analysis of modified or damaged DNA has been shown to be a highly suitable method for assessing exposure to genotoxic contaminants in aquatic environments. In general, the methods developed are sensitive to a range of contaminant concentrations, applicable to a wide range of species and have the advantage of detecting and quantifying exposure to genotoxins without a detailed knowledge of the contaminants present. The Single Cell Gel Electrophoresis (SCGE) or comet assay was first applied to ecotoxicology over 15 years ago, and has since become one of the most widely used tests for detecting DNA strand breaks in aquatic animals[1-5]. The comet assay has many advantages over other methods commonly used to assess genotoxic exposure, including (1) genotoxic damage can be detected in most eukaryotic cell types at the single cell level; (2) only a small number of cells are required; (3) it is a rapid and sensitive technique; (3) Due to the nature of DNA strand break formation it provides an early warning response of genotoxic exposure.

As a consequence of the advantages listed above the comet assay has been used widely in both laboratory and field based studies to assess genotoxic exposure in many fresh water and marine organisms. However, unlike mammalian genotoxicology, where the focus is limited to a small number of model species, efforts in the aquatic field have generally lacked coordination and have used an extensive range of sentinel species[1,3,5]. While guidelines relating to the use of the comet assay have been published for mammalian genotoxicology[6,7], no standard protocols currently exist for environmental studies. Consequently, the variations in protocols can lead to major differences in results and an inability to directly compare studies. Despite these obvious limitations the comet assay provides a well-researched tool for studying genotoxicity in aquatic species.

Confounding factors: protocols, cell types and target organs

The majority of aquatic studies published to date have used circulating blood cells (either haemocytes or erythrocytes), as target cells for comet assay analysis. This is likely to be due to the practical advantage of processing tissues from a ready-made supply of nucleated cells in suspension. Solid tissues such as gill or fish hepatocytes require dissociation prior to analysis, with the potential for introducing damage through enzymatic or mechanical processes. Studies have also demonstrated that different cell types responded with different sensitivities to contaminant exposure. When comparing cells types it is usually reported that circulating cells are less sensitive than hepatocytes or gill cells[8-13]. Blood and to a lesser extent the haemolymph of bivalve molluscs (e.g. mussels) are ''buffered'' tissues, in which contaminants arrive having crossed numerous biological barriers. Gill cells appeared to be the most sensitive following MNNG exposure, while liver and digestive gland were more sensitive to B(a)P, suggesting that uptake routes and bioaccumulation mechanisms need to be taken into account when designing experiment systems[12].

Mammalian studies have demonstrated that certain tissue types may have higher background levels of DNA damage due to presence of alkali sensitive sites in cells with highly condensed chromatin[14]. Similar studies comparing basal levels of DNA migration in mussel gill cells, haemocytes and fish erythrocytes under both mild alkaline (pH 12.1) and alkaline versions (pH > 13) of comet assay have supported this assumption[15, 16]. Indicating that the mild alkaline version of the assay should be employed when dealing with certain cell types (e.g. fish erythrocytes), in order to prevent higher background levels of DNA strand breaks inhibiting data interpretation. Indeed, this problem has been highlighted in other studies using fish species where excessive DNA tail migration has inhibited the interpretation of results[17].

In addition to the variation in response depending on cell type, it is also apparent a range of comet assay protocols (differing in terms of agarose concentrations, lysing and electrophoresis parameters) have been used in studies with aquatic organisms[1-5]. Therefore, effort is required to establish standardized protocols for the main species and cell type commonly used in environmental studies. The production of standard protocols, or the initiation of inter laboratory ring testing workshops focused on aquatic species are essential if the comet assay is to develop further as an environmental monitoring tool.

Ecological relevance

Marine invertebrates

Marine invertebrates have been widely used as sentinel species in environmental monitoring programs. This is mainly due to their sessile nature, ability to bioaccumulate contaminants and general ease of capture[18-20]. The majority of work has focused on coastal and estuarine environments. For example, Hartl *et al*. used the clam (*Tapes semidecussatus*) as an indicator species for the presence of potentially genotoxic substances in estuarine environments, demonstrating an increase in DNA damage in haemocytes, gill and digestive gland cells of animals exposed to contaminated sediments[8]. The study also highlighted the differences in sensitivity between cell types, with gill and digestive gland cells appearing to be the most sensitive target tissues for detecting genotoxic exposure. The Mediterranean mussel (*Mytilus galloprovincialis*) has also been extensively deployed as a sentinel organism to assess the genotoxic effects of crude oil spills[21-23]. Studies have demonstrated the sensitivity of mussels to oil exposure and laboratory studies have clearly linked the total polycyclic aromatic hydrocarbon (TPAHs) content of oils with the level of DNA damage observed[21]. In Northern European studies the Blue mussels (*M. edulis*) has also been used to differentiate sites receiving waste treatment effluent, with positive correlations detected between the presence of selected contaminants and the level of DNA damage[24].

Mussels have also been used extensively in the field as part of transplantation studies[25-27]. The use of indigenous organisms is often hampered by the absence of a suitable sentinel species, or if present, the genotoxic responses obtained may be influenced by local physiological adaptations. Furthermore the use of transplanted organisms also offers advantages over indigenous species, such as ensuring genetic homogeneity, developmental/reproductive status and controlling the precise exposure window. Validation studies have been under taken with the comet assay to assess the time course variations in DNA damage following field transplantation experiments[25, 26]. It was observed that within the first 7 days following transplantation the level of DNA damage can fluctuate, which is likely to be caused by manipulation disturbance, then after 2 weeks the level reaches a plateau. Such data suggests that

transplantation experiments lasting less that 2 weeks may give spurious results, with the levels of DNA damage detected attributable to artefacts associated with the sampling procedure rather than genotoxic exposure. Studies conducted in a coastal area of Denmark, impacted by a disused chemical site, have also highlighted that the levels of DNA damage in mussels can be affected by seasonal variations in baseline levels[25]. Such results are likely to be influenced by the seasonal variations, which are known to exist for a range of physiological and reproductive processes in mussels[28, 29].

The sampling location has also been shown to influence the results of field-based surveys. For example, mussels (*M. edulis*) sampled from the intertidal zone in Reykjavik harbour had higher levels of DNA damage when compared with mussels collected from the sub tidal zone at the same site[30]. While the study supports the use of DNA strand breaks as a measure of environmental pollution it also highlights the high levels of intra site variability of DNA damage that can occur. As such the study further serves to underline the importance of validating experimental protocols and sampling procedures to ensure that non-contaminant related factors (e.g. physiological and biochemical responses to variations in oxygen availability and temperature stress) do not adversely affect biomarkers data.

Marine vertebrates

There are a limited number of comet assay studies utilizing marine fish species compared with those using fresh water species (for detailed review see[1, 4, 5]). This is mainly due to the logistical problems associated with collecting fish at sea (e.g. need for a research vessels) and technical problems inherent within the assay, such as the difficulty of performing electrophoresis reproducibly at sea (e.g. dealing with adverse weather conditions). To date those studies undertaken have mainly focused on flatfish and bottom-feeding species, which due to their close association with sediment bound contaminants are widely used in marine monitoring programmes[31, 32]. *In vivo* studies have been undertaken to investigate oxidative stress in the European eel (*Anguilla anguilla*)[33]. The comet assay has also proven to be a useful tool for studying the genotoxic effects of non bioaccumulating contaminants in the marine environment. For example, the environmental effects of the known mutagen and potential carcinogen styrene has been studied in the mussel (*M. edulis*) and fish (*Symphodus mellops*)[34]. Styrene hasn't previously been considered to be harmful to marine fauna due to its high volatility and low capacity to bioaccumulate. However, it was shown to cause a statistically significant increase in DNA damage in blood cells, probably due to the formation of a radical styrene metabolite, which is thought to have potent oxidative capacity. Hatchery-reared turbot (*Scophthalmus maximus* L.) have been used successfully to investigate the genotoxic potential of PAH and heavy metal contaminated sediment from sites in Cork Harbour (Ireland)[35]. Eelpout (*Zoarces viviparus*) have been used in site-specific investigative monitoring following a bunker oil spill in Goteborg harbour, Sweden. The comet assay was deployed along site a battery of other bioassays and elevated levels of DNA damage were correlated with the presence of PAH metabolites in the bile of fish[36]. The marine flatfish dab (*Limanda limanda*) is a commonly used flatfish species in offshore monitoring programmes and it has been used in a number of studies investigating the impacts of genotoxic contaminants in coastal and estuarine waters[37-39]. Studies have shown that both sex and age of the fish have a significant effect on the presence of DNA strand breaks, which again highlights the influence other factors (i.e. reproductive status) may have on the extent of DNA damage.[37, 38].

Quality assurance

No formal quality assurance programmes are currently run within the marine monitoring community. However, a series of comet assay workshops have taken place with the aim of drafting a common regulatory strategy for industrial genotoxicology screening[6,7]. Final guidelines drafted after the 4th International Workgroup on Genotoxicity testing: Results of the *in vivo* Comet assay workgroup[7] provide a useful starting point for developing quality assurance programmes specifically focused on protocols employed in marine species. These include consideration of 1) cell isolation processes*; 2) cryopreservation processes; 3) concurrent measures of cytotoxicity; 4) Image analysis and scoring method.

Currently data can be reported in a number of formats. % DNA in tail has been reported to be the most linearly related to exposure dose[7]. However there is no clear consensus of which measure of DNA migration should be used (% DNA in tail, Tail moment, Tail length). This difference in scoring criteria hinders our ability to develop a consensus background response and assessment criteria.

***if required**

Background responses and assessment criteria

It is recognized that setting baseline/background response levels have an important role in integrating biological effect parameters into environmental impact assessments of the marine environment. The general philosophy is that an elevated level of a particular biomarker, when compared with a background response, indicates that a hazardous substance has caused an unintended or unacceptable level of biological effect. Therefore, in order to understand and apply the Comet Assay as a biomarker of genotoxic exposure it is of fundamental importance to gain information on the natural background levels in non-contaminated organisms. Table 1 summaries a number of studies that have utilized commonly deployed bio-indicator species collected from reference locations (as supported by chemical and biomarker analyses) or kept under control conditions in the laboratory. While these studies provide a starting point for determining "background" levels of DNA damage they also serve to highlight the number of different tissues, protocols and endpoints currently reported.

Table 1. Assessment of "control DNA damage" by Comet assays after *in vivo* exposure to commonly used biomonitoring organisms.

Organism	Cell type	Agent	Exposure time	Parameter	Control response	Ref.
Invertebrates						
M. edulis	Haemocytes	MMS	0–4 days	Tail Moment	2.08 ± 3.43 2.96 ± 4.60	[25]
M. edulis	Haemocytes	Tritiated water	96 hrs	% DNA Tail	<10	[40]
M. edulis	Haemocytes	TBT	7 days	% DNA Tail	5–10	[41]
M. edulis	Haemocytes	MMS	3–7 days	% DNA Tail	<10	[44]
M. edulis	Gill cells	Cd Cr Cr VI	10 days 7 days injection	% DNA Tail	<15	[42]
M. edulis	Gill cells	MMS		Tail Moment	1.87 ±2.23 0.60 ± 1.05 3.84 ± 3.61 1.22 ± 1.47	[25]
M. edulis	Gill cells	Field site	*In situ*	Tail Moment	<1.5	[45]
M. edulis	Gill cells	Field site	*In situ*	Tail Moment	<5	[46]
M. edulis	Digestive gland	H_2O_2, BaP	1hr	% DNA Tail	<10	[43]
Vertebrates						
L.limanda	Erythrocytes	Field	*In situ*	Tail Moment	<5	[39]
L.limanda	Erythrocytes	Field	*In situ*	% DNA Tail*	4–6	[37]
P. olivaceus	Erythrocytes	Field	*In situ*	Tail length (µm)	<10	[47]
Zoarces viviparus	Erythrocytes	Field	*In situ*	% DNA Tail	<15	[36]

***Mean square root of percent tail DNA measured.**

The requirement now is to establish a common set of protocols for those tissues/species routinely used in biomonitoring programmes. Once established it will be possible to define internationally accepted background levels of DNA damage and from their establish assessment criteria.

Required steps

- Reach consensus on standardized protocol from main species currently used in marine biomonitoring programmes (OSPAR, HELCOM, MEDPOL and MSFD);
- Establish minimum acceptable reporting criteria (cellular toxicity, +/- control, etc.);
- Agree data reporting format to allow cross study comparisons of data (Tail moment, % DNA in Tail, Tail moment);
- Establish international quality assurance;
- Develop assessment criteria.

References

[1] C.L. Mitchelmore and J.K. Chipman, DNA strand breakage in aquatic organisms and the potential value of the comet assay in environmental monitoring, Mutat. Res., 1998, 399, 135–147.

[2] S. Cotelle and J.F. Fèrard. Comet assay in genetic ecotoxicology: a review, *Environ. Mol. Mutagen.*, 1999, 34, 246–255.

[3] R.F. Lee and S. Steinert, Use of the single cell gel electrophoresis/comet assay for detecting DNA damage in aquatic (marine and freshwater) animals, *Mutat. Res.*, 2003, 544, 43–64.

[4] A.N. Jha. Ecotoxicological applications and significance of the comet assay, *Mutagenesis,* 2008, 23 (3), 207–221.

[5] G. Fenzilli, M. Nigro, B.P. Lyons. The Comet assay for evaluation of genotoxic impact in aquatic environments. *Mutat Res*., 2009, 681, 80–92.

[6] R.R. Tice, E. Agurell, D. Anderson, B. Burlinson, A. Hartmann, H. Kobayashi, Y. Miyamae, E. Rojas, J.C. Ryu and Y.F. Sasaki, Single cell gel/comet assay: guidelines for *in vitro* and *in vivo* genetic toxicology testing, *Environ. Mol. Mutagen.* 2000, 35, 206–221.

[7] B. Burlinson, R.R. Tice, G. Speit, E. Agurell, S.Y. Brendler-Schwaab, A.R. Collins, P. Escobar, M. Honma, T.S. Kumaravel, M. Nakajima, Y.F. Sasaki, V. Thybaud, Y. Uno, M. Vasquez and Andreas Hartmann, Fourth International Workgroup on Genotoxicity testing: results of the *in vivo* Comet assay workgroup, *Mutat. Res*., 2007, 627, 31–35.

[8] M.G. Hartl, B.M. Coughlan, D. Sheehan, C. Mothersill, F.N. van Pelt, S.J. O'Reilly, J.J. Heffron, J. O'Halloran and N.M. O'Brien, Implications of seasonal priming and reproductive activity on the interpretation of Comet assay data derived from the clam, Tapes semidecussatus Reeves 1864, exposed to contaminated sediments, *Mar. Environ. Res*., 2004, 57, 295–310.

[9] W.H. Siu, J. Cao, R.W. Jack, R.S. Wu, B.J. Richardson, L. Xu and P.K. Lam, Application of the comet and micronucleus assays to the detection of B[*a*]P genotoxicity in haemocytes of the green-lipped mussel (*Perna viridis*), *Aquat. Toxicol*., 2004, 66, 381–392.

[10] S. Lemiere, C. Cossu-Leguille, A. Bispo, M.J. Jourdain, M.C. Lanhers, D. Burnel and P. Vasseur, DNA damage measured by the single-cell gel electrophoresis (Comet) assay in mammals fed with mussels contaminated by the 'Erika' oil-spill, *Mutat. Res*., 2005, 581, 11–21.

[11] S. Pandey, N.S. Nagpure, R. Kumar, S. Sharma, S.K. Srivastava and M.S. Verma, Genotoxicity evaluation of acute doses of endosulfan to freshwater teleost *Channa punctatus* (Bloch) by alkaline single-cell gel electrophoresis, *Ecotoxicol. Environ. Saf*., 2006, 65, 56–61.

[12] I.Y. Kim and C.K. Hyun, Comparative evaluation of the alkaline comet assay with the micronucleus test for genotoxicity monitoring using aquatic organisms, *Ecotoxicol. Environ. Saf*., 2006, 64, 288–297.

[13] D. Huang, Y. Zhang, Y. Wang, Z. Xie and W. Ji, Assessment of the genotoxicity in toad *Bufo raddei* exposed to petrochemical contaminants in Lanzhou Region, China, *Mutat. Res*., 2007, 629, 81–88.

[14] N.P. Singh, D.B. Danner, R.R. Tice, M.T. McCoy, G.D. Collins and E.L. Schneider, Abundant alkali-sensitive sites in DNA of human and mouse sperm, *Exp. Cell. Res*., 1989, 184, 461–470.

[15] G. Frenzilli, V. Scarcelli, F. Taddei and M. Nigro, Adaption of SCGE as a candidate for monitoring marine ecosystems, *Neoplasma*, 1999, 46, 6–7.

[16] M. Moretti, M. Villarini, G. Scassellati-Sforzolini, A.M. Santroni, D. Fedeli and G. Falcioni, Extent of DNA damage in density-separated trout erythrocytes assessed by the 'comet' assay, *Mutat. Res.* 1998, 397, 353–360.

[17] G. Wirzinger, L. Weltje, J. Gercken and H. Sordyl, Genotoxic damage in field-collected three-spined sticklebacks (*Gasterosteus aculeatus* L.): a suitable biomonitoring tool? *Mutat. Res.*, 2007, 628, 19–30.

[18] B.L.Bayne. Watch on mussels. *Mar. Pollut. Bull*., 1976 7(12), 217–128.

[19] R. Seed. Ecology. In: B.L. Bayne, Marine Mussels, their ecology and physiology, International Biological Programme (13–65), Cambridge University Press, 1976, 13-65.

[20] M.H. Salazar and S.M. Salazar. *In situ* bioassays using transplanted mussels: I. Estimating chemical exposure and bioeffects with bioaccumulation and growth. In Hughes JS, Biddinger GR, Mones E, Eds. Environmental Toxicology and Risk Assessment – Third Volume. Philadelphia: American Society for Testing and Materials STP, 1995 2118, 216–241.

[21] B. Perez-Cadahia, B. Laffon, E. Pasaro and J. Mendez, Evaluation of PAH bioaccumulation and DNA damage in mussels (*Mytilus galloprovincialis*) exposed to spilled Prestige crude oil, *Comp. Biochem. Physiol. C: Toxicol. Pharmacol*., 2004, 138, 453–460.

[22] B. Laffon, T. Rabade, E. Pasaro and J. Mendez, Monitoring of the impact of Prestige oil spill on *Mytilus galloprovincialis* from Galician coast, *Environ. Int.,* 2006 32, 342–348.

[23] I.C. Taban, R.K. Bechmann, S. Torgrimsen, T. Baussant and S. Sanni, Detection of DNA damage in mussels and sea urchins exposed to crude oil using comet assay, *Mar. Environ. Res.,* 2004, 58, 701–705.

[24] J. Rank, K. Jensen and P.H. Jespersen, Monitoring DNA damage in indigenous blue mussels (*Mytilus edulis*) sampled from coastal sites in Denmark, *Mutat. Res./Genet. Toxicol. Environ. Mutagen*., 2005, 585, 33–42.

[25] J. Rank, K.K. Lehtonen, J. Strand and M. Laursen, Aquatic toxicology, DNA damage, acetylcholinesterase activity and lysosomal stability in native and transplanted mussels (*Mytilus edulis*) in areas close to coastal chemical dumping sites in Denmark, *Aquat. Toxicol.,* 2007, 84, 50–61.

[26] F. Regoli, G. Frenzilli, R. Bocchetti, F. Annarumma, V. Scarcelli, D. Fattorini and M. Nigro, Time-course variations of oxyradical metabolism, DNA integrity and lysosomal stability in mussels, *Mytilus galloprovincialis*, during a field translocation experiment, *Aquat. Toxicol.,* 2004, 68, 167–178.

[27] M. Nigro, A. Falleni, I. Del Barga, V. Scarcelli, P. Lucchesi, F. Regoli and G. Frenzilli, Cellular biomarkers for monitoring estuarine environments: transplanted versus native mussels, *Aquat. Toxicol.,* 2006, 77, 339–347.

[28] A. Hines, G.S. Oladirana, J.P. Bignell, G.S. Stentiford, M.R. Viant. Direct sampling of organisms from the field and knowledge of their phenotype: key recommendations for environmental metabolomics, *Environ. Sci. Tech.* 2007, 41, 3375–3381.

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

[30] H.P. Halldorsson, G. Ericson and J. Svavarsson, DNA strand breakage in mussels (*Mytilus edulis* L.) deployed in intertidal and subtidal zone in Reykjavik harbour, *Mar. Environ. Res*., 2004, 58, 763–767.

[31] S.W. Feist, T. Lang, G.D. Stentiford, A. Köhler, A. 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*, 2004 38. 42 pp.

[32] JAMP guidelines for general biological effects monitoring. *Joint Assessment and Monitoring Programme*. Oslo and Paris Commissions, 1998, 38 pp.

[33] F. Regoli, G.W. Winston, S. Gorbi, G. Frenzilli, M. Nigro, I. Corsi and S. Focardi, Integrating enzymatic responses to organic chemical exposure with total oxyradical absorbing ca-

pacity and DNA damage in the European eel *Anguilla anguilla*: toward development of a more holistic biomarker assessment, *Environ. Toxicol. Chem.*, 2003, 22, 2120–2129.

[34] E. Mamaca, R.K. Bechmann, S. Torgrimsen, E. Aas, A. Bjornstad, T. Baussant and S.L. Floch, The neutral red lysosomal retention assay and Comet assay on haemolymph cells from mussels (*Mytilus edulis*) and fish (*Symphodus melops*) exposed to styrene, *Aquat. Toxicol.*, 2005, 75, 191–201.

[35] M.G.J. Hartl, M. Kilemade, D. Sheehan, C. Mothersill, J. O'Halloran, N.M. O'Brien and F.N.A.M. van Pelt, Hepatic biomarkers of sediment-associated pollution in juvenile turbot, *Scophthalmus maximus* L, *Mar. Environ. Res.*, 2007, 64, 191–208.

[36] G. Frenzilli, V. Scarcelli, I. Del Barga, M. Nigro, L. Forlin, C. Bolognesi and J. Sturve, DNA damage in eelpout (*Zoarces viviparus*) from Goteborg harbour, *Mutat. Res.*, 2004, 552, 187–195.

[37] F. Akcha, G. Leday and A. Pfohl-Leszkowicz, Potential value of the comet assay and DNA adduct measurement in dab (*Limanda limanda*) for assessment of *in situ* exposure to genotoxic compounds, *Mutat. Res.*,2003, 534, 21–32.

[38] F. Akcha, F. Vincent Hubert and A. Pfohl-Leszkowicz, Measurement of DNA adducts and strand breaks in dab (*Limanda limanda*) collected in the field: effects of biotic (age, sex) and abiotic (sampling site and period) factors on the extent of DNA damage, *Mutat. Res.*, 2004, 552, 197–207.

[39] B.P. Lyons, G.D. Stentiford, J. Bignell, F. Goodsir, D.B. Sivyer, M.J. Devlin, D. Lowe, A. Beesley, C.K. Pascoe, M.N. Moore and E. Garnacho, A biological effects monitoring survey of Cardigan Bay using flatfish histopathology, cellular biomarkers and sediment bioassays: findings of the Prince Madog Prize 2003, *Mar. Environ. Res.*, 2006 62, S342–S346.

[40] Jha, A.N., Dogra, Y., Turner, A., Millward, G.E. 2005. Impact of low doses of tritium on the marine mussel, *Mytilus edulis*: Genotoxic effects and tissue-specific bioconcentration. Mut. Res.Genet. Toxicol. Environ. Mutagen 586 (1), pp. 47–57.

[41] Hagger, J.A., Depledge, M.H., Galloway, T.S. 2005. Toxicity of tributyltin in the marine mollusc *Mytilus edulis*. Mar. Poll. Bull. 51 (8–12), pp. 811–816.

[42] Emmanouil, C., Smart, D.J., Hodges, N.J., Chipman, J.K. 2006. Oxidative damage produced by Cr(VI) and repair in mussel (*Mytilus edulis* L.) gill. Mar. Environ. Res. 62 (SUPPL. 1), pp. S292–S296.

[43] Mitchelmore, C.L., Birmelin, C., Livingstone, D.R., Chipman, J.K. 1998. Detection of DNA Strand Breaks in Isolated Mussel (*Mytilus edulis* L.) Digestive Gland Cells Using the ''Comet'' Assay. *Ecotox and Environ*. Saf. 41, 51–58. 1998.

[44] Canty, M.N., Hutchinson, T.H., Brown, R.J., Jones, M.B., Jha, A.N. 2009. Linking genotoxic responses with cytotoxic and behavioural or physiological consequences: Differential sensitivity of echinoderms (*Asterias rubens*) and marine molluscs (*Mytilus edulis*). Aquat. Toxicol., 94 (1), pp. 68–76.

[45] Rank J. 2009. Intersex in *Littorina littorea* and DNA damage in *Mytilus edulis* as indicators of harbour pollution. Ecotox. Environ. Saf., 72 (4), pp. 1271–1277.

[46] Rank J. Lehtonen, K.K., Strand, J., Laursen, M. 2007. DNA damage, acetylcholinesterase activity and lysosomal stability in native and transplanted mussels (*Mytilus edulis*) in areas close to coastal chemical dumping sites in Denmark. Aquat.Toxicol., 84 (1), pp. 50–61.

[47] Woo, S., Kim, S., Yum, S., Yim, U.H., Lee, T.K. 2006. Comet assay for the detection of genotoxicity in blood cells of flounder (*Paralichthys olivaceus*) exposed to sediments and polycyclic aromatic hydrocarbons. Mar. Poll. Bull. 52 (12), pp. 1768–1775.

Annex 13: Reproductive success in eelpout (*Zoarces viviparus*)

Version date: 29 January 2010.

Background

The eelpout (*Zoarces viviparus*), also called viviparous blenny, can be used as a bio-indicator of the impact of hazardous substances on reproductive success of fish in the marine environment. The reproductive success in fish is a generic "stress" indicator; causal agents may, however, be identified through a combination of chemical analyses of fish tissue, a knowledge of the history of contamination of the local environment to which the fish have been exposed and/or follow-up laboratory experimentation (Jacobsson *et al*., 1986). Substances such as organochlorines, pesticides, PAH, heavy metals and organometals can affect embryo and larval development in fish (Bodammer, 1993). Several of these substances, which may induce developmental, morphological and/or skeletal anomalies, have also been identified as endocrine disrupting substances (Davis, 1997).

Figure 1. The eelpout is a viviparous fish and the pregnant female bears 20–300 living embryo and larvae in the ovarian cavity (Photo: Jakob Strand).

The eelpout inhabits coastal waters from the White Sea to the southern North Sea. However it is not equally abundant in all areas and it may therefore be difficult to sample adequate numbers throughout the OSPAR area. Use in regional assessments is more appropriate. However, studies of reproductive success in eelpout are recommend by ICES, OSPAR and HELCOM for marine monitoring programmes of biological effects (OSPAR 1997, HELCOM 2006, ICES 2004), and, for instance, Sweden and Denmark have included this method in regional and national monitoring programmes in coastal waters of the Baltic Sea, the Kattegat and the Skagerrak.

It should be noted that eelpouts are protected in the pregnancy period in some areas and an official sampling licence for monitoring activities should be obtained, where necessary.

The methodology is well defined for studies in coastal waters and national guideline exists (Jacobsson *et al*., 1986; Neuman *et al*., 1999, Strand and Dahllöf, 2005). An international guideline is in preparation and to be published in the ICES TIMES series. As method quality assurance, some international and national workshops have been held in relation to the monitoring programmes (e.g. BEQUALM, 2000). A Baltic workshop has been held in 2009 as part of BONUS+-projects BALCOFISH and

BEAST. National workshops in relation to NOVANA monitoring activities have also been held in Denmark (Strand, 2005a).

Elevated levels of adverse developmental effects of embryo and larvae in eelpout broods have been found in populations living in contaminated areas with effluents from cities and industry. In comparison, only low levels of such effects generally occur in populations living in areas regarded as reference sites (e.g. Vetemaa *et al*., 1997; Ådjers *et al*., 2001; Sjölin *et al*., 2003; Strand *et al*., 2004; Kalmarweb, 2005; Gercken *et al*., 2006), however some year -to-year variations can occur (Figure 2). Acute larval mortality has also been observed in eelpout exposed to pulp mill effluents (Jacobsson *et al*., 1986). Other environmental stress factors like increased temperatures and oxygen depletion events may however also affect eelpout reproduction (Veetema, 1999; Fagerholm, 2002; Strand *et al*., 2004). Reproductive success in eelpout is regarded as a general, i.e. non-specific, biological indicator of impaired fish reproduction.

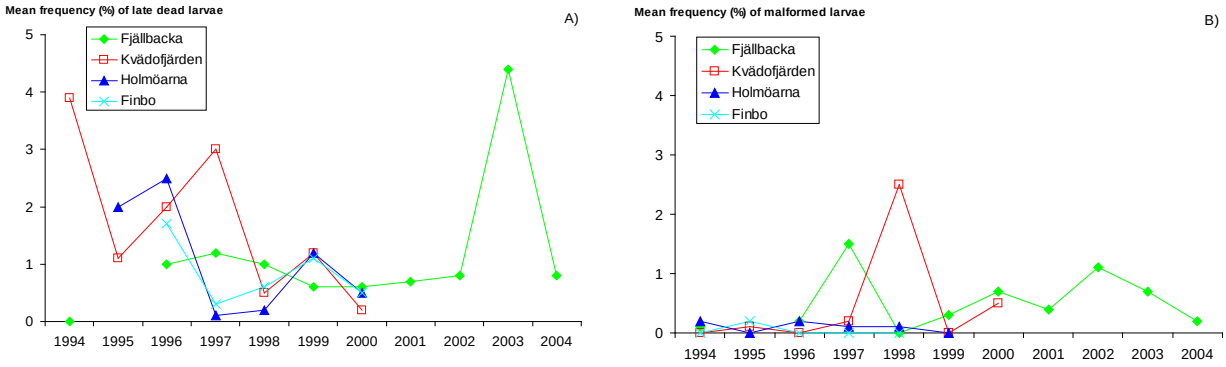


Figure 2. Year-to-year variations in mean frequencies of A) late dead larvae and B) malformed larvae at four Swedish monitoring stations regarded as reference sites.

According to the technical guidelines used in the Swedish and Danish monitoring programmes (Neuman *et al*., 1999; Strand and Dahllöf, 2005), supporting parameters like water temperature and salinity together with general fish physiological and reproductive parameters should be recorded when the reproductive success in eelpout is examined.

For simplifying reasons and as a first step, only the occurrence of abnormal development of embryo and larvae in the broods of pregnant eelpouts has been included in the proposed assessment criteria for impaired reproduction. However, other relevant fish physiological and reproductive parameters must been seen as supplementary parameters and how they can be integrated should be further evaluated.

Abnormal development of embryo and larvae in eelpout broods can, according to the Swedish and Danish guidelines (Neuman *et al*., 1999; Strand and Dahllöf, 2005), be characterized as;

- Malformed larvae: larvae with morphological and/or skeletal gross anomalies. This includes yolk-sac or intestinal defects, bent spine or spiral shapes of the spinal axis, eye defects including rudimentary or missing eye(s), cranio-facial defects and conjoined/Siamese twins more or less separated.
- Late dead larvae: dead larvae without malformations and with a length >15 mm (>10 mm in Denmark).

- Growth retarded larvae: normal developed larvae which are smaller than the three highest length classes in the broods.

Less visible aberrations including altered behavioural aspects are not included in this analysis, although they can be highly ecological relevant effects.

Similar with studies on skewed sex ratio in eelpout broods, although it can be used as an indicator of endocrine disruptions. For instance a Swedish study has found significant male-biased sex ratios of eelpout embryos (53.9%–61.3% males) in an area contaminated with paper mill effluents (Larsson and Förlin, 2002). In eelpout broods the reference conditions are supposed to be 50:50 between females and males.

Proposal for assessment criteria of the reproductive success in eelpout

The approach for deriving the assessment criteria is based on statistical analyses, which imply that the effect level must be significant different from the background response, i.e. where the impact of environmental factors such as contaminants can be regarded as close to zero.

52 datasets from 14 sampling stations regarded as reference sites and 41 datasets from 22 stations not regarded as reference sites in the Baltic Sea, the Kattegat and the Skagerrak from the period 1994–2004 are available for the analyses. However, an important assumption is that adequate reference sites actually can be found in the Baltic Sea, the Kattegat and the Skagerrak, although these waters are generally regarded to be more polluted compared with the North Sea and the North Atlantic.

Both data related to frequencies (mean percents) of abnormal larvae per female and frequencies of broods with >5% abnormal larvae (i.e. related to individual pregnant females) are used in the analyses. However, data of >5% distributions are only available from 37 of the 93 datasets, and there is no information was found available of broods with >5% growth retarded larvae.

Data on frequencies of females with (at least one) abnormal larvae present in the brood is not included in this analysis, because the influence of brood size cannot be discriminated.

Proportion of abnormal larvae per female	Proportion of broods with elevated levels of abnormal larvae
Mean frequency of late dead larvae.	Frequency of broods containing >5% late dead larvae.
Mean frequency of malformed larvae.	Frequency of broods containing >5% malformed larvae.
Mean frequency of growth retarded larvae.	No data

In the assessment criteria the upper level of the background response (class I) is determined by the 90% percentile of all datasets observed in areas regarded as reference sites, i.e. in distance to larger cities and industry.

Assessment criteria related to mean frequencies of abnormal larvae in broods

Most studies on development of eelpout embryo and larvae from the Baltic Sea, the Kattegat and the Skagerrak studies have used mean frequencies of late dead, malformed and growth retarded larvae in the broods as a measure of impaired reproduction in eelpout.

In areas which were considered as reference sites only small frequencies of abnormal larvae have been found, if any. Values of 90% percentiles have been found to be 1% malformed larvae, 2% late dead larvae and 4% growth retarded larvae, respectively.

Table 1. Proposal for assessment criteria for the mean frequencies of malformed larvae, late dead larvae and growth retarded larvae per station.

ASSESSMENT CLASS	CLASS I BACKGROUND RESPONSE	CLASS II
Mean frequency of malformed larvae	0–1%	>1%
Mean frequency of late dead larvae	0–2%	>2%
Mean frequency of growth retarded larvae	0–4%	>4%
	Background response. The upper limit is the 90% percentile of response at reference sites.	Elevated effect levels

Comparisons of datasets shows that class II, i.e. elevated mean frequencies of malformed larvae and late dead larvae, mainly have be found in areas which are not regarded as reference sites, i.e. suspected to be more polluted (Figure 3). However, only one of the datasets shows significantly elevated levels of growth retarded larvae in the broods.

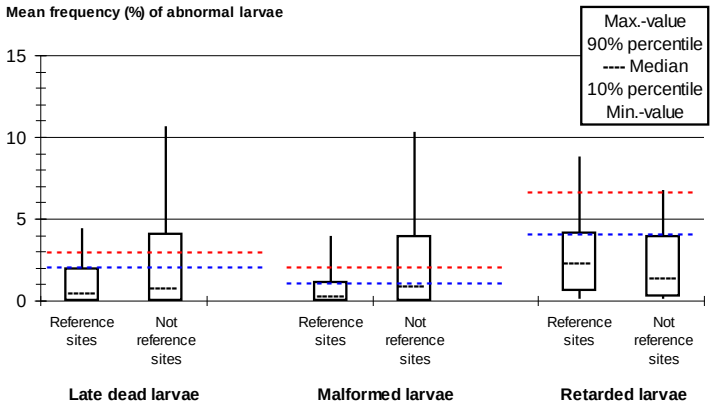


Figure 3. Comparison of data distribution of data on mean frequencies of late dead, malformed and growth retarded larvae in eelpout broods from reference sites and area not regarded as reference sites. The blue dotted line refers to the 90% percentile of data from the reference sites. The red dotted line refers to significantly elevated levels compared with the 90% percentile of the reference sites.

Assessment criteria related to individual broods with >5% abnormal larvae

Some Swedish and Danish eelpout studies from the Baltic Sea, the Kattegat and the Skagerrak studies have also used the frequency of pregnant eelpout containing elevated proportions of late dead or malformed larvae in the broods (e.g. >5%) as a measure of impaired reproduction in eelpout.

In areas which were considered as reference sites, only low frequencies have been found if any (90% percentiles: 5%), of the pregnant eelpout containing elevated frequencies of late dead and malformed larvae in the broods (i.e. >5%).

Table 2. Proposed assessment criteria for the frequencies of pregnant eelpouts, which contain more than 5% malformed larvae and late dead larvae in their broods.

ASSESSMENT CLASS	CLASS I BACKGROUND RESPONSE	CLASS II
Frequency of broods with >5% malformed larvae	0–5%	>5%
Frequency of broods with >5% late dead larvae	0–5%	>5%
	Background response. The upper limit is the 90% percentile of response at reference sites.	Elevated effect levels.

Comparisons of the datasets show that class II, i.e. elevated frequencies of broods containing >5% late dead larvae and malformed larvae can be found in several areas which are not regarded as reference sites (Figure 4).

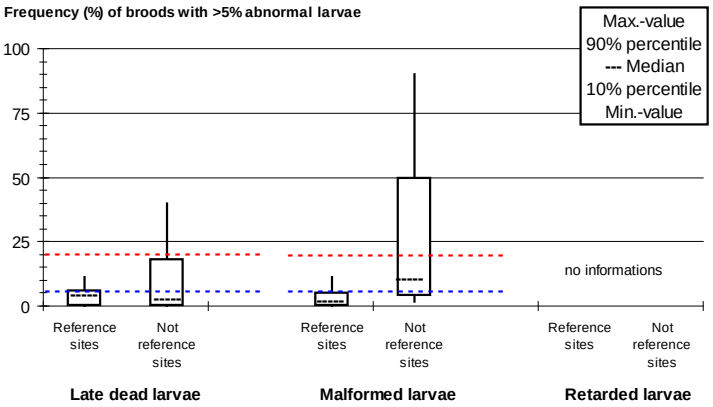


Figure 4. Comparison of data distribution of data on frequencies of broods containing >5% late dead larvae and malformed larvae in from reference sites and area not regarded as reference sites. The blue dotted line refers to the 90% percentile of data from the reference sites. The red dotted line refers to significantly elevated levels compared with the 90% percentile of the reference sites.

The assessment criteria including the existing data material and the statistical analyses will be evaluated and updated by 2011 as part of a BONUS+-project, called Balcofish.

Conclusions

The use of reproductive success of eelpout with focus on the occurrence of abnormal developed embryo and larvae in the broods seems to be a potential tool for assessing environmental impact on fish reproduction, because differences have been shown between areas regarded as reference sites and not.

Proposals for two assessment classes of effect levels (I and II) have been derived based on the 90% percentile of the datasets of mean frequencies as well as broods containing >5% of late dead larvae, malformed larvae and growth retarded larvae, respectively.

These assessment criteria seem especially useful for the data consisting of occurrences of late dead larvae and malformed larvae, where significantly elevated levels can be found in several areas not regarded as reference sites, whereas the occurrence of growth retarded larvae may be less useful.

References

BEQUALM. 2000. Fish Reproductive Success – embryo development and survival in viviparous blenny *Zoarces viviparus*. Biological Effects Quality Assurance in Monitoring Programmes (BEQUALM). BEQUALM Newsletter 2000. http://www.uni-kiel.de/ftzwest/downloads/BQ-Newsletter2.pdf.

Blekingekustens Vattenvårdsförbund. 2005. Hanöbugten 2000–2002. http://www.bkvf.org/huvudsida.htm (In Swedish).

Bodammer, J.E. 1993. The teratological and pathological effects of contaminants on embryonic and larvae fishes exposed as embryos: a brief review. Am. Fish Soc. Symp. 14: 77–84.

Davis, W.P. 1997. Evidence for developmental and skeletal responses as potential signals of endocrine disrupting compounds in fishes. *In* Chemically induced alterations in functional development and reproduction of fishes. Edited by Rolland, R.M., Gilbertson, M., Peterson, SETAC Press, Pensacola FL.

Fagerholm B. 2002. Integrerad fiskövervakning i kustreferensområden, Fjällbacka. Årsrapport för 2001. Fiskeriverket, Kustlaboratoriet. Arbetsrapport. 10p. (In Swedish).

Gercken J., Förlin L., Andersson J. 2006. Developmental disorders in larvae of eelpout (*Zoarces viviparus*) from German and Swedish Baltic coastal waters. Mar Pollut Bull. 53(8-9):497–507.

HELCOM. 2006. Manual for Marine Monitoring in the COMBINE Programme of HELCOM. Part D. Programme for monitoring of contaminants and their effects. http://www.helcom.fi/groups/monas/CombineManual/en_GB/Contents/, updated 31.03.2006.

ICES. 2004. Recommended techniques for biological monitoring programmes at the national or international level – methods for fish. Report of the Working Group on Biological Effects of Contaminants. WGBEC meeting 22–26 March 2004 Ostend, Belgium. www.ices.dk/reports.

Jacobsson, A., Neuman, E., Thoresson, G. 1986. The viviparous blenny as an indicator of environmental effects of harmful substances. Ambio 15(4): 236–238.

Kalmarweb. 2005. Samordnad kustvattenkontroll i Kalmar län. www.smhi.se/miljo/kalmarweb (In Swedish).

Larsson, D.G.J, Forlin, L. 2002. Male-biased sex ratios of fish embryos near a pulp mill: Temporary recovery after a short-term shutdown. Environmental Health Perspective 110(8): 739–742.

Neuman, E., Sandström, O., Thoresson, G. 1999. Guidelines for coastal fish monitoring. National Board of Fisheries, Institute of Coastal Research, Ôregrund, Sweden, 44p.

OSPAR. 1997. JAMP guidelines for general biological effects monitoring. Technical Annex 10 – reproductive success in fish. www.ospar.org.

Sjölin, A., Olsson, P., Magnusson, G. 2003. Hälsotillstånd och fortplantning hos tånglake i recipienten til Mörrums bruk hösten 2002. Toxicon rapport 8/03. http://www.bkvf.org/Rapporter/Rapporter/Tånglake/2003/toxicon.pdf (In Swedish).

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

Strand, J, Dahllöf, I. 2005. Biologisk effektmonitering i fisk. NOVANA teknisk anvisning for marin overvågning, kapitel 6.3. Danmarks Miljøundersøgelser. http://www2.dmu.dk/1_Om_DMU/2_Tvaerfunk/3_fdc_mar/programgrundlag/tekanv/tekniskanv.asp (In Danish).

Strand, J. 2005a. Biologisk effektmonitering i havsnegle og fisk. Marine workshop 2004. Research Note from NERI No. 210, 34p. http://www.dmu.dk/Udgivelser/Arbejdsrapporter (In Danish).

Strand, J. 2005b. Bioligiske effekter i ålekvabbe og muslinger. Marine områder 2004. Status og trends i NOVANA overvågning. Faglig rapport fra DMU. www.dmu.dk. (In Danish).

Vetemaa, M. 1999. Reproduction biology of the viviparous blenny (*Zoarces vivparus* L.). Rep. Nat. Board Fish. Swed. 2: 81–96, Fiskeriverket, Göteborg, Sweden.

Vetemaa, M., Förlin, L., Sandström, O. 1997. Chemical effluent impacts on reproduction and biochemistry in a North Sea population of viviparous blenny (*Zoarces viviparus*). J. Aquat. Ecosys. Stress Rec. 6: 33–41.

Ådjers, K., Appelberg, M., Eschbaum, R., Lappalainen, A., Lozys, L. 2001. Coastal fish monitoring in Baltic reference areas 2000. Kala- ja riistaraportteja 229: 1–14, Helsinki, Finland. Agenda item 3 at the HELCOM MONAS meeting 22–24 March 2004 in Tallinn, Estonia. www.helcom.fi.

Annex 14: Measurement of supporting metrics for fish: condition indices, GSI, HSI

Version date: 29 January 2010; draft version.

Remark: this BG document needs further work. there is confusion about the somatic weight to be used in the various calculations. Total weight, gutted weight (only extracting stomach content? or removing guts completely in all cases), somatic weight without all viscera?) We need to back to the original datasets to evaluate the most appropriate method).

Background

For all biological effect techniques within the OSPAR JAMP there is a requirement to report supporting parameters, and these include fish length, whole fish weight, liver weight and gonad size. The measurement of gonad size and liver weight is used to provide an indication of reproductive state and liver weight may also give an indication of general health and well-being. These measurements are used in indices relating gonad weight to whole body weight (GSI) and liver weight to whole body weight (LSI or HIS), explanations of these are described below. Both gonad and liver weight will change markedly throughout the year and for comparative purposes these seasonal variations must be taken into account for the interpretation of biomarker responses such as EROD and VTG for example.

ICES WGBEC recently reviewed the measurement of these metrics and their role and importance in fish monitoring programmes and this is described below.

General Overview: Organ size and adjacent measurements as biomarkers

Organ sizes constitute a very elementary biomarker. The measurements can be performed with a minimum of equipment, and the procedures are easy to undertake. At least for some species it is possible to analyse these variables on frozen material. With minimum instruction these biomarkers can be performed among personnel not regularly involved in biomarker analysis, although it is preferable to use personnel familiar with handling fish and able to perform simple dissection of fish.

Data of this type may be of relevance either in their own right, indicating adverse effects of various kinds where the toxic mechanisms are not fully understood as a result of xenobiotic exposure and/or, partly as a supporting variable to biomarkers conducted at the tissue, cellular and subcellular levels. As for all biomarkers in use today there is a strong need for quality assurance when these measurements are carried out. For instance abiotic factors of concern are the time of the year (photoperiod), water temperature and fishing technique. Among the biotic factors there are species, sub population, feeding, gender, size, age, developmental stage, natural parasitic infections as well as other diseases.

One of the most important biomarker in this field may be the development of gonads among female fish. This variable is best expressed as gonad size relative to the somatic body weight (Gonad Somatic Index - GSI) and expressed as a percentage value. The best species to use are those where the gonads of juvenile and immature fish are different from adult fish and where there are distinct differences in the genders. For

example, it is much easier when the morphology of the female ovary is a single structure while the male testes are paired bilaterally.

This offers the opportunity to investigate when the fish in relation to size and/or age are sexually immature or adult, or indeed have retarded gonad development (often termed sexually immature-SIM) as compared with normal sexual development. This can be expressed as a percentage of sexually immature females among the adult females, and represents the portion of fish with the extreme low value of the GSI value (usually below ~1%) and they have therefore a gonad with no or neglected development.

Analogous to the analysis of the gonad size is the liver size relative to the somatic body weight (Liver Somatic Index-LSI, or sometimes referred to as Hepato Somatic Index-HSI). It may be regarded as a biomarker in its own right and also as a supporting variable for other biomarkers such as EROD.

Furthermore, growth (e.g. gramme/year) as shown in Kiceniuk and Khan, 1986; McMaster *et al*., 1991 and in Ericson *et al*., 1998, as well as the Condition Factor (CF), are relatively straightforward to perform and may be used as biomarkers for adverse effects due to xenobiotic exposure. The measurement of condition factor has not often been used as a valuable biomarker in short exposure laboratory experiments, however, field observations over longer time periods indicate that it may be a valuable biomarker (see review by van der Oost *et al*., 2003.)

Feeding status in fish may be reflected in the condition factor, and may be important for a number of different biomarker responses, and as such can be included in biomonitoring investigations. Similarly visceral lipid stores are an additional variable that has been suggested as having biomarker potential (McMaster *et al*., 1991). For the calculation of the condition factor (CF) the following formula is used for dab, flounder and cod: 100 × (gutted somatic weight in milligrams) ÷ (total length in cubic millimetres)] Other sentinel species may require slightly altered formula's for the calculation of CF.

Gonad size in fish–GSI

The reproductive process constitutes (one of) the most essential health signals for the individual animal, and when missing or impaired indicates an obvious risk for adverse effect both genetically and for population survival. Therefore, decreased sizes of the gonad, of one or both of the genders, indicate an apparent risk for a reduced reproductive potential.

Gonad size is measured as a percentage of somatic body weight, gonadosomatic index (GSI*), It has been demonstrated to be a variable that can be influenced by contaminants in a number of different polluted field studies. It should be underlined that the toxicological response observed for this variable could have originated from a number of different toxicological reasons such as, tissue or cell death to more sophisticated regulatory endocrine mechanisms.

Measurement of GSI: record whole body weight of fish and gonad weight to two decimal places.

*GSI = (gonad weight x 100)/(total body weight[#]-gonad weight)

[#]subtract, if any, stomach content when significant

Deviation in GSI levels could represent a permanent effect or impairment for the reproductive cycle for one or more years (Janssen *et al*., 1997; Vallin *et al*., 1999). Both

scenarios will seriously affect reproductive potential. Examples of different pollution gradients were reduced gonads have been observed are in bleached kraft pulp mill effluents (Andersson *et al*., 1988; Sandström *et al*., 1988; McMaster *et al*., 1991; Balk *et al*., 1993; Förlin *et al*., 1995), including using chlorine-free processes (Karels *et al*., 2001) and general pollution (Johnson *et al*., 1988; Noaksson *et al*., 2001). Laboratory exposure experiments where effect on the GSI value has been documented include petroleum mixtures (Truscott *et al*., 1983; Kiceniuk and Khan, 1986), specific PAHs (Thomas, 1988; Singh, 1989; Thomas and Budiantara, 1995), PCB mixture (Thomas, 1988), pesticides (Ram *et al*., 1986; Singh, 1989), and cadmium (Singh, 1989; Pereira *et al*., 1993).

There is no doubt that xenobiotics can affect gonad size through a number of different toxicological mechanisms. However, as for most biomarkers, a variable that shows a (annual) natural biological cycle it is essential that the normal background values are well known, and that the appropriate control material is used for comparison. For the GSI value it should be pointed out that during certain time times of the year the gonad development is very fast and that different GSI values are obtained only within a period of a few days/weeks. Analysis of the GSI in these time periods should be avoided. Baseline studies are important in order to evaluate suitable time periods for this variable (Förlin and Haux, 1990; Larsen *et al*., 1992).

A state of complete disruption of sexual maturation reflects an extreme situation of low GSI values, e.g. a state of condition when the adult (based on age and/or size) fish are unable to develop from the prepubertal condition to the sexually mature stage. Field observations demonstrating a delay or lack of gonad development has been observed include the following species; burbot (*Lota lota*) in the north coast of the Bothnian bay (Pulliainen *et al*., 1992), English sole (*Parophrys vetulus*) in generally polluted areas in Puget sound, USA (Johnson *et al*., 1988), perch (*Perca fluviatilis*) in the effluent water from pulp and paper mills in Baltic waters (Sandström *et al*., 1988; Sandström *et al*., 1994) as well as white sucker (*Catostomus commersoni*) in corresponding effluents in Ontario, Canada (McMaster *et al*., 1991). Studies have also shown that perch, roach (*Rutilus rutilus*), and brook trout (*Salvelinus fontinalis*) exposed to leachate from a public refuse dump in a Swedish fresh-water system show corresponding adverse effects (Noaksson *et al*., 2001; Noaksson *et al*., 2002). Although the above cited field investigations are not all related suspected PAH contamination, these kinds of disorders has been created in laboratory experiment using petroleum products and a pure naphthalene (Thomas and Budiantara, 1995).

GSI confounding factors

Although the measurement is robust and easy to perform there is a need to characterize and avoid confounding factors. For example female perch populations, do not naturally spawn every year and the spawning frequency are affected by water temperature as indicated in Luksiene *et al*., 2000 and Sandström *et al*., 1995. Moreover, in the closely related yellow perch (*Perca flavescens*) both photoperiod and temperature been suggested to be of importance (Dabrowski *et al*., 1996). Therefore, for the interpretation of GSI data it is important to have prior knowledge of the reproductive cycle and triggers for gametogenesis for each species under investigation.

Liver size of female and/or male fish–LSI (HSI)

Liver size is measured in relation to somatic body weight, and is known as Liver Somatic Index (LSI* or HSI (see above)).

Measurement of LSI: record whole body weight of fish and gonad weight to two decimal places.

*LSI = (liver weight x 100)/(total body weight[#]-gonad weight)

[#]subtract, if any, stomach content when significant

LSI may be regarded as a relevant biomarker because it has been documented to be affected by contaminants in a number of different polluted field studies. For example, in pollution gradients of paper and pulp mill effluents where increased LSI vales were observed (Andersson *et al*., 1988; Lehtinen *et al*., 1990; Hodson *et al*., 1992; Kloepper-Sams and Owens, 1993; Huuskonen and Lindström-Seppa, 1995; Förlin *et al*., 1995), as well as decreased LSI levels as reported by Balk *et al*. (1993), and Förlin *et al*. (1995). Other complex effluents shown to affect liver size in various fish species are: leakage water from public refuse dumps (Noaksson *et al*., 2001; 2002) and effluent from wastewater treatment plant (Kosmala *et al*., 1998).

Field situations where PAHs and/or organochlorines are suspected contaminants for increased liver size in various fish species are documented by: Sloff *et al*. (1983); Goksoyr *et al*. (1991); Kirby *et al*. (1999); Kirby *et al*. (1999); Beyer *et al*. (1996); Leadly *et al*. (1998); Stephensen *et al*. 2000). Laboratory experiments shown to affect liver size among different fish species from exposure to organochlorines have been documented by: Adams *et al*. (1990); Newsted and Giesy (1993); Otto and Moon (1995); Arnold *et al*. (1995); Gadagbui and Goksoyr (1996); Åkerblom *et al*. (2000), and for two-stroke outboard engine exhaust extract (Tjärnlund *et al*., 1996) and PAHs (Celander *et al*., 1994) as well as pesticides (Singh, 1989; Åkerman *et al*., 2003) and cadmium (Singh, 1989).

LSI Confounding factors

Although there is no doubt that xenobiotics could affect liver size as a result of different toxicological mechanisms but it should be emphasized that, as for most biomarkers, control/reference fish should be analysed in close/direct parallel with the exposed site(s). In addition, seasonal variation is observed in different fish species (Koivusaari *et al*., 1981; Förlin and Haux, 1990; Larsen, 1992), and must be taken into account at all times. Besides the time of the year, factors (i.e. parameters) such as feeding behaviour, gender, maturity, age, size, temperature (George *et al*., 1990), photoperiod, parasites, among others, needs to be taken into considerations. Baseline studies are an important strategy to finally evaluate confounding factors (Balk *et al*., 1996).

References

Adams, S.M., Shugart, L.R., Southworth, G.R., and Hinton, D.E. 1990. Application of bioindicators in assessing the health of fish populations experiencing contaminant stress. In: McCarthy, J.F., Shugart, R.L. (Eds.). Biomarkers of Environmental Contamination. Lewis Pub., pp 333–353.

Åkerblom, N., Olsson, K., Berg, A.H., Andersson, P.L., Tysklind, M., Förlin, L., and Norrgren, L. 2000. Archives of Environmental Contamination and Toxicology 38:225–233.

Åkerman, G., Amcoff, P., Tjärnlund, Fogelberg, K., Torrissen, O. and Balk, L. 2003. Paraquat and menadione exposure of rainbow trout (*Oncorhynchus mykiss*) –Studies of effects on the pentose-phosphate shunt and thiamine levels in liver and kidney. Chemico-Biological Interactions 142, 269–283.

Andersson, T., Förlin, L., Härdig, J., and Larsson, Å. 1988. Physiological disturbances in fish living in coastal water polluted with bleached kraft pulp mill effluents. Canadian Journal of Fisheries and Aquatic Science 45: 1525–1536.

Arnold, H., Pluta, H.-J., and Braunbeck, T. 1995. Simultaneous exposure of fish to endosulfan and disulfoton in vivo: ultrastructural, stereological and biochemical reactions in hepatocytes of male rainbow trout (*Oncorhynchus mykiss*). Aquatic Toxicology 33: 17–43.

Balk, L., Andersson, T., Förlin, L., Söderström, M., and Larsson, Å. 1993. Indications of regional and large-scale biological effects caused by bleached pulp mill effluents. Chemosphere 27: 631–650.

Balk, L., Larsson, Å. and Förlin, L. 1996. Baseline studies of biomarkers in the feral female perch (*Perca fluviatilis*) as tools in biological monitoring of anthropogenic substances. Marine Environmental Research, 42 (1–4), 203–208.

Beyer, J. 1996. Fish biomarkers in marine pollution monitoring; evaluation and validation in laboratory and field studies. Academic thesis, University of Bergen, Norway.

Celander, M., Näf, C., Broman, D., and Förlin, L. 1994. Temporal aspects of induction of hepatic cytochrome P450 1A and conjugating enzymes in the viviparous blenny (*Zoarces viviparus*) treated with petroleum hydrocarbons. Aquatic Toxicology 29: 183–196.

Dabrowski, K., Ciereszko, R.E., Ciereszko, A., Toth, G.P., Christ, S.A., El-Saidy, D., and Ottobre, J.S. 1996. Reproductive physiology of yellow perch (*Perca flavescens*): Environmental and endocrinological cues. Journal of Applied Ichtyology :12: 139–148.

Ericson, G., Lindesjöö, E. and Balk, L. 1998. "DNA adducts and histopathological lesions in perch (*Perca fluviatilis*) and northern pike (*Esox lucius*) along a polycyclic aromatic hydrocarbon gradient on the Swedish coastline of the Baltic Sea." Canadian Journal of Fisheries and Aquatic Sciences 55, 815–824.

Förlin, L., Andersson, T., Balk, L., and Larsson, Å. 1995. Biochemical and physiological effects in fish exposed to bleached kraft mill effluents. Ecotoxicology and Environmental Safety 30: 164–170.

Förlin L, Haux C. Sex differences in hepatic cytochrome P-450 monooxygenase activities in rainbow trout during an annual reproductive cycle. J Endocrinol 1990; 124: 207–213.

Fulton, T. 1902. Rate of growth of sea fishes. Sci Invest. Fish. Div. Scot. Report 20.

George S, Young P, Leaver M, and Clarke D. Activities of pollutant metabolizing and detoxification systems in the liver of the plaice, *Pleuronectes platessa*: sex and seasonal variations in non-induced fish. Comp Biochem Physiol 1990; 96C: 185–192.

Gadagbui, B.K.-M., and Goksoyr, A. 1996. CYP1A and other biomarker responses to effluents from a textile mill in the Volta river (Ghana) using caged tilapia (*Oreochromis niloticus*) and sediment-exposed mudfish (*Clarias anguillaris*). Biomarkers 1: 252–261.

Goksoyr and 10 co-authors. 1991. Environmental contaminants and biochemical responses in flatfish from the Hvaler archipelago in Norway. Archives of Environmental Contamination and Toxicology 21: 486–496.

Hodson and 8 co-authors. 1992. Effects of bleached kraft mill effluent on fish in the St. Maurice river, Quebec. Environmental Toxicology and Chemistry 11: 1635–1651.

Huuskonen, S., and Lindström-Seppä, P. 1995. Hepatic cytochrome P4501A and other biotransformation activities in perch (*Perca fluvitialis*): the effects of unbleached pulp mill effluents. Aquatic Toxicology 31: 27–41.

Janssen, P.A.H., Lambert, J.G.D., Vethaak, A.D., and Goos, H.J.Th. 1997. Environmental pollution caused elevated concentration of oestradiol and vitellogenin in the female flounder, *Platichthys flesus* (L.). Aquatic Toxicology 39: 195–214.

Johnson, L.L., Casillas, E., Collier, T.K., McCain, B.B., and Varanasi, U. 1988. Contaminant effects on ovarian development in English sole (*Parophrys vetulus*) from Puget Sound, Washington. Canadian Journal of Fisheries and Aquatic Science 45: 2133–2146.

Karels, A., Markkula, E., and Oikari, A. 2001. Reproductive, biochemical, physiological, and population responses in perch (*Perca fluviatilis* L.) and roach (*Rutilus rutilus* L.) downstream of two elemental chlorine-free pulp and paper mills. Environmental Toxicology and Chemistry 20: 1517–1527.

Kiceniuk, J.W., and Khan, R.A. 1987. Effect of petroleum hydrocarbons on Atlantic cod, *Gadus morhua*, following chronic exposure. Canadian Journal of Zoology 65:490–494.

Kirby, M.F., Matthiessen, P., Neall, P., Tylor, T., Allchin, C.R., Kelly, C.A., Maxwell, D.L., and Thain, J.E. 1999. Hepatic EROD activity in flounder (*Platichys flesus*) as an indicator of contaminant exposure in English estuaries. Marine Pollution Bulletin 38:676–686.

Kirby, M.F., Neall, P., and Tylor, P. 1999. EROD activity measured in flatfish from the area of the Sea Empress oil spill. Chemosphere 38: 2929–2949.

Kloepper-Sams P.J. and Owens J.W. 1993. Environmental biomarkers as indicators of chemical exposure. Journal of Hazardous Material. 35: 283–294.

Kosmala, A., Migeon, B., Flammarion, P., and Garric, J. 1998. Impact assessment of a wastewater treatment plant effluent using the fish biomarker ethoxyresorufin-O-deethylase: field and on-site experiments. Ecotoxicology and Environmental Safe 41: 19–28.

Koivusaari U, Harri M, Hanninen O. Seasonal variation of hepatic biotransformation in female and male rainbow trout, *Salmo gairdneri.* Comp Biochem Physiol, 1981; 70C: 149–157.

Larsen HE, Celander M, Goksøyr A. The cytochrome P450 system of Atlantic salmon (*Salmo salar*): II. Variations in hepatic catalytic activities and isozyme patterns during an annual reproductive cycle. Fish Physiol Biochem 1992; 10: 291–301.

Leadly, T.A., Balch, G., Metcalfe, C.D., Lazar, R., Mazak, E., Habowsky, J., Haffner, D. 1998. Chemical accumulation and toxicological stress in three brown bullhead (*Ameiurus nebulosus*) populations of the Detroit River, Michigan, USA. Environmental Toxicology Chemistry 17: 1756–1766.

Lehtinen, K.-J., Kierkegaard, A., Jakobsson, E., Wändell, A. 1990. Physiological effects in fish exposed to effluents from mills with six different bleaching processes. Ecotoxicology and Environmental Safe 19: 33–46.

Luksiene, D., Sandström, O., Lounasheimo, L. and Andersson, J. 2000. The effects of thermal effluent exposure on the gametogenesis of female fish. Journal of Fish Biology 56: 37–50.

McMaster, M.E., Van Der Kraak, G.J., Portt, C.B., Munkittrick, K.R., Sibley, P.K., Smith, I.R. and Dixon, D.G. 1991. Changes in hepatic mixed-function oxygenase (MFO) activity, plasma steroid levels and age at maturity of a white sucker (*Catostomus commersoni*) population exposed to bleached kraft pulp mill effluent. 1991. Aquatic Toxicology 21: 199–218.

Newsted, J.L., and Giesy, J.P. 1993. Effect of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) on the epidermal growth factor receptor in hepatic plasma membranes of rainbow trout (*Oncorhynchus mykiss*) Toxicology and Applied Pharmacology 119: 41–51.

Noaksson E, Tjärnlund U, Bosveld ATC, Balk L. Evidence for endocrine disruption in perch *Perca fluviatilis* and roach *Rutilus rutilus* in a remote Swedish lake in the vicinity of a public refuse dump. Toxicol Appl Pharmacol 2001; 174: 160–176.

Noaksson, E., Linderoth, M., Bosveld, A.T.C., Norrgren, L., Zebuhr, Y. and Balk, L. 2002. "Endocrine disruption in brook trout (*Salvelinus fontinalis*) exposed to leachate from a public refuse dump." The Science of the Total Environment, In Press.

Otto, D.M.E., and Moon, T.W. 1995. 3,3′,4,4′-tetrachlorbiphenyl effects on antioxidant enzymes and glutathione status in different tissues in rainbow trout. Pharmacology and Toxicology 77: 281–287.

Pereira, J.J., Mercaldo-Allen, R., Kuropat, C., Luedke, D., and Sennefelder, G. 1993. Effect of cadmium accumulation on serum vitellogenin levels and hepatosomatic and gonadosomatic indices of winter flounder (*Pleuronectes americanus*). Archives of Environmental Contamination and Toxicology 24: 427–431.

Ram, R.N., and Sathyanesan, A.G. 1986. Effect of a mercurial fungicide on the gonadal development of the teleostean fish *Channa punctatus* (Bloch). Ecotoxicology and Environmental Safety 11: 352–360.

Sandström O. Incomplete recovery in a coastal fish community exposed to effluent from a modernised Swedish bleached kraft mill. Can J Fish Aquat Sci 1994; 51: 2195–2205.

Sandström, O., Neuman, E., and Karås, P. 1988. Effects of bleached pulp mill effluent on growth and gonad function in Baltic coastal fish. Water Science Technolology 20: 107–118.

Sandström O, Neuman E, Thoresson G. Effects of temperature on life history variables in perch, *Perca fluviatilis*. J Fish Biol 1995; 47: 652–670.

Singh, H. 1989. Interactions of xenobiotics with reproductive endocrine functions in a Protogynous teleost, *Monopterus albus*. Marine Environmental Research 28: 285–289.

Sloff W., van Kreijl, C.F., Baars, A.J. 1983. Relative liver weights and xenobiotic-metabolizing enzymes of fish from polluted surface waters in the Netherlands. Aquatic Toxicology: 4: 1–14.

Stephensen, E., Svavarsson, J., Sturve, J., Ericson, G., Adolfsson-Erici, M., and Förlin. 2000. Biochemical indicators of pollution exposure in shorthorn sculpin (*Myoxocephalus scorpius*), caught in four harbours on the southwest coast of Iceland. Aquatic Toxicology: 48: 431–442.

Thomas, P. 1988. Reproductive endocrine function in female Atlantic croaker exposed to pollutants. Marine Environmental Research 24: 179–183.

Thomas, P., and Budiantara, L. 1995. Reproductive life history stages sensitive to oil and naphthalene in Atlantic Croaker. Marine Environmental Research 39: 147–150.

Truscott, B., Walsh, J.M., Burton, M.P., Payne, J.F., and Idler, D.R. 1983. Effect of acute exposure to crude petroleum on some reproductive hormones in salmon and flounder. Comparative Biochemistry and Physiology 75C: 121–130.

Tjärnlund, U., Ericson, G., Lindesjöö, E., Petterson, I., Åkerman, G. and Balk, L. 1996. Further studies of exhaust from two-stroke outboard engines on fish. Marine Environmental Research, 42: 267–271.

Vallin, L., Nissling, A., and Westin, L. 1999. Potential factors influencing reproductive success of Baltic cod, *Gadus morhua*: a review. Ambio 28: 92–99.

Van der Oost, R., Beyer, J., and Vermeulen, N.P.E. 2003. Fish bioaccumulation and biomarkers in environmental risk assessment: a review. Environmental Toxicology and Pharmacology 13: 57–149.

Annex 15: Stress on Stress (SoS) in bivalve molluscs

Version date: 29 January 2010.

Background

Stress on Stress (SoS) is defined as the reduction of survival of bivalves in air after removal from the sea.

Contaminant exposure may alter the ability of organisms to survive environmental stress (De Zwaan *et al*., 1995; Viarengo *et al*., 1995). Laboratory and field studies have demonstrated the applicability of anoxic/aerial survival as an early warning indicator of contaminant induced stress. The effects of xenobiotics, including heavy metals, organometals and organics, as well as contaminated field sediments, on survival in air in invertebrates have been demonstrated (De Zwaan *et al*., 1996). Bivalve molluscs have been used in most studies, with marine mussel (*Mytilus* sp.*)* being the most common organism (Viarengo *et al*., 1995; Eertman *et al*., 1995; Smaal *et al*., 1991; Veldhuizen-Tsoerkan *et al*., 1991).

The reduction of survival in air, or Stress on Stress (SoS) is a simple and low cost whole organism response and can show pollutant induced alterations in the organism's physiology that render the animal more sensitive to further environmental changes. The method for determining SoS of mussels is being applied routinely to both toxicant-exposed mussels in laboratory studies and mussels collected in national monitoring programmes from polluted environments and along pollution gradients.

Laboratory studies have been conducted to establish relationships between toxicant concentrations in tissue and SoS. For example, it was demonstrated that short-term exposure to sublethal concentrations (less than μM) of pollutants, such as Cu^{2+}, DMBA (9,10-dimethyl 1,2 benzanthracene), Aroclor 1254, significantly reduced the capacity of mussels to survive in air. This effect was markedly dose-dependent, and was strongly increased by pollutant mixtures (Viarengo *et al*., 1995). The accuracy of air exposure as a monitoring tool has been reported to reflect smaller differences between contaminant groups than other physiological measurements in mussels, such as byssal thread production rate (Moles and Hale, 2003). The measurement of survival in air appeared also to be a sensitive and statistically significant parameter for monitoring the effect of long-term exposure to crude oil (Thomas *et al*., 1999).

Short description of methodology

Bivalve molluscs can survive for a long time in air, but individuals stressed by pre-exposure to pollutants show greater mortality than controls or individuals collected from a reference location. Both caged and native mussels can be used to assess the SoS response. The size of individuals for survival profiles must be selected from frequency distributions of the whole population under study. The individuals must be of a size approximating to the mean shell length for the population. When mussels are collected from the intertidal zone, it is important to sample them when they are submerged (i.e. just before they are uncovered as the tide recedes or conversely when they are covered as the tide comes in).

For spatial/temporal studies, the same size range should be selected (ideally 4–5 cm). Forty mussels (four replicates of ten) are used for each determination of SoS. The mussels must be collected from the sampling site and immediately transported to the

laboratory in insulated containers at a temperature of 5–10ºC, in humid conditions (e.g. damp paper or seaweed). Information that must be recorded includes:- the total number of animals sampled; the date and time of sampling; sampling location and position (e.g. Lat–Long); and seawater temperature.

Upon arrival at the laboratory, the 40 mussels are selected, placed on filter paper in a humidity chamber with continuous humidity of approximately 100%, at a temperature of 15–18°C. Mortality is recorded daily, each 24 hour after the time of sampling, until 100% of mortality is reached. This may take up to 25 days. Mussels are considered alive when closed individuals resist forcible valve separation. Dead mussels are always removed from the chamber and the humidity chamber cleaned and refreshed daily with clean filter paper. The Lethal Threshold for 50% mortality (LT50) and Time To Maximum Mortality (TTM), both in days is reported.

Confounding factors

The water and air temperature at the time of sampling should not be extreme i.e. collected when environmental temperatures are close to zero or above 25°C as this may influence the measurement. As a supporting parameter, condition index (CI) should be measured; spawned out mussels with a low CI tend to be weak and will die quickly when measured for SoS. If information on spawning state is not known, then do not undertake SoS at or immediately after the main spawning season. Tolerance of small mussels to air exposure has been demonstrated to be significantly greater than large mussels (Thomas *et al*., 1999). To date, there is no evidence to suggest that there will be differences in SoS response for different species of mussels or hybrids.

Applicability across the OSPAR maritime area

The SoS was also successfully utilized as a biomarker of stress in biomonitoring programmes such as RAMOGE (Mediterranean sea). To date, SoS has not routinely been applied in the OSPAR maritime area, except for some research activities (i.e. Labarta *et al*., 2005) and the ICON ICES/OSPAR Demonstration Programme (currently in progress). However, mussels are available throughout the OSPAR area, and there are no apparent significant constraints on its use. Several recent papers have emphasized the importance of utilizing this simple biomarker to evaluate the effects of aromatic polycyclic hydrocarbons, as for Exxon Valdez (Thomas, 1999) or in the Halifax Harbour biomonitoring (Hellou and Law, 2003), as well as in studying the effects of the pollutants present in untreated sewage (Moles and Hale, 2003).

SoS has been adopted as general stress biomarker in the UNEP MAP Mediterranean Biomonitoring Program UNEP/MAP (UNEP, 2003). However, it can easily and equally be applied in the OSPAR maritime area as a whole organism biomarker.

The measurement does not require sophisticated equipment, is low cost in terms of manpower to undertake the work. Mussels can be collected from the shoreline or close to the shore, avoiding the high cost of research vessels. Therefore, the applicability across the OSPAR maritime area is highly recommended. Most importantly, it can be used as an index of a general stress syndrome within the integrated mussels monitoring framework proposed by OSPAR. Furthermore, SoS shows a sensitivity which is in the same range of other commonly used general stress indices at the cellular level (e.g. lysosomal membrane stability).

For the new organization of biological effects monitoring in Phase IV of the MEDPOL Programme, a two-tier approach has been proposed. This approach considers Ly-

sosomal Membrane Stability (LMS), stress on stress (SoS) and mortality as core biomarkers to be applied in the first tier.

Ecological relevance

This extremely simple biomarker is able to provide evidence of effects of pollutants at the whole organism response level.

The response of this biomarker shows a typical dose-response curve, characterized by a continuous decrease of the parameter (LT50) with increasing pollutant concentrations, although in some experiments in the presence of low concentrations of contaminants a slight increase was also observed, possibly due to an hormetic effect (Eertman, 1995).

Quality assurance

LT50 values have been reported to show comparability with stress indices determined at the cellular level (Hellou and Law, 2003).

Because of the simplicity of the method, data quality assurance has not been tested by national or international programmes and is not considered to be necessary (ICES WGBEC 2010, pers com). However, an intercalibration/workshop exercise for other mussel techniques has been proposed in 2010 and it is proposed to include SoS in this initiative in order to harmonize the SOP and identify any AQC issues.

Background responses and assessment criteria

Background response times and response thresholds corresponding to unintended/unacceptable levels of response have yet to be defined for SoS. Until more data are available, then values should be interpreted from existing national datasets. It should be noted that these values are provisional and require further validation.

- Animals may be considered healthy if SoS is higher than 10 days;
- Animals may be considered to be stressed but compensating if SoS is between 5 and 10 days;
- Animals may be considered as severely stressed if the SoS is less than five days.

Background SoS responses may be as high as 18 days observed in *M. galloprovincialis*, (Personal communication, C. Martínez-Gómez, ES) and 16 days *Mytilus edulis* (Personal communication, John Thain, UK).

The added value of SoS in mussels is that the response measures the overall impact of multiple stressors on an organism, yet the response can be correlated quantitatively to contaminant tissue concentrations, a "true" integrated biological effect – chemical monitoring tool (see ecological relevance above).

References

Bayne, B.L. 1986. Measuring the effects of pollution at the cellular and organism level. In: Kullemberg G, editor. The Role of the Oceans as a Waste Disposal Option. New York, NY: Riedel, 617–634.

Eertman, R.H.M., Groenink, C.L.F.M.G., Sandee, B., Hummel, H. 1995. Response of the blue mussel *Mytilus edulis* following exposure to PAHs on contaminated sediments. Mar. Environ. Res. 39: 169–173.

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

Labarta, U., Fernández-Reiriz, M.J., Garrido, J.L., Babarro, J.M.F., Bayona, J.M., Albaigés, J. 2005. Response of mussel recruits to pollution from the 'Prestige' oil spill along the Galicia coast. A biochemical approach. Marine Ecological Progress Series 302: 135–145.

Moles, A., Hale, N. 2003. Use of physiological responses in *Mytilus trossul*us as integrative bioindicators of sewage pollution. Mar. Poll. Bull. 46, 954–958.

Moles, A., Hale, N. 2003. Use of physiological responses in *Mytilus trossulus* as integrative bioindicators of sewage pollution. Marine Pollution Bulletin, 46:954–958.

Smaal, AC., Wagenvoort, A., Hemelraad, J., Akkerman, I. 1991. Response to stress of mussels (*Mytilus edulis*) exposed in Dutch tidal waters. Comp. Biochem. Physiol. 100C, 197–200.

Thomas, R.E., Harris, P.M., Rice, S.D. 1999. Survival in air of *Mytilus trossulus* following long-term exposure to spilled Exxon Valdez crude oil in Prince William Sound. Comp Biochem Physiol C 122: 147–152.

UNEP. 2003. Mediterranean Action Plan. Strategy for the development of Mediterranean Marine Pollution Indicators. Meeting of the Med Pol National Coordinators. Sangemini, Italy, 27– 30 May 2003. UNEP(DEC)/MED WG.231/17.

Veldhuizen-Tsoerkan M. B., Holwerda, D. A., Zandee, D. I. 1991. Anoxic survival time and metabolic parameters as stress indices in sea mussels exposed to cadmium or polychlorinated biphenyls. Arch. Environ. Contam. Toxicol. 20, 259–265.

Viarengo A., Canesi, L., Pertica, M., Mancinelli, G., Accomando, R., Smaal, A.C., Orunesu, M. 1995. Stress on Stress response: A simple monitoring tool in the assessment of a general stress syndrome in mussels. Marine Environmental Research, 39: 245–248.

Viarengo, A., Canesi, L., Pertica, M., Mancinelli, G., Accomando, R., Smaal, A.C., Orunesu, M. Stress on stress response: a simple monitoring tool in the assessment of general stress syndrome in mussels. Mar. Environ. Res. 39: 245–248; 1995.

de Zwaan, A. and 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 Part C: Pharmacology, Toxicology and Endocrinology 113(2): 299–312.

de Zwaan, A., Cortesi, P., Cattani, O. 1995. Resistance of bivalves to anoxia as a response to pollution-induced environmental stress. Sci Total Environ 171:121–125.

Annex 16: Technical Annex on sampling and analysis for integrated chemical and biological effects monitoring in fish and shellfish

Version date: 29 January 2010.

Introduction

ICES/OSPAR WKIMON and associated groups have progressively developed an integrated approach to the use of biological effects and chemical measurements in environmental monitoring and assessment to meet the objectives of the OSPAR Strategy for Hazardous Substances. In relation to hazardous substances, the OSPAR Joint Assessment and Monitoring Programme seeks to addresses the following questions:

- What are the concentrations in the marine environment, and the effects, of the substances on the OSPAR List of Chemicals for Priority Action ("priority chemicals")? Are they at, or approaching, background levels for naturally occurring substances and close to zero for manmade substances?
- Are there any problems emerging related to the presence of hazardous substances in the marine environment? In particular, are any unintended/unacceptable biological responses, or unintended/unacceptable levels of such responses, being caused by exposure to hazardous substances?

Integration of chemical and biological effects measurements in OSPAR CEMP

The primary means of addressing these questions on an OSPAR wide basis is the Coordinated Environmental Monitoring Programme (CEMP; OSPAR Agreement 2005-5). Guidelines for the Integrated Monitoring and Assessment of Contaminants and their Effects were presented to ASMO 2007 (ASMO 07/6/8).

The integrated approach described in the Guidelines is been based around recommendations of sets of measurements that could be used to investigate the effects of contaminants on either fish or shellfish (mussels). These reflect the wide experience of the monitoring of the concentrations of priority contaminants in sediment and biota, and the benefits of combining this with the developing experience of the use of biological effects measurements in monitoring programmes. The fish and shellfish integrated monitoring schemes are reproduced below (Figures 1 and 2) from the JAMP Guidelines for the Integrated Monitoring and Assessment of Contaminants and their Effects (as revised at ICES/OSPAR SGIMC 2010).

As indicated in the Guidelines, the contribution made by an integrated programme, involving both chemical and biological effects measurements, is primarily that the combination of the different measurements increases the interpretive value of the individual measurements. For example, biological effects measurements will assist in the assessment of the significance of measured concentrations of contaminants in biota or sediments. When biological effects measurements are carried out in combination with chemical measurements (or additional effects measurements) this will provide an improved assessment due to the possible identification of the substances contributing to the observed effects.

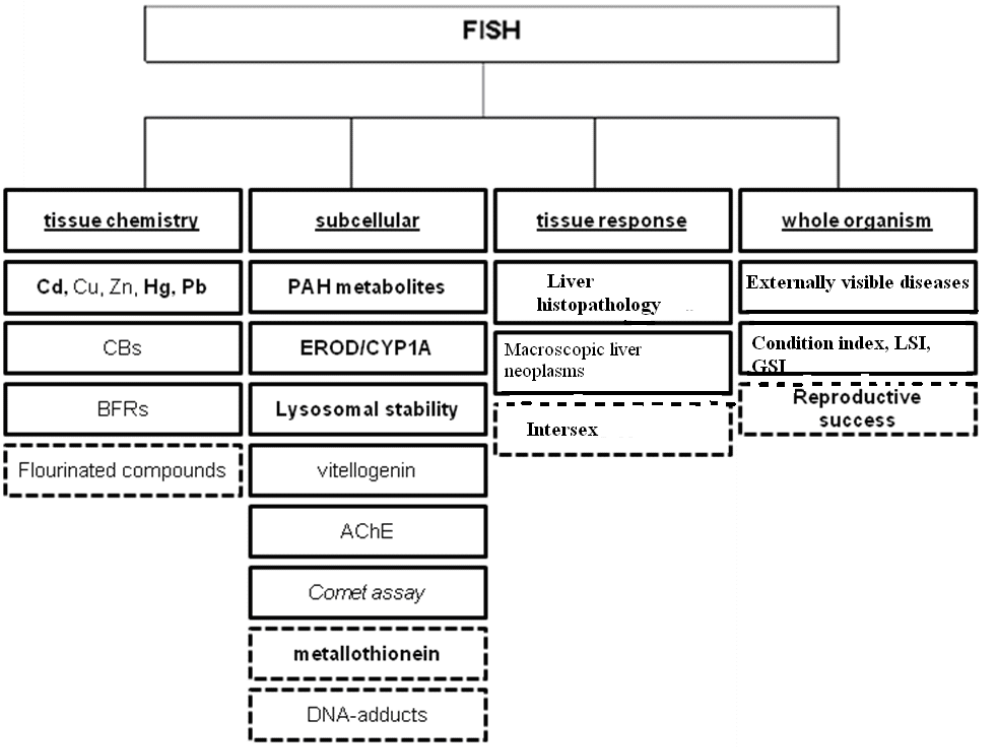


Figure 1. Overview of methods to be included in an integrated programme for selected fish species. (Bold: included in CEMP; solid-line boxes: prioritized components (only applies to tissues and subcellular responses); italics: ICES WGBEC promising method.

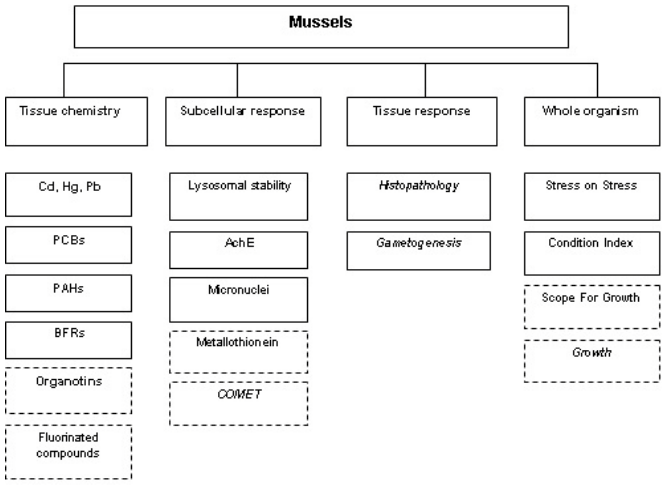


Figure 2. Overview of methods to be included in an integrated programme for selected blue mussel. (Bold: included in CEMP; solid-line boxes: prioritized components (only applies to tissues and subcellular responses); italics: ICES WGBEC promising method.

The structure of each of the schemes recognizes that a fully integrated assessment requires the integration of a variety of chemical measurements (concentrations of contaminants in the fish or mussels) and biological effects data.

It is well recognized that some particular contaminants or groups of contaminants can have characteristic biological effects. The classic example of a highly specific response to a contaminant is that of the effects of tributyltin (TBT) compounds in inducing imposex or intersex in gastropod mollusc species. These responses have been

widely used as an assessment of the environmental significance of tributyltin compounds, and are the topic of an OSPAR EcoQO. While it is theoretically possible for other substances to disrupt the hormonal systems of snails in a similar way, it is generally accepted that TBT is the primary marine contaminant responsible for the effects.

There is clearly great attraction in the recognition of a highly specific response to a particular narrow class of contaminants, particularly if chemical analysis at concentrations known to be associated with the effects is difficult. However, generally such close relationships are rare. For example, a range of effects measurements have been applied to the effects of planar organic contaminants in the sea, i.e.

- the concentration of PAH-metabolites in fish bile;
- CYP1A/EROD induction;
- Indices of genotoxicity (e.g. DNA adducts of PAH, COMET assay, micronucleus assay, etc.);
- liver (microscopic) neoplasms;
- liver histopathology.

However, these effects show varying degrees of specificity for PAH as opposed to other planar organic contaminants such as planar CBs, or dioxins. The concentration of PAH-metabolites in fish bile is clearly specific to the PAH compounds detected, but CYP1A/EROD induction is a property of a range of groups of compounds.

In general, it is found that while subcellular responses can commonly be linked to substances that have the potential to induce the response, measurements of whole organism effects are much less contaminant-specific. However, they are often more closely linked to the potential for cause effects at population level, through reduction in survival or reproductive capacity. This gradation is reflected in the grouping of the effects measurements in Figures 1 and 2 under the headings of subcellular responses, tissues responses and whole organism responses. Sub-cellular responses such as EROD, bile metabolite concentrations and metallothionein are recognized as biomarkers of exposure to contaminants, while whole organism and tissue level responses are more clearly markers of effect.

Sampling and analysis strategies for integrated fish and bivalve monitoring

The integration of contaminant and biological effects monitoring requires a strategy for sampling and analysis that includes the

1) sampling and analyses of same tissues and individuals;

2) sampling of individuals for effects and chemical analyses from the same population as that used for disease and/or population structure determination at a common time;

3) sampling of water, the water column and sediments at the same time and location as collecting biota; and

4) more or less simultaneous sampling for and determination of primary and support parameters (e.g. hydrographic parameters) at any given location.

Examples of sampling strategies for the integrated fish and shellfish schemes are shown in Figures 3 and 4). The numbers of individual organisms required are driven primarily by the assessment of external diseases and macroscopic liver nodules (fish)

and histopathology (bivalves), because these require the largest number of individuals. A subsample of individuals within the primary sample is further sampled for liver histopathology (fish) and biomarkers (fish and bivalves) to meet Requirements 1 and 2 above.

In the specified target species, further subsampling of the same individuals for chemical analysis is often restricted by insufficient remaining tissue, e.g. liver in fish. In order to meet Requirement 2, subsamples for chemical analysis are taken from the same combined hauls/population as those for disease/biomarkers.

In order to integrate sediment, water chemistry and associated bioassay components, with the fish and bivalve schemes, sediment and water samples should be collected at the same time as fish/bivalve samples and from a site or sites that are representative of the defined station/sampling area.

Additional integrated sampling opportunities may arise from trawl/grab contents, for example, gastropods for imposex or benthos, and these should be exploited where possible/practicable.

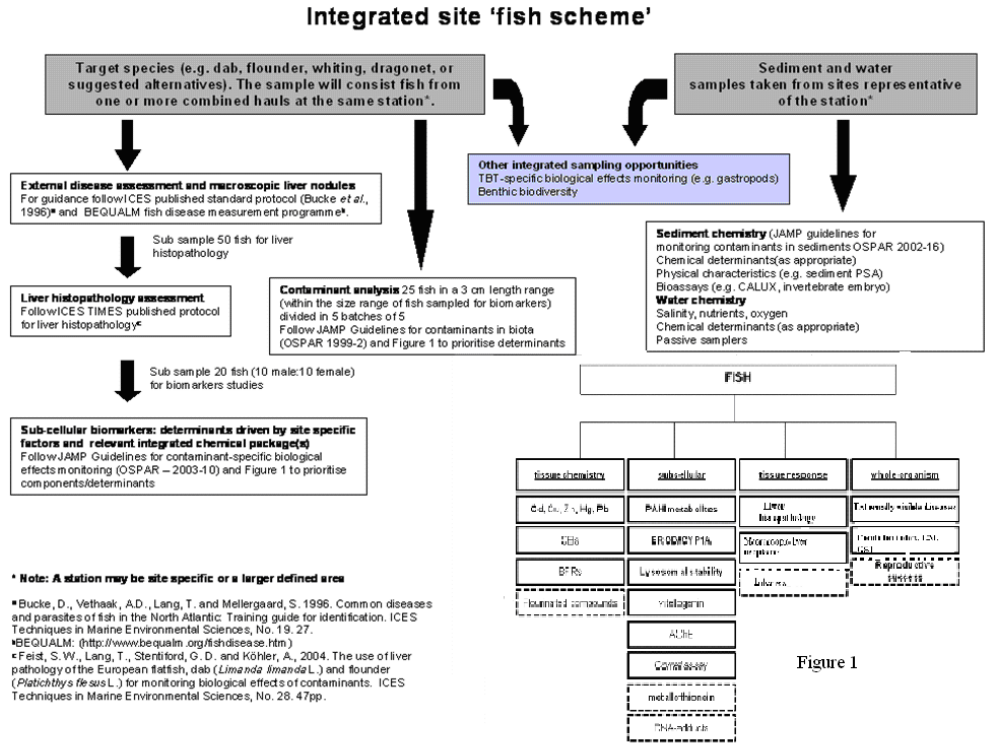


Figure 3. Sampling strategy for integrated fish monitoring.

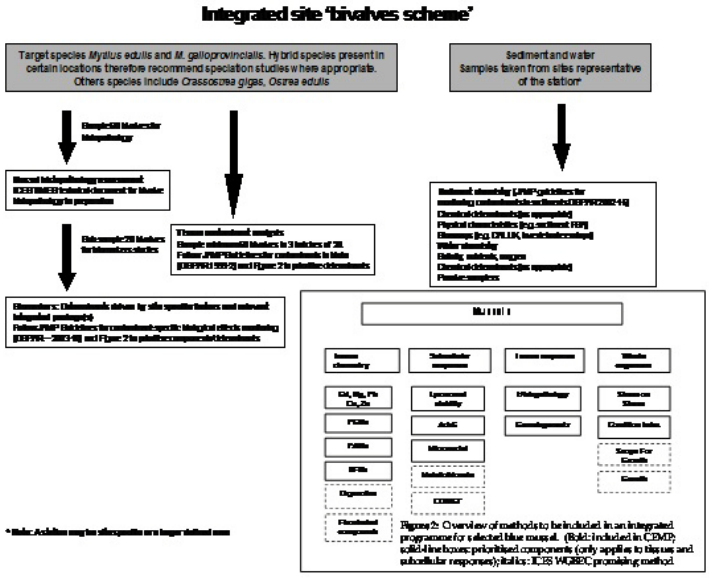


Figure 4. Sampling strategy for integrated bivalve monitoring.

Guidance on sampling and analysis for integrated monitoring of biological effects and chemical measurements

Some aspects of the details of fish and shellfish sampling and analysis are covered in the OSPAR JAMP Guidelines. Integration of chemical and biological effects data in coordinated monitoring programmes was not a primary consideration when the components of these Guidelines were developed. Some revisions have therefore been made to ensure that the information correctly covers the requirements for integrating chemical and biological effects sampling.

The following tables address aspects of technical guidance on sampling design and supporting parameters.

Tables 1–3 cover methods to be used for integrated fish, bivalve and gastropod monitoring, Tables 4 and 5 cover methods for monitoring of water and sediments.

A new scheme for organotin-related integrated monitoring using marine gastropods is added Figure 7.2.

Table 1. Overview of selected methods for integrated fish monitoring (2007 WKIMON Report, revised).

SUBJECT	PARAMETER	COMMENT
Species	Primary species: dab, flounder, Whiting, eelpout Alternative species: plaice, cod, herring, eelpout, hake, dragonet or other	Alternative species may be used if primary species are not available.
Sex	females and/or males	For certain biomarkers or chemical measurements, only females or only males are used (see relevant JAMP guidelines)
Health condition	Specimens free of external visible diseases should be used for chemical and biomarker analysis.	Certain biomarkers are affected by disease conditions.
Size ranges	Dab: ≥ 15 cm (according to suggested new JAMP guidelines for externally visible diseases). Flounder: ≥ 20 cm (according to suggested new JAMP guidelines for externally visible diseases). Whiting: ≥ 15 cm (according to suggested new JAMP guidelines for externally visible diseases). Dragonet: ≥ 10 cm (according to suggested new JAMP guidelines for liver histopathology) Eelpout: Pregnant females 15–30 cm , 50 fish per station.	For integrated monitoring encompassing chemistry, histopathology and biomarkers, the mid size groups are preferable which are: 20–24 cm (dab) 20–29 cm (flounder)20–24 cm (whiting)10–15 cm (dragonet).
Sample size	Depending on the parameter measured, according to JAMP Guidelines.	Sample sizes have to fulfill statistical requirements for spatial and/or temporal trend monitoring. Preferably, all measurements should be done in individual fish and pooling should be avoided (with the possible exception of contaminant measurements).
Sampling time and frequency	Sampling for all parameters should be carried out at the same time, outside the spawning season, and at least once a year in the same time window	Justification is provided in the OSPAR JAMP Guidelines
Sampling location	Sampling for all parameters should be carried out at the same site	The location, size and number of sampling sites depend on the purpose of the monitoring. For offshore sampling targeted at fish, it is recommended to use ICES statistical rectangles as sampling sites. A number of repeated samplings (= hauls) (replicates) should be carried out in each of these rectangles. For coastal and estuarine waters, sites should be selected based on existing WFD and other chemical/biological monitoring sites, taking account of potential hot-spot areas or areas at risk. The number of sampling sites should be sufficient to reflect the environmental conditions ij the survey area, and meet the purposes of the monitoring programme.
Chemical determinands	Metals: Hg, Cd, Pb, Cu, Zn CBs: ICES 7 CBs + CB77, CB81, CB126, CB169 + CB105, CB114, CB123, CB156, CB157, CB167, CB189. Brominated flame retardants: congeners of the penta mix, octa mix and deca mix PBDE formulations; hexabromocyclododecane, tetrabromobisphenol A. Lindane. TBT	In addition, in situ PAH measurements (e.g.., using UV fluorescence spectrometry) may be employed under specific circumstances (e.g. after oil spill or PAH related point source discharges). Besides the contaminants already covered by the OSPAR CEMP, there are a number of other compounds from the OSPAR List of Chemicals for priority action that should be monitored because of their toxicity and environmental relevance. The list provided is, therefore, not complete.

SUBJECT	PARAMETER	COMMENT
Biological effects measurements	Biological effect techniques as specified in the OSPAR Guidelines for the Integrated Monitoring and Assessment of Contaminants and their Effects, as in Figures 1 and 2 above	Additional opportunities for the inclusion of new methods is likely to emerge through the implementation of MSFD and as science develops. Potential examples include indicators of immunocompetence, and embryo-malformation.
Supporting parameters	Length, weight, gender, age, somatic indices, stage of gonadal maturation, grossly visible anomalies, lesions, parasites, hydrography (temperature, salinity, oxygen content)	In the list, parameters are provided that are known to affect both the biological effects responses and the concentration of contaminants. The data can be of assistance in data interpretation.
Haul duration	Haul durations should be harmonized between monitoring authorities. An appropriate value would be 30 minutes, but may be less than this if conditions require.	The purpose is to standardize the stress experienced by fish during capture
Duration and conditions of storage of live fish prior to dissection	Fish should be maintained alive in flowing seawater on the sampling vessel for periods not exceeding 8 hours.	Storage for longer periods or under poor conditions can stress the fish and alter some biomarker responses.

Table 2. Overview of selected methods for integrated shellfish monitoring (2007 WKIMON Report, revised).

SUBJECT	PARAMETER	COMMENT
Species	Primary species: *Mytilus edulis* Alternative species: *Mytilus galloprovincialis, Crassostrea gigas, Ostrea edulis*	The first choice shellfish species is not available in all parts of the OSPAR area. In such cases, other species should be selected, such as oysters. For *Mytilus* sp., speciation studies are recommended in order to confirm species identity.
Sex	Females and/or males	For certain biomarkers or chemical measurements, only females or only males are used (see relevant JAMP guidelines)
Size range	Mussel: >=40 mm, ideally in the range between 40-55mm. Pacific oyster: 9–14 cm	Based on JAMP Guidelines for chemical monitoring
Sample size	Depending on the parameter measured, according to JAMP Guidelines.	Sample sizes have to fulfil statistical requirements for spatial and/or temporal trend monitoring. For some parameters, sample size has still to be defined. Preferably, all measurements should be done in individual mussels and pooling should be avoided (except where recommended, for example for the measurement of contaminant concentrations).
Sampling time and frequency	Sampling for all parameters should be carried out at the same time, outside the spawning season, and at least once a year in the same time window	Justification is provided in the OSPAR JAMP Guidelines
Sampling location	Sampling for all parameters should be carried out at the same site.	The location, size and number of sampling sites depend on the purpose of the monitoring. For coastal and estuarine waters, sites should be selected based on existing sites used for WFD or other purposes, taking account of hot-spot areas and areas at potential risk. The number of sampling sites should be sufficient to reflect the environmental conditions ij the survey area, and meet the purposes of the monitoring programme. For coastal and offshore studies, caging of mussels should be considered.
Chemical determinands	Metals: Hg, Cd, Pb, Cu PAHs: EPA 16 + NPD CBs: ICES 7 + CB 77,81,126,169 + CB 105,114,123,156,157, 167,189 Brominated flame retardants: congeners of the penta mix, octa mix and deca mix PBDE formulations; hexabromocyclododecane, tetrabromobisphenol A. Lindane Organotin compounds	In addition, total hydrocarbon measurements (e.g.., using UV fluorescence spectrometry) may be employed under specific circumstances (e.g. after oil spill or PAH related point source discharges). Besides the contaminants already covered by the OSPAR CEMP, there are a number of other compounds from the OSPAR List of Chemicals for priority action that should be monitored because of their toxicity and environmental relevance. The list provided is not complete.
Biological effects measurements	Biological effect techniques as specified in the OSPAR Guidelines for the Integrated Monitoring and Assessment of Contaminants and their Effects, as in Figures 1 and 2 above	Additional opportunities for the inclusion of new methods is likely to emerge through the implementation of MSFD and as science develops. Potential examples include indicators of immuno-competence, and embryo-malformation.

Supporting parameters	Shell length, shell and soft body weight, gender, stage of gonadal maturation, grossly visible anomalies, lesions, parasites, sampling depth, hydrography (temperature, salinity, oxygen content, turbidity), nutrients/eutrophication	In the list, parameters are provided that are known to affect both the biological effects responses and the concentration of contaminants. The data can be of use for normalization.
Sampling depth	Subtidal or intertidal mussels can be used. Deployed mussels offshore can be positioned at depths 0–8m	Intertidal specimens may be subject to greater biomarker variability. Subtidal specimens are less robust post-sampling and effects measurements may be more susceptible to post-sampling stress.
Storage and transport of bivalves	Transport of bivalves should be completed within than 24 hours. They should be transported in an insulated container at 4°C in a damp atmosphere maintained by absorbent materials (such as seaweed and/or paper towel) wetted with seawater.	

Table 3. Overview of methods and species for integrated gastropod/organotin monitoring (2007 WKIMON Report, revised).

SUBJECT	PARAMETER	COMMENT
Species	Intertidal species: *Nucella lapillus* *Nassarius reticulata* *Littorina littorea* Offshore species: *Buccinum undatum* *Neptunea antiqua*	
Sex	Females and/or males	
Size range	Size ranges are to be selected in accordance with the JAMP Guidelines	
Sample size	Depending on the parameter measured, according to JAMP Guidelines.	All measurements should be done in individual gastropods and pooling should be avoided.
Sampling time and frequency	Sampling for all parameters should be carried out at the same time. Sampling frequency according to JAMP Guidelines.	
Sampling location	Sampling for all parameters should be carried out at the same site.	For coastal and estuarine waters, sites should be selected based on existing WFD sites (where they are established) and TBT hot-spot areas like harbours and major shipping routes (see relevant JAMP guidelines).
Chemical determinands	Organotin compounds in tissue	Guidelines for chemical measurements in biota will be published shortly in ICES TIMES series, and in a Technical Annex to the JAMP Guidelines.
Biological effects measurements	Imposex or intersex (species dependent endpoints, as in the JAMP Guideline) ICES TIMES document on intersex in *Littorina* provides methodological advice.	
Supporting parameters	Shell length, organotin compounds in sediment.	

Table 4. Environmental parameters for inclusion in monitoring programmes (water) (2007 WKIMON Report, revised).

SUBJECT	PARAMETER	COMMENT
Chemistry	Salinity, nutrients, oxygen	
Chemical determinands	Metals: Hg, Cd, Pb, Cu, Zn PAHs: EPA 16 + Naphthalene, phenanthrene, dibenzothiophene and their alkylated derivatives CBs: ICES 7 CBs Brominated flame retardants: congeners of the penta mix, octa mix and deca mix PBDE formulations; hexabromocyclododecane, tetrabromobisphenol A. Lindane Organotin compounds	Consideration should be given to bioavailability. To answer the JAMP question relating to concentrations approaching background or zero, there may be a requirement to measure a broader range of chemicals.
Physical	Temperature, content of suspended matter	
Biology	Phyto and zooplankton	Information might be useful for specific events, such as blooms affecting fish health

Table 5. Environmental parameters for inclusion in monitoring programmes (sediment) (2007 WKIMON Report, revised).

SUBJECT	PARAMETER	COMMENT
Chemistry	TOC, water content, Al, Li	Al and Li (or other elements as appropriate to the sediment type) are used for normalization of contaminant concentrations.
Chemical determinands	Metals: Hg, Cd, Pb, Cu, Zn PAHs: EPA 16 + Naphthalene, phenanthrene, dibenzothiophene and their alkylated derivatives CBs: ICES 7 CBs+ CB77, CB81, CB126, CB169 + CB105, CB114, CB123, CB156, CB157, CB167, CB189. Brominated flame retardants: congeners of the penta mix, octa mix and deca mix PBDE formulations; hexabromocyclododecane, tetrabromobisphenol A. Lindane Organotin compounds	Consideration should be given to bioavailability. To answer the JAMP question relating to concentrations approaching background or zero, there may be a requirement to measure a broader range of chemicals.
Physical	Sediment type, particle size, colour index, information on anthropogenic disturbances, sedimentation rates, current flow rates	Anthropogenic disturbance such as trawling or sand and gravel extraction may affect the sediment structure.

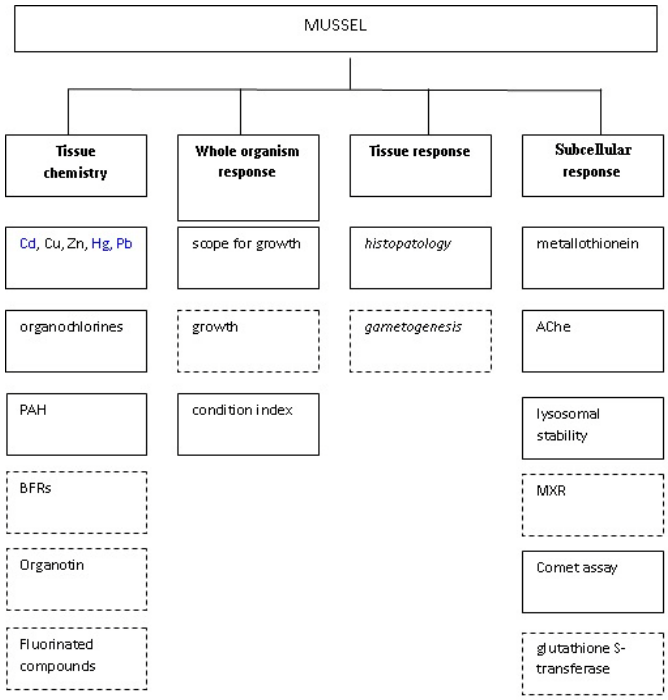


Figure 7.3.1. (p. 13 in 5.1 OSPAR DRAFT integrated guidelines.doc). ***Revised figure.***

Annex 17: Technical Annex for Mussel (*Mytilus* sp.) OSPAR Integrated Monitoring

Version date: 29 January 2010.

Background

The basis for the technical annex is the mussel integrated monitoring strategy incorporating biological effect techniques at the subcellular, tissue and whole organism responses and tissue chemistry. This is outlined below (Figure1).

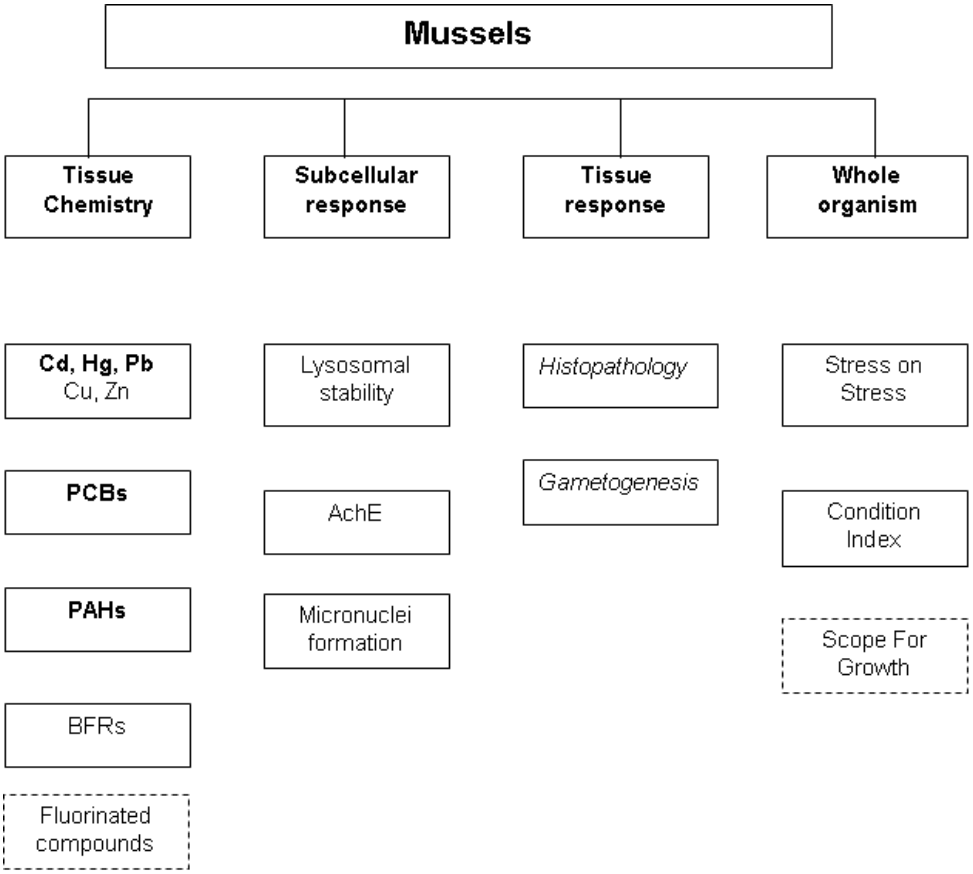


Figure 1. Overview of methods to be included in a programme for blue mussel; solid-lined boxes – core components. Bold text – included in CEMP; italics – recommended but additional supporting documentation required.

In any mussel integrated monitoring programme the core components as indicated should be included as a bare minimum.

Purpose of work

The integrated approach described above can be used for:

- *Status and trend monitoring*; contaminant and biological effect responses are measured over geographic areas and repeated over time. The purpose here may be to compare biological effect responses between sites, to compare changes in response with time and to observe if the "health status" is improving, at a steady state or declining.

- *Investigative monitoring*; most frequently used as a screening step to assess if biological effects are occurring in relation to a suspected contaminant gradient, pollution event or if biological effects are suspected for any reason (e.g. tissue chemical residues have been observed to be high).
- *Hot spot – site-specific monitoring*; usually in relation to risk assessment at pollution sites e.g. oil platform investigations.

Offshore and coastal

Mussels (*Mytilus* species) are infrequently found in the sub littoral zone. But populations do exist in shallow waters and are found on the seabed, usually close to the coastline, in general within the 12 mile limit. They may also be found offshore attached to navigation buoys, chains, and oil and gas platforms. For monitoring purposes these mussels can be used but care needs to be exercised in sampling the organisms, to ensure that they are not damaged during sampling and that the correct size range can be obtained. For offshore monitoring purposes it is usually more applicable to use *in situ* caging methods (see below). Advantages of using caged organisms are; choice of site deployment (including reference sites), selection of depth of deployment (e.g. may be critical for oil platform studies, but generally within 8 m of the sea surface); standardization of origin (same source/supply), size and species. Disadvantages are: cost of deployment in respect of mooring systems and ship time for deployment and retrieval; in addition some techniques require immediate sampling and analysis which may not be feasible on a research vessel offshore.

If caging is used then hydrographical conditions must be considered with special attention given to water currents and stratification.

Shoreline

Mussels may be regarded as ubiquitous on rocky shore coastlines and therefore, ideal for monitoring purposes. Sampling sites can be selected easily, organisms collected with little cost and reference sites located without difficulty. In addition, if mussels are not present at a site of interest then organisms can be caged on the sea shore or in estuaries on piers or similar structures.

Sampling Information

Details required

- Date, time and location on the shoreline (if applicable e.g. low water) and exposure (e.g. highly exposed Atlantic rocky shore or enclosed sheltered bay).
- Position in Lat. Long.
- Type of site; reference, pollution gradient, status or trend.
- At caging sites information on water temperature, depth of deployment, time of immersion, water column depth and information on currents and stratification if available, water temperature and salinity.
- Source of mussels for caging studies; for any caging study it is important that the mussels are sourced from a clean site, and that day 0 values are determined for tissue contaminant chemistry and biological effect responses.
- For shoreline monitoring, ideally the mussels must be sampled in a uniform manner between sites i.e. tidal height and similar salinity profile.

Confounding factors

For *in situ* transplants/caging the mussels must be deployed for at least three weeks in order to allow sufficient time for contaminants to accumulate in the tissues and reach a state of equilibrium. Failure to do this may produce spurious data. Also of note is that in many countries there are regulations controlling the movement and deposit of shellfish and these must be observed (i.e. prevention of transfer of disease).

Reproductive state and gametogenic cycle; Mussels generally spawn in early spring, with spawning occurring later in more northern populations. At spawning there is a major loss in body lipid and a subsequent fall in condition; therefore sampling in or shortly after this period should be avoided for all aspects of tissue chemistry analysis and biological effect determinations.

Salinity; be aware that low salinities affect the biomarker response, of particular importance for caging work in estuaries.

Temperature; Mussels on the shoreline can be subject to extremes of temperature, cold in winter and extreme heat in summer. Avoid sampling when extremes are likely to occur as this may compromise the biological effects response.

Parasites; Mussels with severe parasite infections should not be used.

Algal blooms; in spring and late summer and autumn intense algal blooms may occur and sampling of mussels at such times should be avoided.

Species; on some coastlines mussels are solely of one species whereas at other locations they are mixed or hybrids. It is unclear whether species difference will affect interpretation of data but wherever possible attempts should be made to determine the species under observation.

In caging studies (shoreline or offshore) care should be taken in sourcing mussels from a "clean site". If rope grown mussels are chosen then particular attention must be given to transporting the mussels as they tend to have weak adductor muscles and easily gape and become stressed during transportation which may give rise to initial mortalities or erroneous biological effect responses. Therefore, the source of mussels should be taken account of in the experimental design.

Supporting measurements

- Condition index; dry meat relative to whole live weight or internal shell volume.
- Gonad state; index of reproductive state.
- Lipid content; usually a determined and measured along with tissue chemistry and useful for interpretation of biomarker responses.
- Real growth; if available measured using growth of marked intervals over time, usually months.
- Water quality measurements; salinity, temperature are recommended, and where possible suspended solids or turbidity, DO, and chlorophyll.
- Chemical analysis of tissues; this is essential to interpretation of biological effects data and for the implementation of the integrated chemical biological effect strategy as outlined above. Prioritised contaminants are Cd, Cu, Hg, Zn, Cd, PAHs and PCBs. As a minimum 50 mussels (>40 mm in length) should be collected, taken to the laboratory and held in running seawater for 24 h to eliminate gut contents (e.g. sediment, etc.). The tissues

should then be extracted from the mussel and placed in acid washed hexane rinsed glass/plastic/metal containers (as appropriate to the particular analysis), stored at -20°C for subsequent chemical analysis using ICES or appropriate protocols.

Sampling for bio-effects

For some methods the samples require immediate processing at the time of sampling whereas for other techniques processing is undertaken in the laboratory. An overview of this is shown in the table below (Table 1), and also includes the number of animals typically sampled for each method. Ideally the size of individual mussels for all methods is >40 mm.

Table 1. Overview of sampling procedures for mussels.

Method and minimum numbers of animals usually sampled per site in brackets	When analytical sampling is undertaken	Acclimation	Comments and aspects that are crucial
SFG (10)	24 h	Ca 10 h	Crucial
Ache (10)	Immediate in field	Not applicable	Stored immediately in liquid nitrogen
Mt (10)	Any time within 24 h on live mussel	Not applicable	Take tissue sample – freeze in liquid nitrogen
COMET	Within 24 h	Store for no more than 24 h in cool damp conditions. Must be consistant in startegy	Do as quickly as possible
Micronuclei (20)	Within 3 days	None	Mussels can be kept out of water but cool
NRR (10)	Within 24 h	Store for no more than 24 h in cool damp conditions. Must be consistant in startegy	Do as quickly as possible
Lysosomal histochemical method (10)	Freeze immediately	Not applicable	In liquid nitrogen
Stress on stress (40)	Not applicable	Transport at low temperatures for no more than 24 h	Analysis done at 18°C
Histopathology and gametogenesis(30–50)	Sample immediately if possible	Anything more than 6 h delay in sampling place in water for 48 h acclimation	Dessication must be avoided, correct dissection to incude all organs
Condition (20)	Within 24 h	If > 24 h place in water and do sample within 48 hrs	
Growth (25)	Within 24 h	If > 24 h place in water and do sample within 48 hrs	
Tissue chemistry (50)	Place in 24 h clean running seawater	Not applicable	Depuration of sediment is crucial

Mussels are attached to each other or to a substratum by a byssal thread. When mussels are sampled care should be used not to pull the mussels and byssal threads too vigorously as this can damage and stress the mussels. If mussels have to be transported this should be kept to a minimum and they should be kept damp and cool and if possible the temperature logged during the transport.

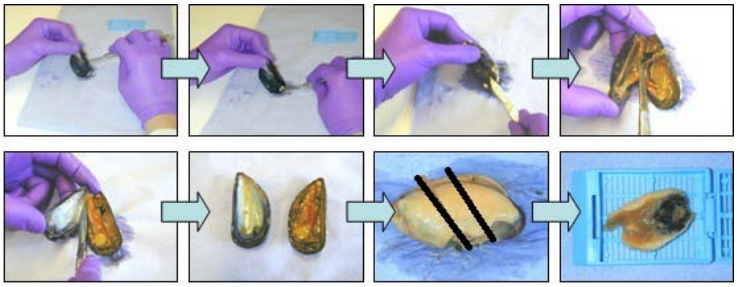
For some techniques such as SFG the mussels will need to be carefully cleaned. It should be noted that there are limitations of analysis for some methods e.g. for SFG

and NRR where time-wise it may be difficult to process more than two samples in a single day.

For histological sampling it is essential that the dissection is conducted in a precise manner and this is described below.

The technical procedure essential to correct mussel sampling for histology (taken from draft TIMES doc. under preparation, provided by J Bignell, UK, Cefas.

- Insert scalpel into ventral byssal cavity and move knife down so it cuts the posterior adductor muscle.
- Open shell and remove byssal thread.
- Remove mussel from one shell half. Repeat for remaining half.
- Analyse tissue for presence of parasites, pearls or other abnormalities.
- Obtain a standardized section as shown in photographs 1–8 in order to include all organs of interest in one section and place into histo-cassette.



- Samples should be preserved for a minimum of 24 hours in Bakers Formal Calcium, and subsequently transferred to 70% alcohol until processed.
- The correct ratio of mussels to fixative is 30 samples per 800 ml (approx.) of fixative. This is the recommended volume of fixative to ensure adequate fixation.
- Samples should be agitated periodically to ensure thorough fixation. A rocker plate facilitates this perfectly.

Methods to be used

These are listed in the mussel integrated strategy above. An overview of the methods is given in the table below (Table 8.2) with references to the analytical procedures.

Table 2. Overview of methods and reference to analytical procedure.

Method	Issue addressed	Biological significance	References
AChE inhibition	Organophosphates and carbamates or similar molecules Possibly algal toxins	Measures exposure to a wide range of compounds and a marker of stress.	1–2
Metallothionein induction	Measures induction of metallothionein protein by certain metals (e.g. Zn, Cu, Cd, Hg)	Measures exposure and disturbance of copper and zinc metabolism.	3–4
Lysosomal stability (including NRR)	Not contaminant-specific, but responds to a wide variety of xenobiotic contaminants and metals	Measures cellular damage and is a good predictor of pathology. Provides a link between exposure and pathological endpoints. Possibly, a tool for immunosuppression studies in white blood cells.	5–19
Scope for growth	Responds to a wide variety of contaminants	Integrative response, a sensitive sublethal measure of energy available for growth.	20–21
Stress on stress	Responds to a wide variety of contaminants and other environmental conditions	Integrative response , a measurew of stress, condition, health and well-being.	26
Micronuclei	Exposure to aneugenic and clastogenic	Exposure to aneugenic and clastogenic	22–23
Histopathology and gametogenesis	Not contaminant-specific	General responses	24–25++
COMET	Genotoxic compounds	DNA strand breaks	See OSPAR Background Document
Condition Index	Quality of tissue	Health status and stress	See OSPAR Background Document
Growth	Conditions of water quality including food availability	Health status and stress	OSPAR Background Document in preparation

Quality assurance

Wherever possible all analytical methods must be supported with quality assurance procedures. These should be through international intercalibration exercises where they exist and through internal quality controls.

The current position with quality assurance is:

- NRR – currently being developed across OPSAR, exists in MEDPOL, for internal QA a dual assessment with a colleague on the same samples is recommended.
- Ache – not yet developed but include internal standard
- Mt – MEDPOL have intercalibration exercises, elsewhere there have been *ad hoc* intercalibrations and additionally an internal standard should be included.
- SFG – none at present.
- Stress on Stress – none at present but will be addressed by MEDPOL/ICES workshop in 2010.

- Histology and gametogenesis – TIMES doc and circulation of reference material.
- Lysosomal histochemical procedures – none currently available but include an internal standard. In addition will be addressed by MEDPOL/ICES workshop in 2010.
- Micronuclei formation – currently being addressed through MEDPOL and may be extended to include a wider participation.
- COMET – none at present but being addressed through ICES WGBEC.

Reporting requirements

Biological effect responses; these should be reported in-line with requirements detailed in each analytical method. When different biological effect measurements are made on the same individual mussel then the data should be identified in the reporting and data assessment.

Contaminants: reported in line with standard analytical procedures.

Supporting parameters:

Essential; date and time of sampling, Lat. Long. position, organism length, whole weight, site characterization (e.g. position on shore, or caging, DO, salinity etc); for caged studies the source of organisms and duration of exposure.

Desirable; identification of species particularly if in a hybrid zone.

References

Al-Sabti K and Metcalfe C. 1995. Fish micronuclei for assessing genotoxicity in water. Mutat Res 343:121–135.

Bocquené, G., Bellanger, C, Cadiou, Y., and Galgani, F. 1995. Joint action of combinations of pollutants on the acetylcholinesterase activity of several marine species. Ecotoxicology, 4: 266–279.

Cajaraville, M.P., Abascal, I., Etxeberria, M., and Marigomez, I. 1995. Lysosomes as cellular markers of environmental pollution time-dependent and dose-dependent responses of the digestive lysosomal system of mussels after petroleum hydrocarbon exposure. Environmental Toxicology and Water Quality, 10: 1–8.

Cajaraville, M.P., Díez, G., Marigómez, J.A., and Angulo, E. 1990. Responses of the basophilic cells of the digestive gland of mussels to petroleum hydrocarbon exposure. Diseases of Aquatic Organisms, 9: 221–228.

Cajaraville, M.P., Marigómez, J.A., Díez, G., and Angulo, E. 1992. Comparative effects of the WAF of three oils on mussels. 2. Quantitative alterations in the structure of the digestive tubules. Comparative Biochemistry and Physiology, 102C(1): 113–123.

Cheung, V.V., Wedderburn, R.J., and Depledge, M.H. 1998. Molluscan lysosomal responses as a diagnostic tool for detection of a pollution gradient in Tolo Harbour, Hong Kong. Marine Environmental Research, 46: 237–241.

Etxeberria, M., Cajaraville, M.P., and Marigomez, I. 1995. Changes in digestive cell lysosomal structure in mussels as biomarkers of environmental stress in the Urdaibai Estuary (Biscay Coast, Iberian Peninsula). Marine Pollution Bulletin, 30: 599–603.

Fernley, P.W., Moore, M.N., Lowe, D.M., Donkin, P., and Evans, S. 2000. Impact of the *Sea Empress* oil spill on lysosomal stability in mussel blood cells. Marine Environmental Research, 50: 451–455.

Grisolia CK and Cordeiro CMT. 2000. Variability in micronucleus induction with different mutagens applied to several species of fish. Genetics and Molecular Biology 23(1):235–239.

Lin, S.Y., and Steichen, D.J. 1994. A method for determining the stability of lysosomal membranes in the digestive cells *of Mytilus edulis.* Marine Ecology Progress Series, 115: 237–241.

Lowe, D.M., and Fossato, V.U. 2000. The influence of environmental contaminants on lysosomal activity in the digestive cells of mussels *(Mytilus galloprovincialis)* from the Venice Lagoon. Aquatic Toxicology, 48: 75–85.

Lowe, D.M., Soverchia, C, and Moore, M.N. 1995. Lysosomal membrane responses in the blood and digestive cells of mussels experimentally exposed to fluoranthene. Aquatic Toxicology, 33: 105–112.

Marigomez, I., and Baybay-Villacorta, L. 2003. Pollutant-specific and general lysosomal responses in digestive cells of mussels exposed to model organic chemicals. Aquatic Toxicology, 64: 235–257.

Regoli, F. 1992. Lysosomal responses as a sensitive stress index in biomonitoring heavy-metal pollution. Marine Ecology Progress Series, 84: 63–69.

Ringwood, A.H., Conners, D.E., and Di Novo, A. 1998. The effects of copper exposures on cellular responses in oysters. Marine Environmental Research, 46: 591–595.

Ringwood, A.H., Connors, D.E., and Hoguet, J. 1998. Effects of natural and anthropogenic stressors on lysosomal destabilisation in oysters, *Crassostrea virginica.* Marine Ecology Progress Series, 166: 163–171.

Rickwood, C.J., and Galloway, T.S. 2004. Acetylcholinesterase inhibition as a biomarker of adverse effect. A study of *Mytilus edulis* exposed to the priority pollutant chlorfenvinphos. Aquatic Toxicology, 67: 45–56.

Ringwood, A.H., Conners, D.E., and Keppler, C.J. 1999. Cellular responses of oysters, *Crassostrea virginica,* to metal-contaminated sediments. Marine Environmental Research, 48: 427–437.

Ringwood, A.H., Conners, D.E., and Hoguet, J. 1998. Effects of natural and anthropogenic stressors on lysosomal destabilization in oysters *Crassostrea virginica.* Marine Ecology Progress Series, 166: 163–171.

Roesijadi, G., Unger, M.E., and Morris, J.E. 1988. Immunochemical quantification of metallothioneins of a marine mollusc. Canadian Journal of Fisheries and Aquatic Sciences, 45: 1257–1263.

Svendsen, C, and Weeks, J.M. 1995. The use of a lysosome assay for the rapid assessment of cellular stress from copper to the freshwater snail *Viviparus contectus* (Millet). Marine Pollution Bulletin, 31: 139–42.

Tremblay, R., Myrand, B., and Guderley, H. 1998. Temporal variation of lysosomal capacities in relation to susceptibility of mussels, *Mytilus edulis,* to summer mortality. Marine Biology, 132: 641–649.

Viarengo, A., Ponzano, E., Dondero, F., and Fabbri, R. 1997. A simple spectrophotometric method for metallothionein evaluation in marine organisms: an application to Mediterranean and Antarctic molluscs. Marine Environmental Research, 44: 69–84.

Viarengo, A., Lowe, D., Bolognesi, C., Fabbri, E., and Koehler, A. 2007.The use of biomarkers in biomonitoring: a 2-tiered approach assessing the level of pollutant-induced stress syndrome in sentinel organisms. Comp Biochem Phys. Accepted.

Widdows, J., and Salkeld, P. 1992. Practical procedures for the measurement of scope for growth. MAP Technical Reports Series, 71: 147–172.

Widdows, J., and Johnson, D. 1988. Physiological energetics of *Mytilus edulis:* Scope for growth. Marine Ecology Progress Series, 46(1-3): 113–121.

Annex 18: Technical annex on recommended packages of chemical and biological methods for monitoring on a determinant basis

Version date 29 January 2010.

Review of CEMP requirements

This agenda item was addressed by reviewing the chemical determinants listed in the OSPAR CEMP and pre-CEMP (ASMO, 2007a) and considering the most appropriate chemical analyses and biological effects techniques that could be applied in an integrated fashion to monitor for these compounds in the marine environment.

Some general points concerning integrated monitoring were noted during this process:

a) In some cases the list of contaminants that should be reported under the CEMP (and pre-CEMP) may be insufficient for an integrated approach. In order to aid interpretation of biological effects measurements, an integrated assessment may require data on related contaminants which would elicit a response on the biological effects components of the methods packages. Determinants additional to those required under the CEMP have therefore been added to the packages below.

b) It was felt that a fully 'integrated' approach to monitoring should include passive sampling of contaminants as part of the package of methods. This will provide information on availability of contaminants in sediments and allows for temporally integrated sampling of contaminants in water. (Guidelines for the application of passive samplers are available from ICES WGMS).

c) The biological effects techniques applied to these packages of methods are listed either in the ICES WGBEC recommended techniques list (WGBEC, 2007) or form part of the fish and shellfish methods packages proposed in the draft JAMP guidelines for integrated monitoring and assessment of contaminants and their effects (ASMO 2007b). The biological effects methods included here are separated into those appropriate to monitoring selected fish species, shellfish (mussels) and bioassays (sediment, water and *in vitro* tests).

d) It should be noted that the biological effects methods listed here are those which may form part of an overall integrated monitoring package and are likely to be affected by the OSPAR priority contaminants in question. Many of the effects measurements listed are 'general' biological effects which are indicative of stress or health status of marine organisms or general toxicity in the sediments and water column. These may be affected by a wide range of contaminants and are not specific to the contaminants in question. Therefore, for each group of substances the most specific and relevant biological effects techniques have also been highlighted.

e) These packages of methods should be considered supplemental to the existing JAMP guidelines for contaminant specific (OSPAR-2003–10) and general (1997–7) biological effects monitoring and the JAMP Guidelines on contaminants in biota (OSPAR 1999–2) and sediment (OSPAR 2002–16). The JAMP guidelines provide more detailed background on the biological effects and chemical analysis methods referred to here and the necessary

cofactors that should be recorded for these techniques. The packages of methods presented here combine contaminant-specific effects with the general biological effects methods that are likely to respond to the contaminants. They also deal with groups of contaminants not addressed by the contaminant specific guidelines and propose further integration of techniques such as passive sampling and invertebrate methods for metals.

The priority chemical determinants from the OSPAR CEMP and pre-CEMP are as follows (taken from ASMO, 2007a). The Appendices referred to are CEMP appendices.

The following components of the CEMP are to be measured on a mandatory basis:

- the heavy metals cadmium, mercury and lead in biota and sediment (Appendix 2);
- the PCB congeners CB 28, CB 52, CB 101, CB 118, CB 138, CB 153, and CB 180 in biota and sediment (Appendix 3);
- the PAHs anthracene, benz[a]anthracene, benzo[ghi]perylene, benzo[a]pyrene, chrysene, fluoranthene, ideno[1, 2, 3-cd]pyrene, pyrene and phenanthrene in biota and sediment (Appendix 4);
- TBT in sediment (biota voluntary/pre-CEMP) (Appendix 5).

The following components are currently part of the pre-CEMP and are to be measured on a voluntary basis:

- the brominated flame retardants HBCD and PBDEs 28, 47, 66, 85, 99, 100, 153, 154 and 183 in biota and sediment, and BDE 209 in sediment (Appendix 8);
- the planar PCB congeners CB 77, 126 and 169 in biota. Monitoring of those congeners in sediment should be undertaken only if levels of marker PCBs are e.g. 100 times higher than the Background Assessment Concentration (Appendix 9);
- the alkylated PAHs C1-, C2-, and C3-naphthalenes, C1-, C2- and C3-phenanthrenes, and C1-, C2- and C3-dibenzothiophenes and the parent compound dibenzothiophene in biota and sediment (Appendix 10);
- PFOS in sediment, biota and water (Appendix 12);
- Polychlorinated dibenzodioxins and furans in biota and sediment (Appendix 13);

Methods package for metals

Although cadmium, mercury and lead are the only mandatory metal determinants under the CEMP, other metal species are needed to interpret the biological effects data as part of an integrated package. Additional metal species needed include copper and zinc. Metals analysis should be performed on sediments and biota collected from the same times and locations where possible. Cofactors for sediment analysis are also required including aluminium and lithium. DGTs present the opportunity to undertake passive sampling for metal species to allow temporally integrated sampling of water and measure availability of metals in sediments.

Metal-'specific' biological effects measurements include metallothionein, ALA-D and oxidative stress, although both metallothionein and oxidative stress responses are known to be affected by other contaminants. ALA-D is lead-specific and can be measured in fish blood, although it has limited use/expertise across the ICES/OSPAR

community and it is recommended that it is applied only in areas where lead contamination is perceived to be a problem or where chemical monitoring indicates that concentrations are e.g. significantly above background.

ALA-D is relevant to fish only. Metallothionein can be applied to fish liver and mussel digestive glands although best results are obtained from mussels. There are a number of oxidative stress measurements that can be made in both fish and mussels which could add value to an integrated package of metals methods, but due to the lack of standardized methods, QA and assessment criteria it is suggested that this method is not an essential part of the metals package.

A number of 'general' biological effects measurements in fish and shellfish will be affected by environmental metal contamination and these are shown in Figure 8.1 below. *In vivo* bioassays are also relevant measurements for the effects of metals.

Metallothionein in mussels and ALA-D in fish are considered the most specific/relevant biological effects methods for metals.

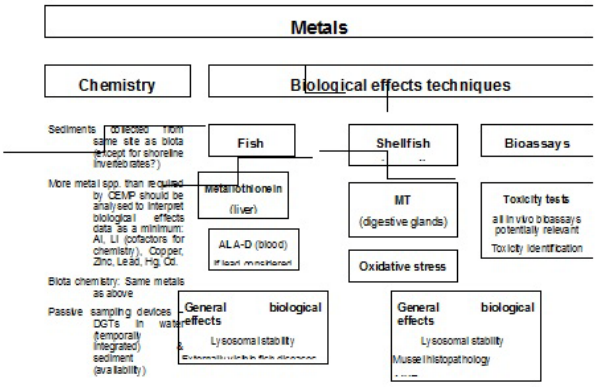


Figure 8.1. Package of chemical and biological effects methods relevant to monitoring for metals. The most specific/relevant biological effects methods are highlighted.

Methods package for PCBs, polychlorinated dibenzodioxins and furans

Due to the similarity of their toxicological effects, a single methods package was proposed for both PCBs and polychlorinated dibenzodioxins and furans. In addition to the OSPAR CEMP required determinants, additional CBs may cause biological effects and their analysis should be included in an integrated monitoring approach. These include co-planar CBs CB105 and CB 156. A variety of passive sampling devices (e.g. silicone rubber) offer the potential for temporally integrated sampling of these compounds from water and investigation of their availability in sediments and these should be employed where possible.

There are no truly specific biological effects measurements available for PCBs, polychlorinated dibenzodioxins and furans. The most relevant are considered to be induction of CYP1A/EROD activity in fish liver and application of the dioxin receptor based in vitro test, DR-CALUX.

Several other general biological effects measurements in fish and shellfish may respond to exposure to these compounds and are given below in Figure 8.2. DR-Calux is considered the most useful *in vitro* bioassay technique although chronic *in vivo* bioassays may also be relevant.

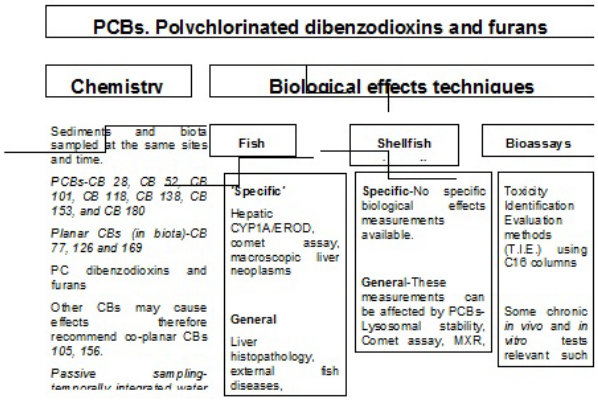


Figure 8.2. Package of chemical and biological effects methods relevant to monitoring for PCBs polychlorinated dibenzodioxins and furans. The most specific/relevant biological effects methods are highlighted.

Methods package for PAH and alkylated PAH

Due to similar toxicological effects, a single package of methods is proposed for PAH and alkylated PAH. The package of methods is similar to Figure 8.2 above although chemical determinants should be analysed in sediment and shellfish for biota only. Due to rapid metabolism in finfish, PAH should be analysed as metabolites in bile rather than parent compounds in liver or flesh. As above, passive sampling should also be applied where possible.

Additional specific biological effects are applicable for PAH/alkylated PAH. These include PAH metabolites in fish bile and DNA adducts in fish liver. The most relevant/specific biological effects techniques are highlighted as induction of hepatic CYP1A/EROD, DNA adducts and the DR-CALUX *in vitro* bioassay.

General biological effects measurements will also respond to exposure to these compounds and are given in Figure 8.3 below.

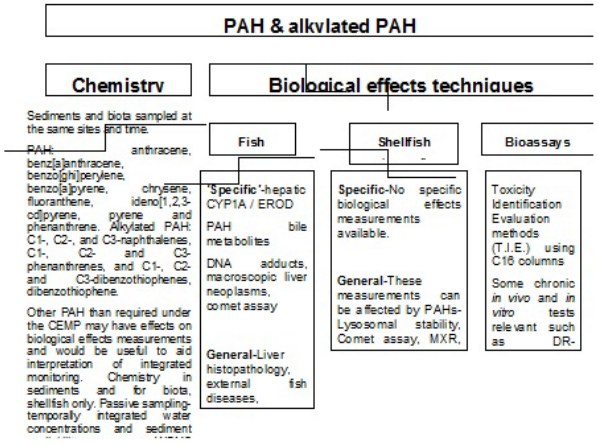


Figure 8.3. Package of chemical and biological effects methods relevant to monitoring for PAH and alkylated PAH. The most specific / relevant biological effects methods are highlighted.

Organotins

It was felt that the package of methods appropriate to organotin monitoring was already very well described by the JAMP guidelines on organotin-specific monitoring and included a suite of parameters relevant to imposex/intersex in gastropods, TBT, DBT, MBT, TPhT, DPhT, MPhT in sediments (for offshore monitoring) and in biota

where appropriate (voluntary). It was noted that passive sampling for organotins may become an option for integrated monitoring of organotins in future. It was also noted that bivalve embryo bioassays are sensitive to dissolved TBT at ng/L level.

BFRs

It was noted that there are currently very few biological effects methods available and tested in a monitoring context for measuring the effects of these compounds. The determinants required for CEMP are HBCD and PBDEs 28, 47, 66, 85, 99, 100, 153, 154 and 183 in biota and sediment, and BDE 209 in sediment. Passive sampling is also relevant as described above in Section 8.2.

There are no specific biological effects techniques available. Thyroid hormone receptor assays in fish blood are relevant but have not been well field tested, nor is this an ICES recommended technique. Recent studies on the toxicological properties of these compounds in fish suggest that there are limited overt effects that can be detected by existing techniques.

PFOS

PFOS analysis in sediment, biota and water is included in the list of pre-CEMP determinants, however no specific biological effects techniques are recommended here. It was noted that the compound may have endocrine disrupting effects and that some ED-relevant endpoints may be appropriate along with general biological effect measurements such as reproductive success. A battery of short-term low volume bioassays (*in vitro* and *in vivo*) using extracts can be used to perform a first screening/assessment of unintended impacts and novel contaminants (see background document on water bioassays). These extracts can be derived from water, sediment, biota and/or passive samplers. Information obtained from bio-analysis can also be used as input for the design of future monitoring programmes and the development of appropriate higher-level biological effects techniques biomarkers. However, a package of methods relevant to PFOS would require further consideration.

References

ASMO. 2007a. Proposal for revising the OSPAR Coordinated Environmental Monitoring Programme (CEMP). ASMO 07/13/1.

ASMO. 2007b Draft JAMP guidelines for integrated monitoring and assessment of contaminants and their effects. ASMO 07/6/8-E.

WGBEC. 2007. Report of the ICES Working Group on Biological Effects of Contaminants, 2007.

Annex 19: Discussion document on survey design for integrated chemical and biological effects monitoring

Version date: 29 January 2010.

Background

OSPAR requested that the joint OSPAR/ICES WKIMON group should develop a draft technical annex on the survey design for integrated chemical and biological effects monitoring be prepared, based on work anticipated to be carried out by ICES WGSAEM. The purpose of the technical annex would be to provide guidance on the selection of representative stations, taking into account requirements under the Water Framework Directive and the proposed Marine Strategy Framework Directive.

In particular, OSPAR recommended that this work should build on work by WGSAEM 2007 relating to the spatial design of monitoring programmes and should take into account the approach taken by the UK in re-designing their station network.

Sampling for a single parameter or for integrated monitoring both require definition of sampling positions, sampling times and the number of cases per sampling. The strategy to do this is essentially the same for both cases, so the general procedure for planning integrated monitoring can follow that for planning single parameter monitoring. Considerations on these steps are detailed in Section 3.

Discussions at ICES/OSPAR SGIMC 2010

The WKIMON Group was discontinued in 2008, and was replaced by the ICES/OSPAR SGIMC. SGIMC 2010 therefore reviewed the opportunities to progress the task to develop a technical annex on survey design.

Following WKIMON IV, SGIMC 2010 noted that survey design had been discussed only briefly by WGSAEM 2007, but it had concluded that it was not possible to take this item forward during the meeting. WGSAEM had not returned to this topic since 2007, and it was not clear that they would be able to do so in 2010. However, effective survey design is heavily dependent on statistical analysis and advice.

SGIMC 2010 discussed the opportunities for further development of advice on survey design. The UK approach to redesign involved the definition of monitoring and assessment regions, and the application of a stratified random sampling scheme within the regions (see Section 3.4 below). Since 2007, there has been an increased interest in assessment of monitoring data on regional bases, for example the presentation of contaminant (CEMP) monitoring data in the OSPAR QSR 2010 documents. The EU Marine Strategy Framework Directive has assessment regions (and subregions) as a core element of its assessment system. In anticipation of this, the Regional Conventions (OSPAR, HELCOM, MEDPOL, etc) are developing proposals for the definition of subregions within their areas.

SGIMC considered therefore, that the development of survey design, including subregions for monitoring and assessment was a considerably wider tissue than just integrated chemical and biological effects monitoring in the context of the OSPAR Hazardous Substances Strategy. For MSFD, it will be necessary to develop coherent and efficient monitoring programmes for a wide range of Descriptors of Good Environmental Status. SGIMC therefore agreed that the overall task was too large for them to address, but that it was possible for the group to offer comment and advice

on aspects of the statistical considerations that will be part of the wider programme of OSPAR to assist in the implementation of MSFD, specifically in the North East Atlantic.

In the light of the current implementation for MSFD and the availability of statistical advice from the relevant environmental WGs in ICES, SGIMC are of the opinion that it is not possible for them to progress this question of survey design further at this time. SGIMC recommend that this discussion document be forward to OSPAR as a contribution to the wider survey design task in relation to harmonization of OSPAR and MSFD programmes.

Some statistical considerations in integrated monitoring

The choice of sampling positions aims at obtaining a sample which reflects the variation of a parameter in the area of interest, i.e. to establish geographic representativeness. Section 3 describes two alternative sampling strategies, adopting either fixed sampling positions or a stratified random position sampling. The stratified random sampling scheme starts from the assumption that there are homogeneous regions, from which samples may be taken at random positions. Fixed position sampling avoids problems that could arise from inhomogeneity that was not anticipated, which could be a problem for stratified random sampling schemes. On the other hand, a fixed station may, by bad luck, be located at an inappropriate position, but will be reused as long as the monitoring programme continues. With random sampling, such a continuously bad positioning is unlikely, instead it can be expected that "good" and "bad" positions compensate another in the long run. These considerations apply to selecting positions for monitoring for a single quantity as well as for an integrated plan. They apply similarly to the choice of sampling times, though the latter is also driven by other considerations (inside / outside the spawning period, etc).

Survey design: general

Survey design is driven by the objectives of the sampling, which are (WKIMON III Report, p. 170).

- to assess status (existing level of marine contamination and its effect) and trends across the OSPAR maritime area;
- to assess the effectiveness of measures taken for the reduction of marine contamination;
- to assess harm (unintended/unacceptable biological responses) to living resources and marine life;
- to identify areas of serious concern/hot spots and elucidate their underlying causes;
- to identify unforeseen impacts and new areas of concern;
- to create the background to develop prediction of expected effects and the verification thereof (hindcasting); and
- to direct future monitoring programmes.

Of course, each choice of sampling points and sample sizes for a survey leads to some data on marine contamination and possible effects (as long as anything at all is measured). However, if the survey is expected to generate statements like:

- an assessment of an absolute level ("level at position A is below/above a critical value") or

- a spatial comparison ("level at position A is lower than/comparable with/higher than at position B) or
- a temporal comparison ("level at position A at time T1 was lower than/comparable with/higher than the level at this position at time T2)
- the level of a parameter has changed in part X of the OSPAR maritime area.

with a defined precision, it is necessary to appropriately organize the survey with respect to sample sizes and sampling positions. The aim is to find a survey design, which is optimal in the sense that with a prespecified effort the most precise map of the spatial parameter distribution is obtained or that a prespecified precision is achieved with the smallest possible effort. To this end, various specifications are needed as input to the survey design, as given in Table 7.2.1 at the beginning of the next section. If the required specifications cannot be given, no *a priori* statement about the quality of the sampling can be made. In this case, a pragmatic way of designing the survey has to be followed as indicated in the last section. Then, however, an *a posteriori* determination of the power of the monitoring scheme should be performed to obtain a quantification of the monitoring quality. This should also be done if the optimal design were formally determined, but could not be followed in reality due to practical restrictions.

Survey design: optimal design for fixed stations

Table 1. Specifications needed as input to the derivation of an optimal survey design.

d, the change of biological interest	numerical specification of the change in parameter level that, if present, is to be detected with safety ß. Must be specified for each parameter.	no standard
ß, the power of test procedures	probability that an existing change at least as large as d is detected	90% or 95%
s_a, the analytical error of the biological/chemical analysis procedure	obtained from analytical experience, e. g. multiple measurements of the same sample	no standard
s_b, the biological variation	obtained from earlier investigation	no standard
D, the geographical area of interest		no standard
F, an initial guess of the spatial distribution of the parameter of interest	may be taken from pilot investigations or derived as educated guess.	if no other information, assume uniform spatial distribution.

An optimal survey design can only be developed in an iterative fashion. Prior to each campaign, an optimal design for that campaign is found by the procedure below. The results obtained from this campaign serve as input information for the optimization of the subsequent campaign.

Assuming that monitoring in a large area is intended, and that *a priori* information on the geographical distribution of the quantity under study is available, the following procedure can be used to derive an initial survey design (size and positions) for a monitoring according to the first part of the first bullet point.

> **Step 1**: Define D, the geographical area of interest (for which the assessment shall be valid) (See 7.1 above).

Step 2: Determine the necessary number of replicates per sampling location (needs knowledge of the sampling variability (analytical + biological, e.g. s_a, s_b), precision requirement plus standard statistics).

Step 3: Take the existing information F about the parameter of interest in this area and generate a map of the parameter level over the area of interest (use a standard geostatistical technique). Subdivide the range of the parameter in "iso-concentration" ranges. Find the corresponding "iso-concentration" areas on the map. If an iso-concentration area is ring-shaped, subdivide the ring into at least 4 sections (e.g. according to compass directions). Ring sections and the non-ring iso-concentration areas define the "sampling cells" addressed below.

Step 4: Define samplings points that are of basic interest or required for formal reasons. These points will not be changed by the following steps.

Step 5: Define an initial number of sampling points (a guess), additional to those from Step 4.

Step 6: Allocate sampling points from Step 5 to initial positions, starting with the geographical means of the sampling cells from Step 3. Define a grid of further candidate positions.

Step 7: For all present sampling points (initially those from Steps 4 and 5), calculate the estimated parameter value from the map of Step 3.

Step 8: Compare the map predictions from Step 3 and Step 7, e.g. by computing the Integrated means square error to characterize the present survey design. Record the IMSE.

Step 9: If there still are unvisited candidate grid locations, change the geographic locations of the free sampling positions to the next grid position (one change per step) and continue with Step 7. Otherwise finish.

The optimal survey design will then be the design that produced the smallest IMSE, e.g. the predictions that best reproduce the initial information. If this IMSE is considered too large, the number of sampling positions has to be increased and Steps 4–9, possibly 3–9, are repeated until a satisfactory result is achieved.

Survey design: a first approach for a fixed stations design

The procedure above may, for various reasons, not be acceptable when designing a monitoring scheme. As an alternative, a simple rule is proposed below.

- Determine the necessary sampling size per sampling position according to precision requirements as above.
- Use at least three sampling positions. Select these such that they include an unimpacted, a heavily impacted and an intermediate situation.
- If more than three sample positions are used, their positions should again cover the whole range of parameter values, preferably along a gradient.

The rationale behind this proposal is that it is necessary to obtain information about the best and the worst situations. The extremes are more likely to exhibit changes in future monitoring campaigns than sampling positions with a mean level.

No attempt should be made to generalize the findings from as few as three sampling positions to a large map. The quality achieved by the chosen design should be investigated by an *a posteriori* power analysis.

UK approach to survey redesign

The UK approach to redesigning its station network moves away from site-specific monitoring of hazardous substances to a more regional approach and uses random stratified sediment sampling to inform on status and trends supplemented by a minimum of one fish sampling site per region (contained within one stratum) to inform on status and to provide supporting information for biological effects monitoring.

Regions and strata have been defined covering the UK continental Shelf. Figure A1.1 shows an example of this, for the Region defined as Humber/Wash. The region consists several strata, which include Water Framework Directive water bodies in the 0–1 nm limit, an intermediate stratum 1–12 nm and two open sea strata, NE open sea and S open sea.

Collecting all samples at the same time and place may be considered to be the 'ideal' survey/sampling strategy for integrated monitoring, however this is usually not achievable in practice due to the seasonal limitations of some parameters, mobility of fish, unsuitable sediment types, etc. and such 'snap-shot sampling often fails to control local temporal and spatial variation in contaminant concentrations.

A regional approach generates more useful management information and can improve the power of the programme to detect trends by controlling local spatial variation.

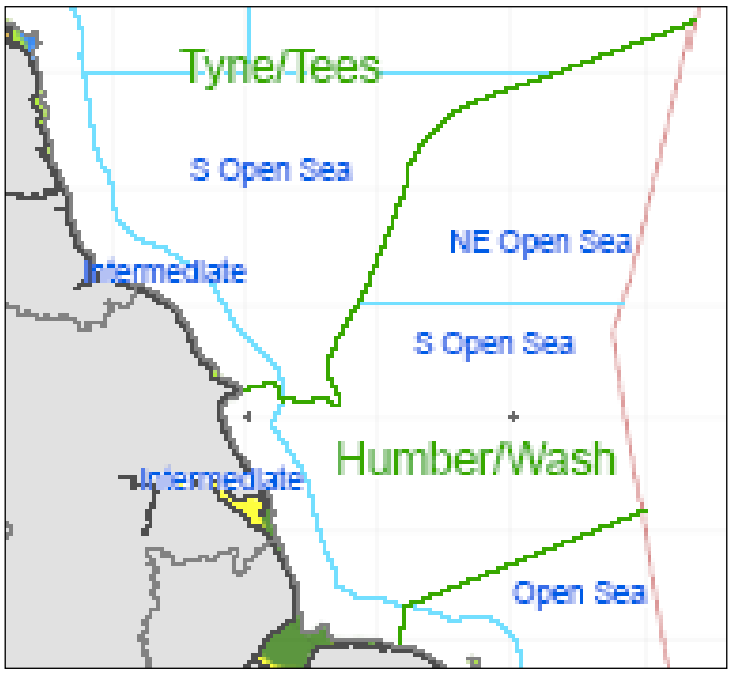


Figure A1.1. An example of the UK regional approach to redesigning the national monitoring station network.

Sample size for integrated monitoring

JAMP Guidelines specify sample sizes for each parameter, though without explicit justification in terms of error probabilities and detectable effects. The sizes given by JAMP seem mainly to be guided by practical considerations (get enough material for analysis and considering the time available to collect it). Additional to these considerations, a formal sample size calculation for single parameters could be done by us-

ing BACs, EACs, knowledge of analytical errors and standard specifications for acceptable error probabilities, in order to ensure that categorizations of parameters in e.g. a traffic light scheme are made with defined precision.

Sample size calculation for integrated assessment starts from the integrative assessment criterion, the value of which is to be determined with a specified precision. To calculate the necessary sample size needed for this precision, the mathematical form by which information on single parameters is accumulated into the integrative criterion is exploited, as the statistical distribution of the integrative criterion is determined by the random variation in the single parameters. As an example, the probability of a single parameter in a reference area exceeding its BAC is by definition 10%. In practice, real world observations from a reference area will not show an exact 10% rate of exceeding values (false positive rate) due to random biological variation and analytical imprecision. In an integrated assessment, each contributing single parameter also contributes a random error, which propagates to the integrative quantity according to the mathematical form by which the integrative quantity is calculated. The distribution of errors in the integrative quantity induced by the single parameters errors would be used for sample size calculation in the usual way. However, at present such a mathematical form is not yet available.

Annex 20: Proposed Terms of Reference for an ICES/OSPAR Workshop on the lysosomal stability data quality and interpretation to be held in association with a MEDPOL training course

Version date: 29 January 2010.

Background

The Hazardous Substances Strategy of the OSPAR Convention seeks to protect the maritime area from unexpected or unacceptable biological effects of contaminants. For some time, this objective has been approached through the development of standard suites of measurements, including both chemical analyses and biological effects assays.

The activity has been progressed through a series of meetings of the ICES/OSPAR Workshop on Integrated Monitoring of Contaminants and their Effects in Coastal and Open-Sea Areas (WKIMON). This process has developed schemes for the integrated monitoring of biological effects and concentrations of contaminants in fish and shellfish. This has been accompanied by the development of supporting documentation on methodology for a range of biological effects measurements.

WKIMON was discontinued in 2008. A new joint ICES/OSPAR Study Group on Integrated Monitoring of Contaminants and Biological Effects (SGIMC) was formed, and held its initial meeting in January 2009. SGIMC identified the need for the improvement of data quality and the further elaboration of data interpretation for some effects measurements to enable greater confidence in the data submitted for international and national programmes as contributions to monitoring programmes and environmental quality assessments.

In order to progress the development of integrated chemical and biological effects monitoring and assessment expeditiously it is necessary that the use of each method is clearly understood by participating laboratories, and that laboratories move towards harmonization of both analytical methods and data interpretation. Lysosomal membrane stability is a central parameter in integrated monitoring schemes developed for both fish and shellfish. It is a general, but sensitive, response to environmental stress which has widespread acceptance in marine and other environmental fields. In addition to its application in ICES/OSPAR programmes, it is likely that lysosomal membrane stability measurements will be required as part of the assessment of Good Environmental Status under the Marine Strategy Framework Directive (particularly Descriptor 8, that contaminant concentrations should not give rise to pollution effects).

SGIMC 2009 included a training Workshop on lysosomal stability in its forward work programme (SGIMC 2009 Report, page 12). SGIMC 2010 confirmed this recommendation and recommended that emphasis should be placed on data quality and interpretation. This workshop will contribute to the timely completion of the SGIMC task, and assist in answering OSPAR request 8/2008 on integrated monitoring and assessment.

It has come to the attention of SGIMC that MEDPOL is planning to hold a methodological training workshop in Alessandria, Italy from 13–17 September 2010. The training workshop would include practical training in cytochemical tests include the neutral red retention assay for lysosomal membrane stability, lipofuscin accumula-

tion, and in the use of micronuclei frequencies as a genotoxicity biomarker. SGIMC 2010 considered that there was an opportunity to link the MEDPOL and ICES/OSPAR workshops together and increase the overall value of the work.

It is proposed that an ICES/OSPAR Workshop on lysosomal stability data quality and interpretation be held on 18 September in association with the MEDPOL training event in Alessandria. This will enable participants to take advantage of the facilities and training being provided through the MEDPOL event. It will also encourage stronger links between ICES, OSPAR and MEDPOL countries, in keeping with the closer relationships that can be anticipated to be required for the implementation of the Marine Strategy Framework Directive. The purpose of the ICES/OSPAR Workshop is to improve the harmonization of the interpretation criteria (e.g. for samples with lysosomal abnormalities, etc.), to elaborate analytical method descriptions accordingly, and to develop procedures and proposals for international quality assurance activities for lysosomal stability.

There is no formal link between the workshop and the MSFD working groups although they deal with similar issues. The results of the workshop should provide science in support of MSFD implementation. The workshop is intended to bring science forward in response to the needs of OSPAR and MSFD groups.

The results are intended to bring the science forward and because there is no OSPAR request there is no review process or ACOM approval of the report involved. In establishing the workshop, ICES and OSPAR support improving ICES's advisory capabilities in this field. The report will be forwarded to ACOM and OSPAR.

Terms of Reference for an ICES/OSPAR Workshop on 18 September on the lysosomal stability data quality and interpretation, in association with a MEDPOL training workshop to be held in Università del Piemonte Orientale "Amedeo Avogadro", Alessandria, Italy on 13–17 September 2010, and convened by Concepción Martínez Gómez (Spain)

Objectives

The objectives of the joint MEDPOL and ICES/OSPAR workshops are:

MEDPOL

i) provide training in the measurement of lysosomal membrane stability in mussel samples.

ii) provide training in other biological effects measurement techniques for marine samples relevant to OSPAR and to MSFD, including of lipofuscin accumulation, micronuclei enumeration and Stress on Stress.

ICES/OSPAR

iii) develop guidance on the interpretation of the results by using neutral red retention assay for lysosomal membrane stability in marine samples, giving particular attention to the variation in sample characteristics experienced through the ICES/OSPAR area.

iv) develop written and graphical material to elaborate, harmonize and clarify criteria of interpretation and the standard analytical method for the neutral red retention assay of lysosomal membrane stability (ICES/OSPAR).

v) develop proposals for effective and repeatable external quality assurance programmes for the measurement of lysosomal membrane stability (ICES/OSPAR).

Activities

The ICES/OSPAR element of the workshop will be organized around the following activities, building on the activities of SGIMC (and WKIMON):

Presentation and discussion of the background information

a) existing experience of the use of the neutral red retention assay for lysosomal membrane stability in mussel samples will be presented, reviewed, and collated.

b) the information will be reviewed in the light of experience during the training elements of the MEDPOL Workshop.

Output from the Workshop

c) enhanced methodological guidance for the neutral red retention assay for lysosomal membrane stability in mussel samples.

d) proposals for effective and repeatable external quality assurance programmes for the measurement of lysosomal membrane stability by the neutral red retention assay.

Organisation

The workshop will be coordinated by Aldo Viarengo (Italy) for MEDPOL and by Concepción Martínez Gómez (Spain) for ICES/OSPAR.

Participants: the ICES/OSPAR component of the workshops should be attended by scientists with experience of the application of the neutral red retention assay for lysosomal membrane stability in marine samples and in the assessment of data, and in the development of QA schemes to ensure that the resulting report to ICES/OSPAR is of high quality.

Preparatory work

The following existing documents are relevant to the Workshop and will be collated and distributed prior to the Workshop

a) Background documents in the OSPAR Strategy for Hazardous Substances;

b) OSPAR Background Documents concerning lysosomal membrane stability, relevant biological effects;

c) ICES TIMES and MEDPOL documents on the measurement of lysosomal membrane stability by the neutral red retention assay in marine samples. Reports of relevant meetings of ICES/OSPAR WKIMON and SGIMC groups, and of ICES WGBEC.

d) Assessment criteria for biological effects currently adopted by OSPAR and MEDPOL

Location

The Workshop will be held at Università del Piemonte Orientale "Amedeo Avogadro", Alessandria, Italy.

Time schedule

Feb–June 2010	Workshop organizer	Develop effective working relations with the MEDPOL organizer and develop programme for the Workshop
July–August 2010	Workshop organizer and host organization	Collate and distribute background material
13–18 September 2010	Workshops	
November 2010		Report of the Workshop finalized and made available to ICES and OSPAR Secretariats
Early 2011	SGIMC 2011	Output from Workshop provide input to SGIMC.
Early 2011	ACOM	ACOM to review Workshop output as a contribution to the further development of guidance on integrated monitoring of chemicals and biological effects (OSPAR request 8/2008).

Supporting information

PRIORITY:	The purpose of this Workshop is to provide a dedicated opportunity to focus on improving the quality of data for a specific biological effects measurement and technique – lysosomal membrane stability by using neutral red retention assay. This will allow the measurement to play a more reliable role in OSPAR assessment processes, and also in the implementation of MSFD GES Descriptor 8 concerning contaminants and pollution effects.
SCIENTIFIC JUSTIFICATION AND RELATION TO ACTION PLAN:	The objectives of the joint MEDPOL/ICES/OSPAR workshops are to: provide training in the measurement of lysosomal membrane stability in mussel samples (MEDPOL). provide training in other biological effects measurement techniques for mussel samples relevant to OSPAR and to MSFD, including of lipofuscin accumulation and micronuclei enumeration (MEDPOL). develop guidance on the interpretation of the neutral red retention assay for lysosomal membrane stability in marine samples, giving particular attention to the variation in sample characteristics experienced through the ICES/OSPAR area. develop written and graphical material to elaborate, harmonize and clarify criteria of interpretation and clarify the standard analytical method for lysosomal membrane stability (ICES/OSPAR). develop proposals for effective and repeatable external quality assurance programmes for the measurement of lysosomal membrane stability (ICES/OSPAR).
RESOURCE REQUIREMENTS:	The research programmes which provide the main input to this Workshop are already underway, and resources are committed. The additional resource required to undertake additional activities in the framework of this event is negligible.
PARTICIPANTS:	The Workshop is expected to attract 8–15 members, in addition to participants from the MEDPOL area. In the light of current and future work related to the Marine Strategy Framework Directive and involvement of the Regional Seas Conventions it is the intention to invite participants from the HELCOM area.
SECRETARIAT FACILITIES:	SharePoint
FINANCIAL:	No financial implications.
LINKAGES TO ADVISORY COMMITTEES:	The Workshop will report to ACOM and OSPAR.
LINKAGES TO OTHER COMMITTEES OR GROUPS:	The report of the Workshop will be of interest to ICES Science Programme and various WGs including MCWG, WGBEC and WGMS.
LINKAGES TO OTHER ORGANIZATIONS:	The Workshop will be a joint ICES/OSPAR group reporting to both bodies, and will meet in association with a MEDPOL training event.

Annex 21: Technical Minutes from the Review Group of the Integrated Monitoring of Contaminants and Biological Effects 2

- (RGMON2)
- By correspondence: deadline 16 April 2010
- Participants: Ellen Kenchington, Canada (Chair), Pekka J. Vuorinen, Finland and Claus Hagebro and Michala Ovens (both from ICES Secretariat)
- Working Group: SGIMC

RGMON2 meeting and remit

RGMON2 met through correspondence during the week of April 12th 2010 to review the work done by the Joint ICES/OSPAR Study Group in Integrated Monitoring of Contaminants and Biological Effects (SGIMC) pertaining to the OSPAR Request OSPAR2008/8:

Further development of guidance on integrated monitoring and assessment of chemicals and biological effects

To complete the development of JAMP guidance for integrated monitoring of chemicals and their biological effects through preparing Technical Annexes on:

i) Survey design. The purpose is to provide guidance on the selection of representative stations, taking into account requirements under the Water Framework Directive and the proposed Marine Strategy Directive, and for the selection of stations for integrated monitoring. This work should build on work by WGSAEM 2007 relating to the spatial design of monitoring programmes and should take into account the approach taken by the UK in re-designing their station network;

ii) Groups of biological effects methods to be deployed to address specific questions. This should provide guidance on recommended packages of chemical and biological effects for monitoring on determinand basis to ensure that chemical and biological methods were well matched and that chemical analysis underpinned biological effects monitoring.

SGIMC began their work on this request in 2009 and will continue through to 2011. In 2010, SGIMC prepared twelve Technical Annexes that provide background material for a suite of biological effects, (i.e., EROD, lysosomal stability, PAH bile metabolites, DNA adducts, liver neoplasms, acetyl cholinesterase levels, DNA damage, reproductive success in eelpout, stress in bivalve molluscs) as well as water bioassays. Four Technical Annexes were produced related to integrated monitoring:

- Annex 16: Technical Annex on Sampling and Analysis for Integrated Chemical and Biological Effects Monitoring in Fish and Shellfish;
- Annex 17: Technical Annex for Mussel (*Mytilus* sp.) OSPAR Integrated Monitoring;
- Annex 18: Technical Annex on Recommended Packages of Chemical and Biological Methods for Monitoring on a Contaminant Basis;
- Annex 19: Discussion Document on Survey Design for Integrated Chemical and Biological Effects Monitoring.

Annex 19 relates directly to Request OSPAR2008/8(i), while Annex 18 relates directly to Request OSPAR2008/8(ii). However, all of the Technical Annexes are being put forward by SGIMC as required background material needed before Request OSPAR2008/8 can be fully addressed in 2011.

The role of RGMON2 was to provide an independent review of the technical text of the SGIMC Report related to Request OSPAR2008/8. Specifically RGMON2 reviewed the SGIMC Report to determine:

- Is it technically correct?
- Is the scope and depth of the science appropriate to the request?
- Does it answer the request (OSPAR2008/8)?

1 Review of Technical Annexes

1.1 General comments on assessment criteria

1.1.1 Clarifying the objectives

The Technical Annexes all provide guidance on Background Response and Assessment Criteria; however, it is unclear what the higher level objectives for setting these criteria are. Annex 10 of the 2009 SGIMC Report indicates that the policy drivers for this request are the Water Framework Directive (WFD) and the Marine Strategy Framework Directive (MSFD).

The general objective of the WFD is to achieve the 'good status' of surface waters by 2015 in all Member States. 'Good status' is defined as the status achieved by a surface water body when both its ecological status (i.e., an expression of the quality of the structure and functioning of aquatic ecosystems, classified in accordance with Annex V of the WFD) and its chemical status (concentrations of pollutants compared with the environmental quality standards established in Annex IX and under Article 16(7) of the WFD, and under other environmental quality standards at Community level) are at least 'good'.

Different prescriptions apply to artificial and heavily modified water bodies where the target is to achieve the 'good ecological potential' (as defined in Annex V of the Directive) and good surface water chemical status at the latest 15 years from the entry into force of the WFD.

The WFD classification scheme for water quality includes five status classes: high, good, moderate, poor and bad. 'High status' is defined as the biological, chemical and morphological conditions associated with no or very low human pressure. This is also called the 'reference condition' as it is the best status achievable - the benchmark. These reference conditions are type-specific, so they are different for different types of rivers, lakes or coastal waters so as to take into account the broad diversity of ecological regions in Europe. Assessment of quality is based on the extent of deviation from these reference conditions, following the definitions in the Directive. 'Good status' means 'slight' deviation, 'moderate status' means 'moderate' deviation, and so on.

The overall goal of the MSFD is the application of an ecosystem-based approach to the management of human activities. This ensures that the collective pressure of such activities is kept within levels compatible with the achievement of "good environmental status", and that the capacity of marine ecosystems to respond to human-induced changes is not compromised, while enabling the sustainable use of marine goods and services by present and future generations. Thus priority should be given to achieving or maintaining good environmental status in the Community's marine environment, to continuing its protection and preservation, and to preventing subsequent deterioration.

While ecological and pressure-related descriptors of good environmental status are provided, the MSFD does not provide indicators or reference levels at present. It states that when devising them, Member States shall take into account the application of relevant existing national, Community or international level environmental targets, ensuring they are mutually compatible and address any transboundary impacts.

The MSFD recognises that due to the transboundary nature of the marine environment, Member States should cooperate to ensure the coordinated development of marine strategies for each marine region or subregion and a subsequent standardized assessment. For this reason the ecological standards which relate to the sustainable use of the sea need to be assessed in a consistent and scientifically sound way.

The first step in setting a reference level for an indicator is to decide what ecological or chemical property the reference level has to represent. This is rooted in the policy objective that the indicator is supposed to serve. It is often the case that the policy objective is not explicit enough about the state intended to be achieved by the objective to make the choice of a reference quantitatively straightforward. In the cases of the WFD and MSFD those objectives are "achieve good status" and "achieve good environmental status" respectively, and both will require substantial interpretation to select a corresponding reference level on an indicator used to measure progress towards the objective. For water quality indicators as those proposed here, an indicator may be applied over wide spatial scales where ecological and physical conditions differ enough that a single absolute or relative value may not capture the desired water quality or ecological conditions in every place. Hence the science advisory process will have to apply sound and systematic reasoning to translate the water quality or ecological status intended by a policy directive into the properties that should be captured by the corresponding reference level on an indicator.

The SGIMC Report delivers a suite of indirect indicators. That is, they are measured in living organisms (mostly fish and shellfish) and are selected to measure *in vivo* physiological responses to *in situ* water chemistry, namely pollutants. They are indirect in the sense that they measure water quality through its response in the fish and shellfish. In this sense the indicators most directly relate to the WFD. However, because they are measured in living organisms they could also be adapted to relate to the MSFD. To do this it would be necessary to link the physiological response of the fish/shellfish more clearly to fish health. For example, the EROD (S9 Fraction) assessment level for Dab is 178 pmol/min/mg prot. The threshold is couched in terms of unexposed/exposed, meaning that fish above the limit are exposed. What does it mean to fish health to be above this limit? Does this have population consequences?

For all of the various indicators reported here there is a need to identify clearly the objective(s), and to equate the assessment criteria (reference points) to the objectives in the language of the policy directives for which they are aimed. This may involve extending the current binary criteria into more categories and/or to provide guidance on the interpretation of changes in the metrics. The excellent work done by SGIMC thus far should allow for this extension to be considered when finalizing the request for Advice in 2011.

1.1.2 Use of reference sites for determining baseline conditions for assessing change

The SGIMC use reference sites to determine the appropriate indicator level for the unexposed or non-responding state. For most species these reference sites were geographically restricted compared to the range of the species, a circumstance that was determined by available data. RGMON2 suggests that SGIMC clearly state whether this is an issue in establishing these thresholds in the context of the indicators presented. This is particularly important when figures showing the proportion of fish above and below the threshold are presented for broad geographic areas (e.g., Section 4.1.2, Figure 3). For the purposes of establishing assessment criteria it might be better to determine the most suitable locations for reference sites for each species *a priori*

and then to subsequently determine their EROD levels. This should include geographic variation consistent with the range area of the species.

It appears that the selection of sites for the reference baseline was determined as either "no known local sources of contamination" or "less influenced from human and industrial activity", but it is not clear how these were handled in the Report. Were data from the two classes combined? In order to align these classes to the WFD and MSFD it is necessary to clearly indicate the location and time frame used, in addition to the endogenous factors that went into the production of the threshold. Note that the WFD has different requirements for heavily modified water bodies where the target is to achieve the 'good ecological potential'.

1.1.3 Persistence of the effects

For each indicator (species/effect measure) it is important to know the persistence of the effect after initial exposure. This is not consistently addressed in the Report but if the indicators are to be used to measure water quality, then it will be important to know how responsive they are to change. Additionally are the effects reversible? Some clearly are reversible, while others perhaps are not (tumours?).

1.1.4 Use of percentiles to calculate threshold values

Most of the indicators appear to use the 90th percentile of the data as the threshold separating response/no response. RGMON2 does not recommend this approach when sample sizes are small and/or the median value is close to P90. Small sample sizes (e.g., N=30 for European flounder EROD) do not provide very robust estimates using percentile statistics. A few large records can widely change the P90 level. In cases were the median is close to the P90 (Four spotted megrim for EROD) a single data point can also widely change the threshold value and this is accentuated when the sample size is small. Although P90 is widely used as a limit measure it should only be applied when the underlying data are robust. Further, there is no biological basis for choosing one threshold over another. Why not the 95th or even the median? In other applications the point of maximum curvature of the cumulative distribution is used to mark the transition from one phase to another (e.g., Stirling and Zakynthinaki, 2008). For DNA adducts the SGIMC state that the P90 "is not associated with significant effects on fitness in whole organisms. There we have also defined a threshold value of significant effects". This approach is much more defensible and should be considered for all of the indicators.

1.1.5 Selection of species to measure effects on

The selection of fish/shellfish to use in biological monitoring is critical and a fundamental knowledge of the physiology, population dynamics and most particularly the migration behaviour of the species is required. Failure to understand the migration behaviour of the species can lead to falsification and misinterpretation of the physiological monitoring data. Some of the selected species, e.g., cod, can move large distances and the selection of species should be reviewed by fish biologists (WGFE?) to determine whether they are appropriately selected and whether the recommended sampling season, size and sex mitigate or compound this issue. When the indicator is meant to be applied at large spatial scales (e.g., the whole of an OSPAR Maritime area) then migration may not be an issue provided there is good geographic coverage, however, when the indicators are meant to represent local areas then the issue is much more pertinent.

1.1.5.1 Cryptic species

Any proposals involving using the mussel, *Mytilus* spp., should recognize that there are at least three species living sympatrically in the OSPAR area: *M. galloprovincialis*, *M. edulis*, and *M. trossulus* with widespread hybridization occurring among them (e.g., McDonald *et al*., 1991 but many others). They are known to have different ecological tolerances to salinity and so it can be expected that they exhibit other differential physical responses to the environment. Most protocols call for testing on *M. edulis* but the diagnostic markers for species determination are usually not considered or elaborated on, and involve nuclear and mitochondrial DNA markers, as external morphology is highly plastic. Many studies assume that the mussel is *M. edulis*. For example, in the SGIMC Report observations on background AChE activity in mussels were derived from mussels collected in the northern Baltic Sea as "*M. edulis"* when they were almost certainly the Baltic Blue Mussel *M. trossulus* or its hybrids (cf. Gardeström *et al*., 2008). It would be far simpler to demonstrate that these congeners have similar threshold levels and so can be used indiscriminately, however, this does not appear to have been done and so it is possible that the species and their hybrids will have differential responses. Some comments on this issue should be included.

1.1.6 Adaptation

Organisms have the capacity to adapt as individuals to their environment and over time to adapt as a population through natural selection. For example, in an acidified lake (due to acid rain) the plasma chloride concentration of whitefish was 90% of that in whitefish from a neutral lake, but the whitefish in the acidified lake seemed to manage; they had adapted (Rask *et al*., 1992). This was just a three-year observation, and thus represented individual adaptation but over time the population as a whole could also adapt. Genetic adaptation is very common and is discussed specifically in Section 1.13 where concerns are raised over its influence on SoS in bivalves, but all of the biomarkers represented have potential for both individual and genetic adaptation which is not discussed in the SGIMC Report.

1.2 Annex 4 SGIMC Report: Cytochrome P4501A activity (EROD)

SGIMC have prepared a very good background document on EROD, building on work done in their 2009 Report. EROD activity in fish is proposed as an indirect indicator of environmental contamination with the strongest inducers of hepatic EROD being dioxins, planar PCBs and PAHs. Saborowski and Bucholz (1997) document coastal migrations in dab and conclude that "the population of females in the Helgoland area was more constant than that of males. Accordingly, it is of advantage in ecophysiological research to concentrate on investigations on female specimens, which are available all through the year for field sampling". They further describe the results of tagging experiments in the southern North Sea which suggest migrations during the recommended sampling season. There was insufficient time to adequately review these issues for the other species but similar concerns were raised over cod, which are known to migrate to feeding and spawning grounds (e.g., Turner *et al*., 2002).

1.3 Annex 5 SGIMC Report: Lysosomal stability as a global health status indicator

This section is well written and well referenced although this indicator does not appear to have been addressed in the body of the Report and therefore there is no back-

ground material other than what is in Annex 5. The Annex explains how the indicator can be applied to a wide range of organisms using the same set of thresholds but does not suggest what species it should be used on in the OSPAR Maritime area.

1.4 Annex 6 SGIMC Report: DNA adducts

This section also is well written and well referenced but again this indicator does not appear to have been addressed in the body of the Report and therefore there is no background material other than what is in Annex 6. For this indicator the reference condition is "pristine" as fish have been collected from such areas as supported by chemical and biomarker analyses. It appears to be a good marker for PAH exposure and is technically sound. Specific reference levels have not been set but provisional levels for flounder, dab and Atlantic cod have been suggested. The presentation of thresholds which produce significant fitness effects on the whole organism is very valuable (see Section 1.1.4). As for other indicators the migration potential of the species used for monitoring should be evaluated (e.g., Atlantic cod).

1.5 Annex 7 SGIMC Report: PAH metabolites in bile

This indicator is responsive on the scale of days or weeks as the compounds are metabolised rapidly and excreted through digestion. It is therefore also well suited for monitoring recovery from spills or other point sources. Assessment criteria have been derived from reference sites in Iceland and the Barents Sea and follow the P90 criteria. Regional differences are larger than species differences and the SGIMC recommends combining data from all fish species when comparing data on a large scale. The points discussed in Section 1.1 should be considered, and some link to significant effects on whole organisms (as per Section1.4) should be made. A pertinent article about biliary PAH metabolites was not cited. Vuorinen *et al*. (2006) compare standardisation methods and sex differences and this should be included in the reference list.

1.6 Annex 8 SGIMC Report: fish diseases and liver neoplasms

This section is well written and addresses many of the concerns expressed by RGMON2 in Section 1.1. The Fish Disease Indicator and reference levels more accurately reflect biological responses than the P90 used elsewhere. We are not sure about the smiley faces though!

1.7 Annex 9 SGIMC Report: water *in vivo* bioassays

No further comments.

1.8 Annex 10 SGIMC Report: AChE activity

See comments in 1.1.5.1 regarding cryptic species for Section on seasonality in mussels in the main body of the Report (Section 5 SGIMC). The Annex itself is well written, however it seems that assessment criteria can only be set for one species (*Mytilus edulis*).

1.9 Annex 11 SGIMC Report: Micronucleus assay

This Annex is not complete. It is missing references and a number of track change comments which were presumably made by SG members persist and have not been addressed. Again this indicator does not appear to have been addressed in the body

of the Report and therefore there is no background material other than what is in Annex. The general sense is that this method may be premature to present as in indicator. It is unclear why SGIMC recommends that this Annex go forward at this time.

1.10 Annex 12 SGIMC Report: Comet assay

No specific technical comments. Again this indicator does not appear to have been addressed in the body of the Report and therefore there is no background material other than what is in Annex 12. This Annex is well written but the indicator is not ready to "roll out". Standardized protocols must be established and assessment criteria have not been developed. It is unclear why SGIMC recommends that this Annex go forward at this time.

1.11 Annex 13 SGIMC Report: Reproductive success in eelpout

Again this indicator does not appear to have been addressed in the body of the Report and therefore there is no background material other than what is in the Annex. The indicator shows potential as an indicator of eelpout health as well as water quality. The use of a percentage of dead larvae is valuable and could be considered as a percent that would affect population growth (5% is probably low for that) extending it use as an indicator of population health for an ecosystem context. The Annex proposes it as an indicator of the "impact of hazardous substances on reproductive success of fish in the marine environment," however no evidence is presented to show that it reflects reproductive success in species other than eelpout. Although reproductive success or embryonic development is convenient to study in eelpout, vivipary is very rare and exceptional among fish and so RGMON2 considers that it is not a good indicator of fish health in general.

1.12 Annex 14 SGIMC Report: fish indices

This Annex was not ready for review. The SGIMC notes that it needs further work and so it was not assessed further by RGMON2.

1.13 Annex 15 SGIMC Report: stress in bivalve mollusks

The evidence presented supporting the use of this indicator is less than compelling. It does not address the many potential confounding effects on the ability of bivalves to withstand exposure to air, not least of which is position on the shoreline relative to EHWS tides. This indicator has parallels to the ability of clams to burrow when exposed to toxic microalgae (Bricelj *et al.*, 2005). Those studies show that natural selection takes place in areas that are repeatedly exposed to red tides, resulting in genetic adaptation. This literature should be considered in this Annex as it is likely that mussels will also show adaptation to chronic levels of contaminants. It is the opinion of the RGMON2 that this Annex not go forward to OSPAR as it requires more critical evaluation.

2 Review of Annexes pertaining to sampling and analyses, and integrated monitoring

The integrated approach has been based around recommendations of sets of measurements that could be used to investigate the effects of contaminants on fish and/or shellfish (Annex 16). The suite of Technical Annexes described in Section 1 provides indicators that affect individuals at the tissue chemistry, subcellular, tissue response and whole organism levels. Some of these indicators may have relevance to human health if the organisms are consumed and if that is the case then that should be clarified. However for them to be useful as indicators in the context of the MSFD or WFD the responses must be shown to affect fitness of the whole organism or populations through disruption to reproductive processes (e.g., imposex). This is noted in Annex 16 but is not developed.

Further, collection of samples through existing resource assessment surveys will provide direct links to additional data on population and/or habitat. For example, many assessment surveys collect information on sex ratio, age/size structure etc. which could be used with the chemical and biological effects monitoring. They usually follow stratified –random site selection procedures or other statistically sound survey designs which could be used to give the same advantages to the effects monitoring. Given the widespread prevalence of such surveys and the fact that many of the species proposed for monitoring are commercial species or bycatch, it might be useful to look at what aspects of biological effects monitoring could not be served by collecting samples through resource assessment surveys. One issue could be spatial coverage or time of year but these would need to be systematically reviewed. In general the steps outlined in Annex 19 are sound and should be applied to each situation prior to looking at the economies of using existing surveys to fulfill these needs.

Both Annexes 16 and 17 show that for both fish and mussels there are several choices of indicators for each of the four response levels (tissue, subcellular, etc.). Annex 18 reverses the information and provides tables of indicators which measure chemistry and biological effects in the water, fish and shellfish for groups of contaminants (e.g., metals, PCBs, etc.). The SGIMC should comment on whether all of these indicators are needed to address the objectives and if not which indicators are recommended and why?

3 References

Bricelj, M., Connell, L., Konoki, K. *et al*. 2005. Nature 434, 763–767.

Gardeström, J., Pereyra, R.T. and Andre, C. 2008. Conservation Genetics 9, 1003–1005.

McDonald, J., Seed, R. and Koehn, R. 1991. Marine Biology 111, 323–333.

Rask, M., Vuorinen, P.J., Raitaniemi, J. *et al*. 1992. Hydrobiologia 243/244: 277–282.

Saborowski, R. and Bucholz, F. 1997. Helgolander Meeresunters. 51, 41–51.

Stirling, J.R. and Zakynthinaki, M. 2008. The point of maximum curvature as a marker for physiological time-series. Journal of Nonlinear Mathematical Physics 15, 396–406.

Turner, K., Righton, D. and Metcalfe, J.D. 2002. Hydrobiologia. 483, 201–208.

Vuorinen, P.J., Keinänen, M, Vuontisjärvi, H. *et al*. 2006. Mar Poll Bull 53: 479–487.