Report of the ICES/OSPAR Workshop on Assessment Criteria for Biological Effects Measurements (WKIMC)

14–16 October 2009
Aberdeen, Scotland, UK
# Contents

1 Opening of the meeting ................................................................. 2

2 Adoption of the agenda ................................................................. 3

3 Discussion of ToR and background to the meeting ......................... 4

4 Completion of the development of supporting documents for a coherent set of biological effects measurements linked to planar organic contaminants, particularly PAHs ........................................... 8
   4.1 Lysosomal stability-WKIMC review of background document and assessment criteria .................................................. 8
   4.2 The in-vitro DR CALUX bioassays for dioxin-like activity .......... 8
   4.3 Background document on DNA adducts ................................... 9
   4.4 Background document on PAH metabolites in fish bile .......... 9
   4.5 Summary of documentation ................................................... 9

5 Clarification of the mutually supporting role of a range of biological effects measurements in relation to OSPAR objectives and GES under MSFD ................................................................. 14

6 Development of assessment criteria for EROD, bile metabolites, DNA adducts of PAHs, DR-CALUX and aspects of fish disease related to contaminants .................................................................. 23
   6.1 Assessment criteria for fish diseases ....................................... 23
   6.2 EROD Assessment criteria .................................................... 24
   6.3 Determination of threshold level of significant effects for DNA adducts in cod .......................................................... 26

7 Presentation of proposals for the adoption and utilisation of assessment criteria for the biological effects measurements indicated above ................................................................. 32

8 Other related business ................................................................ 33
   8.1 SGIMC work’s contribution to a risk based approach of produced water offshore ......................................................... 33
   8.2 Array technologies – Genomic analysis .................................... 33
   8.3 Work programme for SGIMC ................................................ 36

9 Completion of the Workshop report ............................................. 40

10 Closing of the meeting ................................................................. 41

Annex 1: List of participants .......................................................... 42

Annex 2: Agenda ........................................................................... 44

Annex 3: Terms of Reference ........................................................ 45
Annex 4: Assessment criteria for fish diseases .................................................................48
Annex 5: Revised Chapter 6 for OSPAR Background Documents on Biological Effects Monitoring Techniques ........................................................................................................58
Annex 6: Background document (revised) on DNA adducts .........................................185
Annex 7: Revised Background document of PAH metabolites in fish bile ..................195
Annex 8: Revised Background document on Bioassays (Chapter 7) .............................204
Annex 9: Revised Background document on Cytochrome P450IA activity (EROD) (Chapter 3) .........................................................................................................................225
Annex 10: Recommendations ..........................................................................................233
Executive summary

SKIMC met in the Marine Scotland Marine Laboratory, Aberdeen, UK from 14–16 October 2009. The main tasks completed were:

a) Review and updating of OSPAR Background Documents on a range of biological effects measurements.

b) Review and confirmation of assessment criteria for biological effects measurements, and development of new assessment criteria for a range of effects.

c) Elaboration of an integrated scheme for the assessment of biological effects and environmental chemistry data for use in environmental quality assessment.

d) Updating of the forward work programme for SGIMC, and for cooperation with WGBEC.

It is recommended that the output in the form of revised chapters for the OSPAR Background Document on Biological Effects Monitoring Techniques is sent to SGIMC 2010 and WGBEC 2010 for final review prior to communication to OSPAR with recommendations for adoption into the OSPAR Background Document on Biological Effects Monitoring Techniques. The revisions to the Chapters cover a range of issues, but particularly include proposals for assessment criteria.

WKIMC developed initial proposals for an integrated assessment scheme for biological effects measurements that is compatible with the assessment scheme applied by OSPAR to chemical measurements and recommends that it be discussed by WGBEC and SGIMC 2010.
1 Opening of the meeting

The meeting was opened at 0930 hr on Wednesday 14 October 2009 by the co-chairs, Ian Davies (OSPAR) and Dick Vethaak (ICES). The Workshop was formally welcomed to Marine Scotland Marine Laboratory by Professor Colin Moffat, Director of the Aquatic Environment Programme. He emphasized the importance of the Workshop in cementing the role of biological effects into ICES/OSPAR and MSFD programmes, and the need for assessment criteria to allow this to occur. He noted the value of effects measurements in addressing the combined effects of different contaminants, the need to accelerate the introduction of effects-based assessments and the opportunity to take account of emerging environmental technologies, including, for example genomics.
2 Adoption of the agenda

A draft agenda, based on preliminary discussion of the terms of reference, was accepted.
3 Discussion of ToR and background to the meeting

SKIMC recognized that this Workshop was key to the continuing development of the use of biological effects measurements in national, Regional and international monitoring and assessment programmes. It was particularly timely, in that the MSFD GES Descriptor 8 Task Group would meet in Italy the following week.

The Terms of Reference were targeted at improving the availability of assessment criteria for effects measurements, and also to clarifying how they might be used in an integrated assessment context. Ian Davies gave an introductory presentation of the current situation, as summarized below.

Over the last few years, ICES/OSPAR WKIMON and associated groups have progressively developed an integrated approach to the use of biological effects 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 man made 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?

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 (Figure 1) and shellfish (Figure 2) 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 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. (Blue: 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 full 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. 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 reveal 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 a 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 to 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.

SKIMC discussed the potential application of biological effects measurements in the assessment of Good Environmental Status (GES) under the Marine Strategy Framework Directive. The qualitative Descriptors of GES currently being considered by a series of Task Groups set up by the Commission through ICES and the JRC, and their advice on criteria and methodological standards is expected to emerge over the next few months. Of particular relevance to the work of SKIMC is Descriptor 8 of GES – that “contaminant concentrations should not give rise to pollution effects”. SKIMC noted that this objective has considerable parallels with the OSPAR Strategy for Hazardous Substances, and therefore an approach to integrated assessment in an OSPAR context should also be very relevant to GES Descriptor 8.

Early considerations of possible interpretations of GES Descriptor 8 had suggested that there appeared to be two main routes by which the descriptor could be approached. One would rely heavily on the current system for the assessment of environmental chemical data. A variety of assessment criteria or standards are currently used. Examples include the EQSs established under WFD (for example) for concentrations of contaminants in water, Background Concentrations (and BACs) and Environmental Assessment Criteria (EACs) defined through OSPAR for concentrations in sediment or biota, and ERL and ERM concentrations used in the USA for assessment of contaminant concentrations in sediment. These criteria address two main thresholds – whether the concentrations exceed normal/background levels, and whether the concentrations present a significant risk to organisms. In general, Background Concentrations are derived from field observations of samples from areas distant from sources of contamination (or deep sediment cores). EACs and similar values are derived of toxicological data, normally acute toxicity data, often for the effects of single substances in laboratory tests, while ERL/ERM values are derived from statistical analysis of field data linking concentrations of contaminants in sediments which were found to be associated with biological effects.
4 Completion of the development of supporting documents for a coherent set of biological effects measurements linked to planar organic contaminants, particularly PAHs

WKIMC 2009 reviewed the OSPAR Background Documents for:

a) Lysosomal stability  
b) Cytochrome P4501A (EROD)  
c) DNA adducts of PAHs  
d) Bile metabolites of PAHs  
e) Fish disease related to contaminants  
f) (Water) bioassays

In doing so, WKIMC took account of comments made at previous meetings (e.g. WKIMON IV, SGIMC 2009, WGBEC) regarding the need to update and amend these documents. The revised texts are included as Annexes to this report. Specific comments regarding some of the documents are in the subsections of Section 4 below.

4.1 Lysosomal stability-WKIMC review of background document and assessment criteria

The background document was considered concise, comprehensive and well referenced. The links between lysosomal stability, contaminants, other biological effect endpoints and conditions of severe physiological stress are highlighted in the document. Assessment criteria for ‘healthy’, ‘stressed but compensating’ and ‘severely stressed and probably exhibiting pathology’ are provided for both neutral red retention (NRR) and cytochemical methods of measurement. These are applicable across a wide range of species. The workshop agreed that the structure of these assessment criteria were suitable for identifying responses ‘above background’ and ‘significant effect’ levels.

WGBEC 2008 noted that the NRR lower assessment criteria (healthy: stressed but compensating) was often unachievable by some practitioners in some species/regions. This may be due to the background stress levels experienced by organisms at these sites or an intercalibration issue. In the absence of a recent intercalibration exercise this particular assessment criteria should be reviewed following a neutral red retention intercalibration exercise. Such an exercise is planned in the near future. WKIMC would support the use of these draft assessment criteria until such time as this has been reviewed.

4.2 The in-vitro DR CALUX bioassays for dioxin-like activity

The most relevant bioassays and biomarkers that indicate exposure to dioxin and dioxin-like compounds are considered to be induction of CYP1A/EROD activity in fish liver and application of the dioxin receptor based in vitro test, DR-Luc or also named DR-CALUX. The DR CALUX is a suitable in vitro screening assay that allows to measure the toxic potency, expressed as toxic equivalent quotient (TEQs) relative to the biological response in the DR Luc bioassay of the most toxic compound 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD). The TEQ values are calculated on basis of concentrations of individual congeners as determined by high resolution mass spectrometry (HRGCMS). An advantage of the application of these in-vitro bioassays (using extracts) as compared with CYP1A/EROD is that
they are independent of species differences and environmental influences, thus applicable in a generic way. Findings from several studies demonstrate this bioassay to be of value in both inshore and offshore regions, for example a high DR CALUX response was found in surface sediments at the Oyster Grounds, (an offshore region in the SW North Sea) that could be linked with the occurrence of larger PAHs (4–6 rings) (Klamer et al., 2005).

Background information on the in-vitro DR CALUX bioassay is included in the OSPAR BG document for biological effects methods: (Water) Bioassays including protocols for assessment tools and criteria (Chapter 7). Provisional assessment criteria, based on data from the UK and the Netherlands, were assessed and are now included in the background document. ICES TIMES-Series draft manuscripts on the methodology of DR CALUX (Schipper et al., in prep.) and on the extraction protocols for in vitro bioassays (Klamer et al., in prep) were made available to the workshop and are expected to be submitted before March 2010 for review and publication by ICES. ICES is requested to explore their capabilities to include DR CALUX data (contact D. Vethaak) in the ICES databank and to report on this to WGBEC in January 2010.


4.3 Background document on DNA adducts

The background document on DNA adducts was revised and the new version is included at Annex 6.

4.4 Background document on PAH metabolites in fish bile

The background document on PAH metabolites in fish bile was revised and the new version is included at Annex 7.

4.5 Summary of documentation

The current position regarding the preparation and/or revision of Background Documents and assessment criteria for biological effects measurements is summarized in Table B below:
Table B: Summary of current position with OSPAR Background Documents and assessment criteria.

<table>
<thead>
<tr>
<th>Biological Effect</th>
<th>Qualifying Comments</th>
<th>Background Document</th>
<th>Method of Determination / Expression of Assessment Criteria</th>
<th>Above Background Assessment Criteria</th>
<th>Significant Effects Assessment Criteria</th>
<th>Arrangements for Development</th>
</tr>
</thead>
<tbody>
<tr>
<td>VTG in plasma; μg/l</td>
<td>Cod</td>
<td>Thorough and complete</td>
<td>90 percentile from reference area / needs to be amended to zero</td>
<td>Needs to be developed but limited data available</td>
<td>Nothing on EAC equivalent*</td>
<td>Collation of existing data ID - FRS UK</td>
</tr>
<tr>
<td></td>
<td>Flounder</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reproduction in eelpout; mean frequency (%)</td>
<td>Malformed larvae Late dead larvae Growth / retarded larvae</td>
<td>Thorough and complete</td>
<td>95 percentile of ref sites / control</td>
<td>Already in place</td>
<td>Already in place</td>
<td>Further refinement as more data becomes available</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EROD; pmol/min/ mg protein S9/microsomal</td>
<td>Cod</td>
<td>Complete</td>
<td>Background responses as median of Median values from reference stations / 90 percentile as upper limit of background response</td>
<td>Proposed BR values in place</td>
<td>Not possible for EROD</td>
<td>Collect available EROD data and re-assess/ further development of a model with existing dataset</td>
</tr>
<tr>
<td></td>
<td>Dab</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Flounder</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Four-spotted megrim</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dragonet</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Red mullet</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plaice</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bile metabolites; 1-OH pyrene (μg/ml; 341/383 nm fluorescence, Syncr.scan); HPLC</td>
<td>Dab</td>
<td>Now complete for HPLC and FF. (Syncr. Scan for Flounder)</td>
<td>Based on monitoring data from reference areas.</td>
<td>90 percentile values from reference areas added for for cod, haddock, Dab and Flounder</td>
<td>Will not be made for Bile metabolites</td>
<td>Later other methods in use should be considered (Syncr.scan, GC/MS). Long rough dab data will be added later (when open)</td>
</tr>
<tr>
<td></td>
<td>Flounder</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Long rough dab</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biological Effect</td>
<td>Qualifying Comments</td>
<td>Background Document</td>
<td>Method of determination / expression of assessment criteria</td>
<td>Above background assessment criteria</td>
<td>Significant effects assessment criteria</td>
<td>Arrangements for development</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------------------</td>
<td>---------------------</td>
<td>------------------------------------------------------------</td>
<td>--------------------------------------</td>
<td>----------------------------------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>DNA adducts; nm adducts/mol DNA</td>
<td>Dab</td>
<td>Now complete for cod – sign. effect. ass. crit. to be verified for flatfish</td>
<td>Based on monitoring data from reference areas. Dab-data (UK) divided by 10 to be expressed in nm adducts/mol DNA.</td>
<td>90 percentile for Dab, Flounder, Cod, Haddock</td>
<td>Determined by comparison with chronic &amp; larval NOEC levels and added for cod - also indicative for flatfish (to be verified).</td>
<td>BL – UK to update/amend Sign. effects ass.crit. to be reviewed against new data for corresp. chronic &amp; larval NOEC levels. Sign. effects ass.crit. to be done for other monitoring species (dab/flounder other flatfish, haddock)</td>
</tr>
<tr>
<td>Bioassays; % mortality</td>
<td>Sediment Corophium</td>
<td>Thorough and complete</td>
<td>95 percentile of ref sites/control</td>
<td>Already in place</td>
<td>Already in place</td>
<td>Further validation as more data becomes available</td>
</tr>
<tr>
<td></td>
<td>Sediment Arenicola</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Water bivalve embryo</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Water copepod</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Water echinoderm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lysosomal stability; minutes</td>
<td>Cytochemical; all species</td>
<td>Thorough and complete</td>
<td>Best professional judgment from extensive literature and laboratory studies</td>
<td>Already in place</td>
<td>Already in place</td>
<td>Further validation as more data becomes available. Currently based on limited dataset from one country.</td>
</tr>
<tr>
<td></td>
<td>Neutral Red Retention; all species</td>
<td>Thorough and complete</td>
<td>Best professional judgment from extensive literature and laboratory studies</td>
<td>Already in place</td>
<td>Already in place</td>
<td>No further development</td>
</tr>
<tr>
<td><strong>BIological Effect</strong></td>
<td><strong>QUALifying Comments</strong></td>
<td><strong>Background document</strong></td>
<td><strong>Method of determination / expression of assessment criteria</strong></td>
<td><strong>Above background assessment criteria?</strong></td>
<td><strong>SignIFICANT effects assessment criteria?</strong></td>
<td><strong>Arrangements for development</strong></td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------------------------</td>
<td>-------------------------</td>
<td>---------------------------------------------------------------</td>
<td>------------------------------------------</td>
<td>----------------------------------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>Fish Disease Index</td>
<td>Combination of externally visible diseases, macroscopic liver neoplasms and liver histopathology</td>
<td>Details in WGPMO reports 2007, 2008, 2009; A summary is included in the OSPAR Background Document.</td>
<td>Cutpoints of the FDI assessment statistic: 2.5% and 97.5% quantiles of the assessment statistic. &lt; P 2.5%: desirable, P 2.5%-P 97.5%: indifferent, &gt; P 97.5%: raising concern.</td>
<td>Already in place</td>
<td>Already in place</td>
<td>No further development required. May be refined by ICES WGDPMO as appropriate</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>For contaminant-specific liver histopathology and for macroscopic liver neoplasms, WKIMC 2009 has proposed assessment criteria for unacceptable effects that can be used either in addition to the criteria for the FDI approach above or as independent assessment criteria (e.g. in case of exploratory monitoring or in newly launched monitoring programmes).</td>
<td>Proposed at WKIMC 2009</td>
<td>Proposed at WKIMC 2009</td>
<td>Further validation as more data becomes available</td>
</tr>
<tr>
<td>Biological Effect</td>
<td>Qualifying Comments</td>
<td>Background Document</td>
<td>Method of Determination / Expression of Assessment Criteria</td>
<td>Above Background Assessment Criteria</td>
<td>Significant Effects Assessment Criteria</td>
<td>Arrangements for Development</td>
</tr>
<tr>
<td>-------------------</td>
<td>---------------------</td>
<td>---------------------</td>
<td>-------------------------------------------------------------</td>
<td>-------------------------------------</td>
<td>------------------------------------------</td>
<td>-------------------------------</td>
</tr>
<tr>
<td>Scope For Growth</td>
<td>Mussels</td>
<td>Thorough and complete</td>
<td>Best professional judgment from extensive literature and laboratory studies</td>
<td>Already in place</td>
<td>Already in place</td>
<td>Further validation as more data becomes available</td>
</tr>
<tr>
<td>DR-CALUX</td>
<td>Sediment</td>
<td>Included in Chapter 7, but needs to be amended when TIMES Series becomes available</td>
<td>Data sets from 2 studies, using silica extraction</td>
<td>Limited data requires further validation / development</td>
<td>Further development required</td>
<td>DV / JT to update and amend</td>
</tr>
</tbody>
</table>


5 Clarification of the mutually supporting role of a range of biological effects measurements in relation to OSPAR objectives and GES under MSFD

Background

The objectives of OSPAR and GES under MSFD in relation to biological effects of contaminants differ. 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 man made 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?

In the MSFD context of the objective of achieving seas that are clean, safe, healthy, biologically diverse and productive, chemical contaminant concentrations is an aspect of clean seas. A potential consequence of seas that are chemically contaminated is that organisms or processes in the sea demonstrate responses, and in some case may be adversely affected. This is captured in Descriptor 8 of GES under MSFD is that “contaminant concentrations should not give rise to pollution effects”.

The objectives of OSPAR and GES are therefore different, and the contributions that biological effects can make to addressing them are also likely to be different.

It has been recognized for some time that biological effects measurements can be linked with chemical contamination in various ways. For example, biomarker responses at subcellular level can be used to detect exposure to particular types of contaminants and can be used as “early warning” or more serious effects, or can stimulate investigations to determine whether more serious effects are occurring. Conversely, if effects are observed at whole organism or tissue levels, subcellular measurements can be used to assist in the identification of the causative chemical agents, and can indicate where novel or unknown contaminants are having significant effects. These approaches have been widely used in field programmes, building on knowledge gained from laboratory experiments. However, until recently the processes of data interpretation and assessment have tended to be ad hoc and case/situation specific and generalization of the approach has been difficult. The formalization of a data assessment scheme has been an urgent need if effects data are to be integrated in national or international monitoring and environmental assessment programmes.

SKIMC noted that the mutually supporting role of a range of biological effects measurements arises from the adoption of integrated monitoring and data collection schemes, supported by integrated data assessment systems. The JAMP Guidelines for the Integrated Monitoring and Assessment of Contaminants and their Effects embody the current advice on integrated monitoring. A defined system for integrated assessment is currently not available, and therefore a major task of the SKIMC meeting is to develop a data assessment scheme that first integrates across biological effects data, and second has the capability of integrating these data with supporting chemical analytical data.
OSPAR Integrated monitoring of contaminants and their effects

The JAMP Guidelines for the Integrated Monitoring and Assessment of Contaminants and their Effects note that 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. Biological effects measurements can also indicate areas where unintended/unacceptable effects are occurring, which may be caused by chemical contaminants that are not included in the core CEMP chemical monitoring programme. Combination of biological effects with chemical measurements (or additional effects measurements) this will provide an improved assessment due to the ability to address effects that are potentially caused by a wide range of contaminants as well as those that are more clearly linked to specific compounds or groups of compounds.

The JAMP Guidelines contain advice on the appropriate combinations of chemical and biological effects measurements in integrated monitoring programmes of fish and shellfish (mussels). The process of integrated monitoring and assessment of these data leading to conclusions of status and associated feedback loop was conceptualised in the diagram below:
Figure 3: Integrated assessment process.

This scheme envisaged assessment criteria being available to categorize data into a “traffic light” system. To form the basis of assessment schemes, OSPAR has used two forms of assessment criteria in the interpretation of chemical monitoring data: background (assessment) concentrations (BC/BACs) and environmental assessment criteria (EACs). The former correspond to the green/orange boundary and the latter to the orange/red boundary. These have been used by OSPAR MON and other groups to develop assessments of contaminant concentrations in sediment and biota for the QSR 2010 project. Contaminant concentrations in sediment and biota at individual stations have then been integrated across stations and contaminants in various ways to prepare assessments of individual contaminants or groups of contaminants across subregions and regions.

Ultimately, the purpose of an integrated monitoring programme is to provide the necessary data to facilitate integrated assessments so that the status of the marine environment in relation to hazardous substances can be described as a contribution to general assessments of the quality status of the OSPAR maritime area (e.g. OSPAR QSRs). Integrated programmes can also potentially contribute to other objectives, such as the assessment of GES under MSFD.

The application of a similar approach to the assessment of biological effects data also requires assessment criteria to be defined, and this has been clearly recognized in the
preparation of OSPAR Background Documents on effects measurements. The concepts of a background level of response (enzyme activity, metabolite concentration, etc.), representing the response found in organisms which have been exposed to low concentrations of the causative contaminants has been found to be applicable to biological effects measurements. A higher level response that represents harm to the organism has also been found to be potentially applicable to biological effects measurements. However, both forms of assessment criterion are not appropriate to all of the effects measurements recommended in the OSPAR integrated monitoring strategy.

Work has been undertaken through ICES/OSPAR to develop assessment criteria to aid interpretation of some biological effects measurements. Mainly, these have concentrated on developing understanding of normal or background levels of the responses, such as are encountered in areas that are distant from sources of environmental contaminants. They can therefore be considered, to a degree, to be parallel with the Background Concentrations developed for chemical contaminant concentrations, and to address the first objective of the OSPAR Hazardous Substances Strategy.

Further assessment criteria (EACs) have been developed for chemical concentrations. These represent concentrations below which unacceptable biological effects are unlikely to occur, and therefore address the second objective of the OSPAR Hazardous Substances Strategy. Unacceptable biological effects are likely to be observed at whole organism or tissue levels. SKIMC considered that assessment criteria for whole organism and tissue level responses are more likely than subcellular responses to be suitable for the establishment of assessment criteria corresponding to background conditions and also to significant pollution effects.

ICES WGBEC has noted that some biological measurements are indicators of exposure to contaminants, while others are more clearly indicators of effect. In most cases, a background response level has been found to be applicable to both categories of measurements, but the higher level assessment criterion may only be applicable to indicators of effect. Examples of the latter include the imposex response of marine snails to tributyltin, and lysosomal stability. Background degrees of imposex (VDSI) are very low (<<1), whereas clear indications of inhibition of reproduction arte present if VDSI is >4. The OSPAR EcoQO structure for TBT effects in snails recognizes the significance of differing degrees of induction of imposex for snail populations. Lysosomal stability measurements (by two methods) can also be interpreted using two assessment criteria, as illustrated below:

<table>
<thead>
<tr>
<th>NRR</th>
<th>Cyt Ch</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \geq 120 \text{ mins} )</td>
<td>( \geq 120 \text{ mins} )</td>
</tr>
<tr>
<td>( \geq 50 \text{ mins} )</td>
<td>( \geq 20 \text{ mins} )</td>
</tr>
<tr>
<td>( &lt; 50 \text{ mins} )</td>
<td>( &lt; 10 \text{ mins} )</td>
</tr>
</tbody>
</table>

**Figure 4: Assessment criteria for lysosomal membrane stability.**

By contrast, some biological measurements are more appropriately viewed as biomarkers of exposure. Reasons for this differentiation can include that the response is rather transitory and unlikely to persist, or that there is no clear implication of the response for tissue or while organism level responses, or that the response is adaptive. Examples include the induction of metallothionein as a detoxification system for some metals, or the presence of metabolites of PAHs in fish bile. In these cases, while a background level of response may be recognizable, it may be difficult
or impossible to define a higher level of response corresponding to unacceptable harm at higher levels of organization.

This can be illustrated as below:

![Diagram showing relationships between assessment criteria and types of biological effect measurements.](image)

Figure 5: Relationships between assessment criteria and types of biological effect measurements.

Indicators of exposure therefore typically can be assessed against a background activity or response criterion, i.e., in relation to the green/orange boundary. Higher order effects can typically be assessed against both a background criterion and also against a criterion that represents unacceptable levels of biological effect, i.e., against both green/orange and orange/green boundaries. There may be a set of lower order measures of biological effect that are assessable against either both types of assessment criteria, or just a background level of response.

The assessment of data for individual biological effects, for example as part of quality status assessments, therefore requires that appropriate assessment criteria are developed. This has been progressed through various ICES and OSPAR groups including WKIMON, WGBEC, SGIMC and the current Workshop. The range of assessment criteria that have been developed to date is given in Table A, and some have been used in national and international programmes, such as the OSPAR QSR 2010.

It is generally recommended that biological effects measurements are not used individually, but that a suite of measurements should be used to provide a broad view of the influence of environmental quality on local organisms, or to obtain evidence of the effects of a particular class of contaminants on subcellular to whole organism levels of organization. A further degree of integration of assessment is necessary to derive such overall expressions of the status of fish or shellfish in relation to biological effects. Further integration is also necessary to allow combined assessment of biological and chemical monitoring data.

As outlined above, work in preparation for QSR 2010 allowed the integration and expression of chemical monitoring data in three colour traffic light systems. A similar approach to biological effects measurements would offer the potential to integrate across effects measurements and also to combine effects data with chemical data.
It is proposed that the biological measurements be considered in two groups; those to which only a lower (background) assessment criterion can be applied, and those to which an upper criterion can also be applied. The former will include biomarkers of exposure, while the latter will include biomarkers of lower or higher level effects. It is envisaged that integrated monitoring programmes will produce data for both groups of measurements. As indicated above, interpretation of effects data generally has been either based on single determinands, or else has followed a case-specific form of combination of data for different determinands. An important task for the utilization of integrated monitoring in international programmes is to develop a general procedure data integration.

A proposal for such a procedure is outlined in Table 1. This table illustrates data for a range of biomarkers of exposure and effect which have been interpreted against either one or two assessment criteria, and the conclusions are presented as traffic-lights. It is proposed that the first stage in integration is to integrate data for all biomarkers of exposure. These data are summarized as the percentages of the available data that fall within the green or orange categories (because only one assessment criterion is being used). The same process is followed for the biomarkers of effect, and this time the percentages for all three colours will be calculated. The percentages of each colour for exposure and effect markers can be used as a graphical summary of these categories of markers.

The next stage in integration of biological measurements is to combine the summaries for exposure and effects markers into an overall summary. This is calculated as the mean of the two summary values for each colour, doubling the weight of the red value for biomarkers of effect to take account of the absence of red assessments of biomarkers of exposure. The resulting spectrum summarizes the assessments of biomarkers of both exposure and effects.

The use of traffic light assessments based upon two assessment criteria considering background responses and responses is consistent with OSPAR assessments of chemical monitoring data. Table 2 repeats the data in Table 1, together with additional lines summarizing chemical monitoring data on contaminants in biota, or other relevant measurements, such as dissolved concentrations of contaminants in water derived from passive samplers. Chemical concentrations in organisms or surrounding water may be considered as expressions of the contaminant stress to which organisms are exposed. It may therefore be appropriate, as suggested in Table 2 to combine these with an integrated assessment of biomarkers of exposure to give an overall summary expression of exposure, and then as a second step to combine this with an integrated expression of markers of effect to obtain an overall expression of environmental status in relation to contaminants and fish or shellfish.

**Good Environmental Status (GES) under MSFD**

Good Environmental Status (GES) will be assessed under MSFD in relation to a series of qualitative Descriptors. Biological effects measurements are most relevant to Descriptor 8: Contaminant concentrations should not give rise to pollution effects.

The Task Groups organized through ICES and JRC to provide advice on the practical implementation of the Descriptors through criteria and methodological standards have not yet competed their work, and therefore it is not yet possible to have a clear view of the approach the that the Descriptor 8 Group will recommend. However, initial indications are that the Group will propose that use is made of a range of quality standards for contaminants that have already been developed, such as EQS
values for contaminant concentrations in water, and EACs for contaminants in sediment and biota.

If this is the case, this would correspond to the upper assessment criterion on Figure 5 as being of particular importance. However, interpretation of the data for biomarkers of effect is improved by the inclusion of biomarkers of exposure in data assessments, for example in assisting to identify the contaminants giving rise to the higher level effects for which upper assessment criteria are available.

SKIMC considered that the wide range of contaminants that were likely to be declared of interest under MSFD (e.g. substances on the WFD or OSPAR (for example) priority hazardous substances lists) would mean that it would be unlikely that the range of effects measurements currently included in the OSPAR fish and shellfish integrated monitoring schemes would be able, in most cases, to provide sufficiently characteristic responses at whole organism or tissue level that would allow clear linkages to be made between the observed effects and the specific causal substance or substances.

However, SKIMC considered that whole organism or tissue level responses could be used to indicate where contaminants were giving rise to pollution effects, or were contributing to the occurrence of pollution effects. Examples of relevant measurements in this context were the general stress indicator lysosomal stability, a range of fish diseases, some whole organism bioassays that did not rely on high degrees of concentration to derive test solutions from environmental matrices (water, sediment or biota). The role of subcellular responses would then be to direct attention towards potential causative agents, building on knowledge of the sensitivity of these responses to groups of environmental contaminants.

It is likely that the results of chemical monitoring programmes for MSFD will exhibit a range of comparisons with assessment criteria; some values exceeding the assessment criteria, and others not. Similarly, results are likely to vary with location within in the large assessment areas. It is not yet clear how this expected variation in data will be handled in coming to an overall conclusion regarding compliance with GES for Descriptor 8. However, full use of available information on pollution effects will assist in ensuring that data are interpreted reliably.
Table 1: Two worked examples of integrated assessment of biological effects measurements, distinguishing between biomarkers of exposure and biomarkers of effect.

Outline integration process A

<table>
<thead>
<tr>
<th>Biomarkers of exposure</th>
<th>Green</th>
<th>Orange</th>
<th>Red</th>
<th>Green</th>
<th>Orange</th>
<th>Red</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biomarker A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biomarker B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biomarker C</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biomarker D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biomarker E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exp. Biomarker summary %</td>
<td>40</td>
<td>60</td>
<td>0</td>
<td>60</td>
<td>40</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Biomarkers of effect</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Biomarker F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biomarker G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biomarker H</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biomarker I</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Summary %</td>
<td>25</td>
<td>50</td>
<td>25</td>
<td>0</td>
<td>25</td>
<td>75</td>
</tr>
</tbody>
</table>

Overall

Red *2

29  49  22

22  24  55
Table 2: Two worked examples of integrated assessment of biological effects measurements, distinguishing between biomarkers of exposure and biomarkers of effect, together with supporting chemical analytical information.

Outline integration process B

<table>
<thead>
<tr>
<th>Markers of exposure</th>
<th>Green</th>
<th>Orange</th>
<th>Red</th>
<th>Green</th>
<th>Orange</th>
<th>Red</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biomarker A</td>
<td>Green</td>
<td>Orange</td>
<td>Red</td>
<td>Green</td>
<td>Orange</td>
<td>Red</td>
</tr>
<tr>
<td>Biomarker B</td>
<td>Green</td>
<td>Orange</td>
<td>Red</td>
<td>Green</td>
<td>Orange</td>
<td>Red</td>
</tr>
<tr>
<td>Biomarker C</td>
<td>Green</td>
<td>Orange</td>
<td>Red</td>
<td>Green</td>
<td>Orange</td>
<td>Red</td>
</tr>
<tr>
<td>Biomarker D</td>
<td>Green</td>
<td>Orange</td>
<td>Red</td>
<td>Green</td>
<td>Orange</td>
<td>Red</td>
</tr>
<tr>
<td>Biomarker E</td>
<td>Green</td>
<td>Orange</td>
<td>Red</td>
<td>Green</td>
<td>Orange</td>
<td>Red</td>
</tr>
<tr>
<td>Exp. Biomarker summary %</td>
<td>40</td>
<td>60</td>
<td>0</td>
<td>60</td>
<td>40</td>
<td>0</td>
</tr>
<tr>
<td>Chemical analysis</td>
<td>20</td>
<td>45</td>
<td>35</td>
<td>45</td>
<td>20</td>
<td>35</td>
</tr>
<tr>
<td>e.g passive samplers</td>
<td>10</td>
<td>50</td>
<td>40</td>
<td>20</td>
<td>30</td>
<td>50</td>
</tr>
<tr>
<td>Chemical analysis summary %</td>
<td>15</td>
<td>48</td>
<td>38</td>
<td>33</td>
<td>25</td>
<td>43</td>
</tr>
<tr>
<td>Markers of exposure summary %</td>
<td>23</td>
<td>45</td>
<td>32</td>
<td>38</td>
<td>27</td>
<td>35</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Biomarkers of effect</th>
<th>Green</th>
<th>Orange</th>
<th>Red</th>
<th>Green</th>
<th>Orange</th>
<th>Red</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biomarker F</td>
<td>Green</td>
<td>Orange</td>
<td>Red</td>
<td>Green</td>
<td>Orange</td>
<td>Red</td>
</tr>
<tr>
<td>Biomarker G</td>
<td>Green</td>
<td>Orange</td>
<td>Red</td>
<td>Green</td>
<td>Orange</td>
<td>Red</td>
</tr>
<tr>
<td>Biomarker H</td>
<td>Green</td>
<td>Orange</td>
<td>Red</td>
<td>Green</td>
<td>Orange</td>
<td>Red</td>
</tr>
<tr>
<td>Biomarker I</td>
<td>Green</td>
<td>Orange</td>
<td>Red</td>
<td>Green</td>
<td>Orange</td>
<td>Red</td>
</tr>
<tr>
<td>Summary %</td>
<td>25</td>
<td>50</td>
<td>25</td>
<td>0</td>
<td>25</td>
<td>75</td>
</tr>
<tr>
<td>Overall</td>
<td>24</td>
<td>48</td>
<td>28</td>
<td>19</td>
<td>26</td>
<td>55</td>
</tr>
</tbody>
</table>
6 Development of assessment criteria for EROD, bile metabolites, DNA adducts of PAHs, DR-CALUX and aspects of fish disease related to contaminants

WKIMC 2009 reviewed the assessment criteria for:

a) Lysosomal stability
b) EROD
c) DNA adducts of PAHs
d) Bile metabolites of PAHs
e) Fish disease related to contaminants
f) DR-CALUX

In doing so, WKIMC took account of comments made at previous meetings (e.g. WKIMON IV, SGIMC 2009, WGBEC) regarding the need to update and amend these criteria. Specific comments regarding some of the documents are in the subsections of Section 4 below. The revised assessment criteria are summarized in Table A in Section 6 of this report.

6.1 Assessment criteria for fish diseases

T. Lang and W. Wosniok gave an introduction on the development of the Fish Disease Index (FDI) approach in the ICES Working Group on Pathology and Diseases of Marine Organisms (WGPDMO) and outlined its concept, components and major strengths (see Annex 4). A reference was made to the ICES fish disease database (part of the ICES Data Centre: Environment Data) and to previous data assessments, e.g. based on OSPAR requests in relation to the OSPAR CEMP and the OSPAR QSR 2010.

In the framework of ICES/OSPAR WKIMC, suggestions were made for additions to the original FDI approach in order to meet the OSPAR requirements as regards the establishment of assessment criteria for biological effects of contaminants. The new assessment procedure proposed can be seen as a combination of the original FDI approach with the OSPAR assessment criteria approach.

For this purpose, the FDI disease category ‘liver histopathology’ was divided into two new categories, namely ‘liver histopathology: non-specific’ and ‘liver histopathology: contaminant-specific’, resulting in the following four disease categories to be used for assessment purposes:

1)Externally visible diseases
2)Liver histopathology: non-specific
3)Liver histopathology: contaminant-specific
4)Macroscopic liver neoplasms

By applying these categories, it is possible to separate between contaminant-specific and non-specific biological effects in relation to fish diseases.

It was recommended to conduct the contaminant-specific assessment on the basis of the FDIs for the categories ‘liver histopathology: contaminant-specific’ and ‘macroscopic liver neoplasms’. Assessment criteria for these categories have been developed and are contained in Table 2, Annex 4. The assessment criteria for ‘liver histopathology: contaminant-specific’ have been derived based on a dataset from the Cefas Weymouth Laboratory (DEFRA 2009, in press) with some modifications.
The assessment criteria developed cannot only be used when long-term data are available, but also in one-off surveys, e.g. as part of exploratory monitoring, or in newly established fish disease monitoring programmes.

It was further recommended to do a parallel but separate assessment of the non-specific categories ‘externally visible diseases’ and ‘liver histopathology: non-specific’ in order to provide additional information on significant changes in disease status that are useful for a more integrated assessment of contaminant effects. For this purpose, the original FDI strategy as developed by the ICES Working Group on Pathology and Diseases should be used (Lang and Wosniok, 2008; ICES, 2009). The possibility of developing numerical assessment criteria for externally visible diseases and for non-specific liver histopathology will further be investigated. For one-off surveys or newly started monitoring programmes, the FDI approach needs to be modified and ways for this will be explored by the ICES Working Group on Pathology and Diseases of Marine Organisms (WGPDMO).

It was emphasized that the ICES fish disease database is the appropriate data source for the fish disease assessments. However, while the submitted data on externally visible disease are sufficient for a comprehensive assessment, there is still an apparent lack of data on liver histopathology and macroscopic liver neoplasms (repeatedly noted by WGPDMO and, in 2009, also by OSPAR ASMO). ICES Member Countries/OSPAR Contracting Parties are therefore encouraged to generate appropriate data according to established guidelines as part of their national marine monitoring programmes and to submit the data obtained (as well as the historical ones) to ICES.

It was further emphasized that fish diseases should not only be seen in the context of monitoring and assessing biological effects of contaminant, but have much wider applications and perspectives because they are appropriate indicators of general ecosystem health. As such, they are of relevance not only for the MSFD Task Group Descriptor 8 (contaminants) but certainly also for other MSFD Task Groups.

Because the ICES WGPDMO has developed and adopted the FDI approach in its original form, the modifications proposed to the FDI assessment approach should be reviewed by WGPDMO at its next meeting in February 2010 and further action needed should be taken. Possible changes required should subsequently be considered by OSPAR.

**Literature cited**


### 6.2 EROD Assessment criteria

Different fish species used as target species in different monitoring programmes of biological effects in European waters and the background reference values in cod (*Gadus morhua*), flounder (*Platichthys flesus*), dab (*Limanda limanda*), dragonet (*Callionymus lyra*), plaice (*Pleuronectes platessa*), four-spotted megrim (*Lepidorhombus boscii*), red mullet (*Mullus barbatus*) and haddock (*Melanogrammus aeglefinus*), have been estimated (Table 1). 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 for these species.

Background responses (BR) for cod, flounder and dab were developed in ICES/OSPAR WKIMON III meeting (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 (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 Science Institute (Scotland, UK) during the workshop. Part of these results derived from the midterm 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 Science Institute from 2001 to 2009. Because significant sexual differences were observed in these two datasets, only males were considered.
Table 1. EROD Background responses in fish target species used in biomonitoring programmes around European waters.

<table>
<thead>
<tr>
<th>EROD ASSESSMENT CRITERIA</th>
<th>SAMPLING SEASON</th>
<th>BOTTOM WATER TEMPERATURE RANGE</th>
<th>SIZE LENGTH</th>
<th>SEX</th>
<th>EROD BR ACTIVITY</th>
<th>UPPER LIMIT OF EROD BR ACTIVITY</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>°C</td>
<td>cm</td>
<td></td>
<td></td>
<td></td>
<td>Median (pmol/min/mg prot)</td>
<td></td>
</tr>
<tr>
<td>Dab (Limanda Limanda)</td>
<td>August-November</td>
<td>[10-18]</td>
<td>12-25</td>
<td>Females and/or males</td>
<td>&lt;30°</td>
<td>&lt;152°†</td>
<td>1034</td>
</tr>
<tr>
<td>European flounder</td>
<td>August-November</td>
<td>[10-18]</td>
<td>20-25</td>
<td>Females and/or males</td>
<td>&lt;14°</td>
<td>&lt;24°†</td>
<td>30</td>
</tr>
<tr>
<td>Gadus morhua</td>
<td>August-November</td>
<td>[10-18]</td>
<td>30-45</td>
<td>Females and/or males</td>
<td>&lt;78°</td>
<td>&lt;151°†</td>
<td>74</td>
</tr>
<tr>
<td>Four spotted megrim</td>
<td>September-October</td>
<td>[11.7-12.7]</td>
<td>18-22</td>
<td>Females and/or males</td>
<td>&lt;12°</td>
<td>&lt;13°†</td>
<td>317</td>
</tr>
<tr>
<td>Dragonet (Callionymus</td>
<td>September-October</td>
<td>[12.0-12.8]</td>
<td>15-22</td>
<td>Females and/or males</td>
<td>&lt;144°*</td>
<td>&lt;202°*</td>
<td>159</td>
</tr>
<tr>
<td>lyra)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Red mullet (Mullus</td>
<td>April</td>
<td>[13.3-15.3]</td>
<td>12-18</td>
<td>Males</td>
<td>&lt;85°*</td>
<td>&lt;208°*</td>
<td>40</td>
</tr>
<tr>
<td>barbatius)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pleuronectes platessa</td>
<td>January</td>
<td>[17.5-22.5]</td>
<td>7-10</td>
<td>Males</td>
<td>&lt;3.71°</td>
<td>&lt;9.49°†</td>
<td>116</td>
</tr>
<tr>
<td>Haddock (Melanogrammus</td>
<td>August</td>
<td>[5-10]</td>
<td>33-55</td>
<td>Females and/or males</td>
<td>&lt;7.2°/&lt;215°</td>
<td>&lt;162°/&lt;421°</td>
<td>20 / 23</td>
</tr>
<tr>
<td>aeglefinus)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saithe (Pollachius</td>
<td>September</td>
<td>[5-10]</td>
<td>40-100</td>
<td>Females and/or males</td>
<td>&lt;57°†</td>
<td>&lt;142°†</td>
<td>21</td>
</tr>
<tr>
<td>virens)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Herring (Clupea harengus)</td>
<td>November</td>
<td>[5-10]</td>
<td>22-33</td>
<td>Females and/or males</td>
<td>&lt;10°†</td>
<td>&lt;23°†</td>
<td>24</td>
</tr>
</tbody>
</table>

6.3 Determination of threshold level of significant effects 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:response 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). These data have 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 herring and halibut exposed to oil (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 turbot at this oil concentration (Jonsson et al., in prep.). This value may be revised as new data to determine chronic effect levels in cod emerge.

References


The output from the detailed review of assessment criteria carried out at SKIMC 2009, including existing values where they have been recommended for retention, and revised values arising from the work at this meeting is tabulated in Table A below:

**Table A: Summary of current proposals for assessment criteria (WKIMC October 2009).**

<table>
<thead>
<tr>
<th>Biological Effect</th>
<th>Qualifying Comments</th>
<th>Background Response Range</th>
<th>Elevated Response Range</th>
<th>High and Cause for Concern Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>VTG in plasma; μg/l</td>
<td>Cod</td>
<td>LOD to 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Flounder</td>
<td>LOD to 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reproduction in eelpout; mean frequency (%)</td>
<td>Malformed larvae</td>
<td>0 - 1</td>
<td>&gt; 1-2</td>
<td>&gt; 2</td>
</tr>
<tr>
<td></td>
<td>Late dead larvae</td>
<td>0 - 2</td>
<td>&gt; 2-3</td>
<td>&gt; 3</td>
</tr>
<tr>
<td></td>
<td>Growth / retarded larvae</td>
<td>0 – 4</td>
<td>&gt; 4-6</td>
<td>&gt; 6</td>
</tr>
<tr>
<td>EROD; pmol/min/ mg protein S9* pmol/min/ mg microsomal protein</td>
<td>Cod</td>
<td>≤ 151*</td>
<td>&gt;151*</td>
<td>not relevant</td>
</tr>
<tr>
<td></td>
<td>Dab (males)</td>
<td>≤ 97</td>
<td>&gt;97</td>
<td>not relevant</td>
</tr>
<tr>
<td></td>
<td>Dab (females)</td>
<td>≤ 142</td>
<td>&gt;142</td>
<td>not relevant</td>
</tr>
<tr>
<td></td>
<td>Flounder</td>
<td>≤ 34</td>
<td>&gt;34</td>
<td>not relevant</td>
</tr>
<tr>
<td></td>
<td>Four-spotted megrim</td>
<td>≤ 13*</td>
<td>&gt;13*</td>
<td>not relevant</td>
</tr>
<tr>
<td></td>
<td>Plaice (males)</td>
<td>≤ 3.7</td>
<td>&gt; 3.7</td>
<td>not relevant</td>
</tr>
<tr>
<td>Bile metabolites; 1-OH pyrene (μg/ml; 341/383 nm fluorescence) *synchronous scan fluorescence</td>
<td>Dab</td>
<td>≤ 0.15*</td>
<td>&gt; 0.15*</td>
<td>not relevant</td>
</tr>
<tr>
<td></td>
<td>Cod</td>
<td>≤ 1.1</td>
<td>&gt; 1.1</td>
<td>not relevant</td>
</tr>
<tr>
<td></td>
<td>Flounder</td>
<td>≤ 1.3</td>
<td>&gt; 1.3</td>
<td>not relevant</td>
</tr>
<tr>
<td></td>
<td>Haddock</td>
<td>≤ 1.9</td>
<td>&gt; 1.9</td>
<td>not relevant</td>
</tr>
<tr>
<td>Bile metabolites; HPLC (fluorescence detection; ng/g)</td>
<td>Dab</td>
<td>≤ 13</td>
<td>&gt; 13</td>
<td>not relevant</td>
</tr>
<tr>
<td></td>
<td>Cod</td>
<td>≤ 13</td>
<td>&gt; 13</td>
<td>not relevant</td>
</tr>
<tr>
<td></td>
<td>Flounder</td>
<td>≤ 232</td>
<td>&gt; 232</td>
<td>not relevant</td>
</tr>
<tr>
<td></td>
<td>Haddock</td>
<td>≤ 10</td>
<td>&gt; 10</td>
<td>not relevant</td>
</tr>
<tr>
<td>DR-CALUX</td>
<td>Sediment</td>
<td>&lt; 10</td>
<td>&gt;10 - &lt;40</td>
<td>&gt; 40</td>
</tr>
<tr>
<td>DNA adducts; nm adducts / mol DNA</td>
<td>Dab</td>
<td>≤ 1.0</td>
<td>&gt; 1.0</td>
<td>(&gt; 6)</td>
</tr>
<tr>
<td></td>
<td>Flounder</td>
<td>≤ 1.0</td>
<td>&gt; 1.0</td>
<td>(&gt; 6)</td>
</tr>
<tr>
<td></td>
<td>Cod</td>
<td>≤ 1.6</td>
<td>&gt; 1.6</td>
<td>&gt; 6</td>
</tr>
<tr>
<td></td>
<td>Haddock</td>
<td>≤ 3.0</td>
<td>&gt; 3.0</td>
<td>(&gt; 6)</td>
</tr>
<tr>
<td>Biological Effect</td>
<td>Qualifying comments</td>
<td>Background Response Range</td>
<td>Elevated Response Range</td>
<td>High and Cause for Concern Response</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-----------------------------------------------</td>
<td>---------------------------</td>
<td>-------------------------</td>
<td>-------------------------------------</td>
</tr>
<tr>
<td>Bioassays; % mortality</td>
<td>Sediment, Corophium</td>
<td>0-30</td>
<td>&gt; 30&lt; 60</td>
<td>&gt; 60</td>
</tr>
<tr>
<td></td>
<td>Sediment, Arenicola</td>
<td>0-10</td>
<td>&gt; 10&lt; 50</td>
<td>&gt; 50</td>
</tr>
<tr>
<td></td>
<td>Water, copepod</td>
<td>0-10</td>
<td>&gt; 10&lt; 50</td>
<td>&gt; 50</td>
</tr>
<tr>
<td>Bioassays; % abnormality</td>
<td>Water, bivalve embryo</td>
<td>0-20</td>
<td>&gt; 20&lt; 50</td>
<td>&gt; 50</td>
</tr>
<tr>
<td></td>
<td>Water, sea urchin embryo</td>
<td>0-10</td>
<td>&gt; 10&lt; 50</td>
<td>&gt; 50</td>
</tr>
<tr>
<td>Bioassay; % growth</td>
<td>Water, sea urchin embryo</td>
<td>0-20</td>
<td>&gt; 20&lt; 50</td>
<td>&gt; 50</td>
</tr>
<tr>
<td>Lysosomal stability; minutes</td>
<td>Cytochemical; all species</td>
<td>&gt; 20</td>
<td>≤ 20≤ 10</td>
<td>&lt; 10</td>
</tr>
<tr>
<td></td>
<td>Neutral Red Retention: all species</td>
<td>&gt; 120</td>
<td>≤ 120&lt; 50</td>
<td>≤ 50</td>
</tr>
<tr>
<td>Fish Disease Index – Externally visible diseases</td>
<td>Cutpoints of the FDI assessment statistic are the 2.5% and the 97.5% quantiles of the assessment statistic. Their numerical values depend on the amount of data and are determined by simulation. Values below the 2.5% quantile are considered as “desirable”, values between and including the quantiles are considered as “indifferent”, values above the 97.5% quantile are considered as “undesirable” (raising concern).</td>
<td>&lt; 2.5% quantile</td>
<td>2.5-97.5 % quantiles</td>
<td>&gt; 97.5% quantile</td>
</tr>
<tr>
<td>Fish Disease Index – non-specific liver histopathology</td>
<td>Cutpoints of the FDI assessment statistic are the 2.5% and the 97.5% quantiles of the assessment statistic. Their numerical values depend on the amount of data and are determined by simulation. Values below the 2.5% quantile are considered as “desirable”, values between and including the quantiles are considered as “indifferent”, values above the 97.5% quantile are considered as “undesirable” (raising concern).</td>
<td>&lt; 2.5% quantile</td>
<td>2.5-97.5 % quantiles</td>
<td>&gt; 97.5% quantile</td>
</tr>
<tr>
<td>Fish Disease Index – contaminant-specific liver histopathology</td>
<td>Dab, Flounder Cutpoints of the FDI assessment statistic are the 2.5% and the 97.5% quantiles of the assessment statistic. Their numerical values depend on the amount of data and are determined by simulation. Values below the 2.5% quantile are considered as “desirable”, values between and including the quantiles are considered as “indifferent”, values above the 97.5% quantile are considered as “undesirable” (raising concern).</td>
<td>&lt; 2.5% quantile</td>
<td>2.5-97.5 % quantiles</td>
<td>&gt; 97.5% quantile</td>
</tr>
</tbody>
</table>
### Biological Effect

At WKIMC 2009, the following assessment criteria where proposed that should ideally be used in addition to the above FDI approach. However, they can also be used independently:

Dab, Flounder
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.

<table>
<thead>
<tr>
<th>Biological Effect</th>
<th>Qualifying Comments</th>
<th>Background Response Range</th>
<th>Elevated Response Range</th>
<th>High and Cause for Concern Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fish Disease Index – macroscopic liver neoplasms</td>
<td>Dab, Flounder</td>
<td>Cutpoints of the FDI assessment statistic are the 2.5% and the 97.5% quantiles of the assessment statistic. Their numerical values depend on the amount of data and are determined by simulation. Values below the 2.5% quantile are considered as “desirable”, values between and including the quantiles are considered as “indifferent”, values above the 97.5% quantile are considered as “undesirable” (raising concern).</td>
<td>&lt; 2.5% quantile</td>
<td>2.5-97.5% quantiles</td>
</tr>
</tbody>
</table>

FDI < 2

FDI ≥ 2

FDI ≥ 2
<table>
<thead>
<tr>
<th>Biological Effect</th>
<th>Qualifying Comments</th>
<th>Background Response Range</th>
<th>Elevated Response Range</th>
<th>High and Cause for Concern Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fish Disease Index – contaminant-specific liver histopathology</td>
<td>Dab, Flounder Cutpoints of the FDI assessment statistic are the 2.5% and the 97.5% quantiles of the assessment statistic. Their numerical values depend on the amount of data and are determined by simulation. Values below the 2.5% quantile are considered as “desirable”, values between and including the quantiles are considered as “indifferent”, values above the 97.5% quantile are considered as “undesirable” (raising concern). At WKIMC 2009, the following assessment criteria where proposed that should ideally be used in addition to the above FDI approach: However, they can also be used independently: Dab, Flounder 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.</td>
<td>&lt; 2.5% quantile</td>
<td>2.5–97.5% quantiles</td>
<td>&gt; 97.5% quantile</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FDI &lt; 2</td>
<td>FDI ≥ 2</td>
<td>FDI ≥ 2</td>
</tr>
</tbody>
</table>
7 Presentation of proposals for the adoption and utilization of assessment criteria for the biological effects measurements indicated above

The main areas of progress made by SKIMC have been the development of a wider range of assessment criteria for biological effects measurements, the proposal of an approach to an integrated assessment procedure for biological effects measurements, and an indication of how these assessments could be integrated with chemical monitoring data. SKIMC recommended that these outputs should be forwarded to OSPAR, and suggested that the principles might be of use to the GES Descriptor 8 Task Group.

There was not time during the meeting to apply the proposed assessment procedure to field data. SKIMC suggested that this could be trialled through the ICON project, which had as an objective the field demonstration of the OSPAR approaches to integrated monitoring.

As an additional exercise, it was suggested that opportunities should be sought to apply the assessment procedure to national datasets, and that the results of these be made available for review by WGBEC and SGIMC.
8 Other related business

8.1 SGIMC work’s contribution to a risk based approach of produced water offshore

Oil and gas industry operating offshore in the OSPAR region use a risk based approach to manage discharges of produced water. This is especially the case in the Norwegian sector of the continental shelf where the regulation implies a no harmful effect goal. Risk tools have been developed (DREAM, EIF) to assess by predictions how this goal can be achieved by implementing different produced water management measures (cleaning, re-injection etc.). For the control monitoring of the achievement of no harmful effect goals a programme using biological effect measurements has been implemented (Water Column effect Monitoring programme; WCM). Work is ongoing to link the field measured biological effect responses to the predicted risk values, so that the achievement of the no harmful effect goal can be assessed. (The biological effect measurements will serve as risk indicators). Because this programme includes several of the biological effect methods that are treated in SGIMC work, the assessment criteria established here will contribute well to strengthen the interpretation basis for the data obtained in the WCM. The WCM programme is based on caged organisms, while the SGIMC assessment criteria are aimed at wild-caught organisms. Nevertheless, the assessment criteria will undoubtedly be useful in the WCM context.

In addition, the oil and gas operators in the Norwegian sector have also the responsibility to carry out another type of survey for the so-called ‘condition monitoring’ which covers larger areas than the WCM. This is based on wild-caught fish, and the biological effect methods included in this programme are well covered by the SGIMC work. The SGIMC assessment criteria in the relevant fish species will be fully applicable for this condition monitoring programme.

The environmental management of the maritime areas where oil and gas production will take place in Norway in future will be ‘risk based’, and the state-of-the environment will be monitored by ‘environmental indicators’ (White paper no.8, 2005-6). The biological effect methods and their assigned assessment criteria included in SGIMC work may also have great relevance as such environmental indicators, because they both include species and assessment values that have been derived from these areas (e.g. the Barents Sea).

In the other oil and gas producing areas within the OSPAR region there are not at present so clear links between risk based management of produced water and biological effect monitoring based on the methods in the SGIMC work. To the extent that the produced water management is risk based, the risk indicators used by the industry are here mainly related to chemical concentrations in the discharged water. The SGIMC methods will in these areas also be relevant and applicable for state-of-the environment assessments on corresponding scales as in the ‘condition monitoring’ in the Norwegian sector described above.

8.2 Array technologies – Genomic analysis

Current biological effects techniques used in environmental health assessment are an assemblage of bioassays, assays for specific inhibition of enzymes, induction of proteins, pollutant metabolites, DNA adducts, physiological responses and pathology. These utilize a wide variety of techniques and instrumental platforms and require staff with expertise in many disciplines. In human disease and pharmacological diag-
nosis many are not possible because they are too invasive and also contradict the 3R's policy, and so many tests have been replaced by the application of 'omics techniques, including genomics (transcriptomics and population genetics), metabolomics and proteomics. 'Omic tools allow simultaneous analysis of a wide range changes in gene expression, protein (e.g. enzyme) and physiological metabolite (e.g. lipid, aminoacid) profiles within the cells of an organism. When applied to animals exposed to toxic chemicals, these can provide a "fingerprint" which can be used to identify the underlying biochemical mechanisms of pathology and toxicity.

A number of national (UK NERC, USEPA [2004] and International workshops (NERC PGP [Van Aggelen et al., 2009] SETAC Pellston ) have made recommendations for the applicability of 'omic analyses, especially genomics, in human and ecotoxicological health assessments. The current consensus of opinion is that genomics data can usefully contribute to a weight of evidence approach, but used alone are not sufficient for risk assessment in regulatory toxicology.

In environmental toxicology, analysis of gene expression profiles (transcriptomics) is the only procedure that is sufficiently well developed at present to be considered in environmental health or risk assessments. Gene expression profiles (or "fingerprints") associated with chemical exposure and the ensuing toxicity and pathology can be determined by use of DNA microarrays. High density DNA oligonucleotide microarrays (which are species-specific) have been developed for Atlantic salmon as well as for some “laboratory test species” (including medaka and zebrafish). Whilst not having such a comprehensive gene coverage, cDNA arrays have the advantage that they can be utilized across species within a genus e.g. all Pleuronectidae with the flounder array, all Sparidae with the sea bream array and all Cyprinidae with the zebrafish array (Cohen et al., 2007). Medium density cDNA arrays have been developed for a number of fish species, including sentinel species such as carp, the European flounder and stickleback, as well as for other environmentally or commercially relevant species such as cod, halibut and sea bream. Many of the fish arrays developed so far have been described recently in a special issue of the Journal of Fish Biology (2008, Vol. 72).

Responses to laboratory exposures of fish to a number of classes of toxic chemicals (e.g. Cd, Cu, estrogenic compounds, PAHs, PCBs, etc.) and simple mixtures have successfully identified modulation of important biochemical pathways (e.g., cholesterol and sex steroid metabolizm, immune response) and have also identified numerous diagnostic biomarker genes of both exposure and effect, thus confirming the validity of the technique and providing reference data (e.g. Williams et al., 2008). Such an approach is also extremely useful both for biomarker discovery and for determining modes of action (MOAs) of novel compounds.

However, animals sampled from the environment have usually been exposed to multiple uncharacterized stimuli, such that individual classical biomarkers are unlikely to summarize such a complex response. Indeed, it is well recognized that estrogen and estrogen-like compounds modulate the PAH-response of CYP1A (the EROD catalyst) and the response of metallothionein to Cd. Thus one of the key needs in ecotoxicology is to relate responses in laboratory exposures to those discovered in environmental samples. cDNA arrays offer an approach to this problem, but a major, and so far unresolved, challenge is the interpretation of the very large datasets involved, especially when other data such as histopathology are also incorporated in the analysis.

In the EU-funded FP7 Genipol project, some progress has been made toward demonstrating such a linkage between field and laboratory data in immature male flounders
sampled from a number of sites in the UK and FRG. Gene expression profiles in livers of field sampled flounders have been compared with the responses observed in controlled laboratory exposures to individual compounds by multivariate statistical methods, and the probabilities of exposure to these compounds have been assessed (UK DEFRA Charting Progress 2, 2009). Whilst the method used was relatively crude and requires refinement, it nevertheless does imply that bioavailability and responses are in accord with the known pollutant profiles of these environments, allowing a preliminary status assessment of the flounder. A more sophisticated analysis of the gene expression profiles was performed using a genetic algorithm (Galgo) to classify the sites of origin of the fish. By incorporating prior knowledge of biomarker genes obtained from the laboratory toxicant-exposure experiments, a final model of 17 genes was developed to distinguish between the sampling sites on the basis of their gene expression profiles. Probes for these genes were then incorporated into a qPCR array which successfully classified independent flounder samples by site (Falciani et al., 2008). Both approaches indicate that changes in gene expression derived from short-term laboratory exposures can, in some cases, be used to interpret responses of fish from the environment despite these being a combination of both acute and chronic responses. Of additional interest was that accompanying genetic analysis using neutral microsatellite markers and microsatellites in biomarker and nuclear transcription factor genes failed to reveal any marked population differences.

Whilst these studies may be considered a “proof of principle”, it is not envisaged that current DNA microarray technology should be utilized in routine environmental impact assessment. Use of high density microarrays is likely to be confined to the characterization of biological responses to novel pollutants and pre-screening of compounds, or in case studies of environmental sites of concern. It will allow identification of toxicant-affected pathways and processes and elucidation of modes of toxicant action, as well as helping to elucidate the effects of mixtures of toxicants.

For routine application, both qualitative and quantitative analysis of specific transcripts of established and emergent biomarkers with a qPCR array promises a useful advance in environmental risk assessment because it has a number of notable improvements over current classical methods. Sampling is simplified because one small sample (e.g. 20 larvae up to 50 mg tissue) which is freeze clamped on necropsy can be used for all the determinands. Deep frozen samples are stable for years. There is only one analytical platform, and analyses are carried out simultaneously and costs are low. The method has vastly improved sensitivity over current procedures, and has been used for pools of 20 larvae. The range of biomarker genes can easily be extended, currently validated biomarkers for flounder include CYP1A, MT, VTG (Dixon et al., 2002; George et al., 2004) as well as GST and genes related to effects on animal health (e.g. hepcidin (inflammatory response), Diab et al., 2008) and Diablo (a pre-apoptotic gene) Leaver et al., 2009). Two variations are envisaged, one a qualitative array for screening and diagnostic use and the second a quantitative array for status assessment.

Two preliminary validation studies with flounders from UK estuaries have recently been conducted. In the first, agreement with established biomarker measurements was obtained (George et al., 2004) and in the second, expression of well-established biomarker genes correlated with the WFD risk categorization of UK estuaries by the UK EA (Williams et al., NERC KT contract report). It is proposed that this technique should now be taken forward to a full European validation phase.
Literature


8.3 Work programme for SGIMC

The work programme developed at SGIMC 2009 was reviewed at SKIMC in Oct 2009, and the revised programme is shown below as Table C:
Table C: Proposed Work Programme for SGIMC from January 2009 to January 2011.

<table>
<thead>
<tr>
<th>Effect</th>
<th>Task</th>
<th>Responsible member</th>
<th>When</th>
<th>Report to</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>EROD</td>
<td>Organising Workshop</td>
<td>Ian Davies</td>
<td>For Oct 2009</td>
<td>SGIMC 2010</td>
<td>Aberdeen, October 2009</td>
</tr>
<tr>
<td>EROD</td>
<td>Update the initial approach to Background Response assessment criteria and review available EROD data. Investigate conversion factor S9 microsomal. Seasonal cycle approach (Werner)</td>
<td>Ian Davies, Thomas Maes, Kris Cooreman and Concepción Martínez-Gómez</td>
<td>January 2010</td>
<td>SGIMC 2010</td>
<td>To be addressed by and in SGIMC WGBEC in January 2010</td>
</tr>
<tr>
<td>PAH bile metabolites</td>
<td>Done</td>
<td>Dick Vethaak and Ketil Hylland</td>
<td>Mar-09</td>
<td>WGBEC 2009</td>
<td>To be available for Workshop in Oct 2009</td>
</tr>
<tr>
<td>PAH bile metabolites</td>
<td>Expand to SF method, other species</td>
<td>Ketil Hylland (and Ian Davies)</td>
<td>Oct 2009</td>
<td>SGIMC 2010</td>
<td>To be addressed in a workshop in Aberdeen in October 2009</td>
</tr>
<tr>
<td>DNA adducts</td>
<td>Done</td>
<td>Brett Lyons (and Ian Davies)</td>
<td>Oct 2009</td>
<td>SGIMC 2010</td>
<td>To be addressed in a workshop in Aberdeen in October 2009</td>
</tr>
<tr>
<td>DNA adducts</td>
<td>Done for some species, expand to others.</td>
<td>Brett Lyons (and Ian Davies)</td>
<td>Oct 2009</td>
<td>SGIMC 2010</td>
<td>To be addressed in a workshop in Aberdeen in October 2009</td>
</tr>
<tr>
<td>DR-CALUX</td>
<td>Ack grd doc in Chapter 7 now. Refine assessment criteria when more data available. TPC extraction method criteria needed.</td>
<td>Dick Vethaak</td>
<td>Dick &amp; John</td>
<td>WGBEC Jan 10</td>
<td>SGIMC Feb 10</td>
</tr>
<tr>
<td>DR-CALUX</td>
<td>Done cf. above</td>
<td>Dick Vethaak</td>
<td>Oct 2009</td>
<td>SGIMC 2010</td>
<td>To be addressed in a workshop in Aberdeen in October 2009</td>
</tr>
<tr>
<td>DR-CALUX</td>
<td>Complete TIMES series method document</td>
<td>Dick Vethaak</td>
<td>March 2010</td>
<td>For ICES Pub Com</td>
<td></td>
</tr>
<tr>
<td>Liver nodules (neoplasm)</td>
<td>Done</td>
<td>Thomas Lang and</td>
<td>Oct 2009</td>
<td>SGIMC 2010</td>
<td>Report to WGPDMO 2010</td>
</tr>
<tr>
<td>Effect</td>
<td>Task</td>
<td>Responsible Member</td>
<td>When</td>
<td>Report to</td>
<td>Comments</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
<td>--------------------</td>
<td>------</td>
<td>-----------</td>
<td>----------</td>
</tr>
<tr>
<td>Extraction procedures for bioassay methods</td>
<td>Complete TIMES series method document</td>
<td>John Thain</td>
<td>March 2010</td>
<td>WGBEC 2009</td>
<td>Revised draft to be reviewed by WGBEC 2009</td>
</tr>
<tr>
<td>VTG</td>
<td>Establish BAC in monitoring species</td>
<td>MSS, Cefas</td>
<td>For WGBEC 2010</td>
<td>SGIMC 2010</td>
<td>To be reviewed at SGIMC 2010</td>
</tr>
<tr>
<td>VTG</td>
<td>Develop EAC equivalent for monitoring species</td>
<td>Ian Davies and Dick Vethaak</td>
<td>WGBEC 2010</td>
<td>SGIMC 2010</td>
<td></td>
</tr>
<tr>
<td>VTG mRNA</td>
<td>Need values</td>
<td>SGG</td>
<td>Put forward at WGBEC?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intersex in fish</td>
<td>Review Background document</td>
<td>Steve Feist</td>
<td>WGBEC2009</td>
<td>SGIMC 2010</td>
<td>In progress, remind SF.</td>
</tr>
<tr>
<td>Intersex in fish</td>
<td>To develop Background Response and EAC-equivalent assessment criteria</td>
<td>Steve Feist</td>
<td>WGBEC 2010</td>
<td>SGIMC 2010</td>
<td></td>
</tr>
<tr>
<td>Fish Disease Index</td>
<td>Was addressed at WKIMC – address liver histopath.</td>
<td></td>
<td></td>
<td>WGPDMO 2010</td>
<td>There is good progress. OSPAR guideline on fish diseases has been amended in 2008. Future updates by Germany (TL, WW). No action required by SGIMC.</td>
</tr>
<tr>
<td>Lysosomal stability (Neutral Red)</td>
<td>Organising training workshop Draft proposal Permission from ICES/OSPAR</td>
<td>Concepcion Martinez-Gomez</td>
<td>June 2010</td>
<td>SGIMC 2011</td>
<td>Subject of a Workshop to be organized in Spain in June 2010</td>
</tr>
<tr>
<td>Reproductive success (eelpout)</td>
<td>No action required by SGIMC</td>
<td></td>
<td></td>
<td></td>
<td>No action required by SGIMC</td>
</tr>
<tr>
<td>Scope for Growth, Acetyl cholinesterase</td>
<td>Wait for ICON results Update Background Document</td>
<td>JT, Concepcion Martinez-Gomez</td>
<td>WGBEC2010</td>
<td>SGIMC 2010</td>
<td>No action currently required by SGIMC</td>
</tr>
<tr>
<td>Acetyl cholinesterase</td>
<td>To develop Background</td>
<td>Thierry Burgeot</td>
<td>WGBEC 2010</td>
<td>SGIMC 2011</td>
<td></td>
</tr>
<tr>
<td>EFFECT</td>
<td>TASK</td>
<td>RESPONSIBLE MEMBER</td>
<td>WHEN</td>
<td>REPORT TO</td>
<td>COMMENTS</td>
</tr>
<tr>
<td>----------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>-------------------------------</td>
<td>--------</td>
<td>-------------</td>
<td>----------</td>
</tr>
<tr>
<td>Mussel histopathology</td>
<td>ICES Times manuscript including BAC in preparation</td>
<td>Steve Feist + Miren Cajaraville</td>
<td>Dec 2009</td>
<td>WGBEC 2010</td>
<td></td>
</tr>
<tr>
<td>Micronucleus assay + comet assay</td>
<td>Background document and draft BAC</td>
<td>Brett Lyon</td>
<td>WGBEC2010</td>
<td>SGIMC 2010</td>
<td></td>
</tr>
<tr>
<td>MT &amp; ALA-D</td>
<td>Develop BC using recent data</td>
<td>Ketil Hylland</td>
<td>Feb 2010</td>
<td>SGIMC 2010</td>
<td></td>
</tr>
<tr>
<td>Chapter 7</td>
<td>In vitros YES/YAS, ER CALUX</td>
<td>JT / DV</td>
<td>Jan</td>
<td>WGBEC 2010</td>
<td></td>
</tr>
<tr>
<td>Chapter 8</td>
<td>Add Sed &amp; SW elutriate bioassays for invert embryos. Further validate others as more data becomes available</td>
<td>Ricardo Beiras</td>
<td>Jan 2010</td>
<td>WGBEC 2010</td>
<td></td>
</tr>
<tr>
<td>Chapter 9</td>
<td>As above with copepods</td>
<td>As above</td>
<td>As above</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chapter 10</td>
<td>Update BG and ass cri for Whole sediments with amphipods as more data becomes available</td>
<td>As above</td>
<td>As above</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
9 Completion of the Workshop report

It was agreed that the Report would be completed over the next 2–3 weeks (!) by a written procedure, managed by the Co-chairs.
10 Closing of the meeting

The meeting was closed at 1600 hr on 16 October 2009. Matt Gubbins was particularly thanked for the effectiveness of his hosting of the Workshop.
### Annex 1: List of participants

<table>
<thead>
<tr>
<th>Name</th>
<th>Address</th>
<th>Phone/Fax</th>
<th>Email</th>
</tr>
</thead>
</table>
| Kris Cooreman         | ILVO – Fisheries
Ankerstraat 1
8400 Oostende
Belgium               | Tel: +3259569820
Fax: +3259330629         | kris.cooreman@ilvo.vlaanderen.be                                     |
| Ian Davies Co-chair   | Marine Scotland
Marine Laboratory
PO Box 101
375 Victoria Road
Aberdeen AB11 9DB
UK                     | Tel: +44(0)122 429 5468
Fax: +44 (0)1224 295511     | i.m.davies@marlab.ac.uk                                           |
| Stephen George        | Institute of Aquaculture
University of Stirling
Stirling FK9 4LA          | Tel: +44 (0) 1786467922
Fax: +44 (0)1224 295511     | s.g.george@stir.ac.uk
sgg1@stir.ac.uk          |
| Matt Gubbins          | Marine Scotland
Marine Laboratory
PO Box 101
375 Victoria Road
Aberdeen AB11 9DB
UK                     | Tel: +44(0)122 429 5681
Fax: +44 (0)1224 295511     | m.gubbins@marlab.ac.uk                                           |
| Ketil Hylland         | Norwegian Institute for Water Research (NIVA)
PB 173 Kjelsås
N-0411 Oslo
Norway                 | Tel: +47 22185170
Fax: +47 22185200         | ketil.hylland@niva.no                                        |
| Thomas Lang           | VTI Institute of Fisheries
Ecology,
Deichstr 12
27472 Cuxhaven
Germany               | Tel: +49 4721 38034
Fax: +49 4727 33583      | Thomas.lang@vti.bund.de                                      |
| Thomas Maes           | CEFAS
Pakefield Road
NR330HT
 Lowestoft
UK                   | Tel: +44 (0)1502 524433     | thomas.maes@cefas.co.uk                                      |
| Concepción Martinez   | Instituto Español de Oceanografía
Centro Oceanográfico de Murcia
Varadero 1, Lo Pagán
30740 San Pedro del Pinatar (Murcia)
Spain               | Tel: +34 968180500
Fax: +34 9681844441     | cmartinez@mu.ieo.es                                           |
| Steinar Sanni         | IRIS Biomiljø
Mekjarvik 12
N-4070 Randaberg
Norway             | Tel: +47 5187 5504
Fax: +47 5187 5540       | Steinar.Sanni@iris.no                                       |
| John Thain            | CEFAS
Weymouth Laboratory
The Nothe
Barrack Road Weymouth
Dorset DT4 8UB          | Tel: 44 (0) 1305 206600
Fax: 44 (0) 1305 206601     | j.e.thain@cefas.co.uk                                      |
<table>
<thead>
<tr>
<th>Name</th>
<th>Address</th>
<th>Phone/Fax</th>
<th>Email</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dick Vetheaak</td>
<td>Deltares, Unit Marine and Coastal Systems,</td>
<td>Tel: +31 15-2858659 / +31 651232412</td>
<td><a href="mailto:dick.vetheaak@deltasres.nl">dick.vetheaak@deltasres.nl</a></td>
</tr>
<tr>
<td>Co-chair</td>
<td>Section</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ecosystem Analysis and Assessment (ESA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rotterdamseweg 185</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2629 HD Delf</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>The Netherlands</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Werner Wosniok</td>
<td>University of Bremen</td>
<td>Phone +49 42 12 1834 71</td>
<td><a href="mailto:wvosniok@math.uni-bremen.de">wvosniok@math.uni-bremen.de</a></td>
</tr>
<tr>
<td></td>
<td>Institut für Statistik, FB 3</td>
<td>Fax +49 42 12 18 8944</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PO Box 330 440</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>DE -28344 Bremen</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Germany</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Annex 2: Agenda

ICES/OSPAR Workshop on Assessment Criteria for Biological Effects Measurements

Marine Scotland Marine Laboratory, Aberdeen, UK.
14–16 October 2009

Co-Chairs: Ian Davies (UK, OSPAR), and Dick Vet Haak (The Netherlands, ICES).

1) Opening of the meeting;
2) Adoption of draft agenda;
3) Discussion of ToR and background to the meeting;
4) Completion of the development of supporting documents for a coherent set of biological effects measurements linked to planar organic contaminants, particularly PAHs;
5) Clarification of the mutually supporting role of a range of biological effects measurements in relation to OSPAR objectives and GES under MSFD;
6) Development of assessment criteria for EROD, bile metabolites, DNA adducts of PAHs, DR-CALUX and aspects of fish disease related to contaminants;
7) Presentation of proposals for the adoption and utilization of assessment criteria for the biological effects measurements indicated above;
8) Other related business;
9) Completion of the Workshop Report;
10) Closing of the meeting.
Annex 3: Terms of Reference

Terms of Reference for an ICES/OSPAR Workshop on Assessment Criteria for Biological Effects Measurements

Background

1) The Hazardous Substances Strategy of the 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 assessment criteria for chemical monitoring, namely Background Concentrations and Environmental Assessment Criteria. The latter are concentrations of contaminants in sediment or biota below which unacceptable biological effects are unlikely to occur.

2) A parallel activity has been in progress 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 a series of background documents, which in most cases have included proposals for background values (concentrations, activities, scales of effect) for a range of biological effects measurements.

3) 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 additional assessment criteria for biological effects measurements to allow these measurements to be brought into greater play in data assessment in monitoring programmes and environmental quality assessments.

4) In order to progress the development of integrated chemical and biological effects monitoring and assessment expeditiously it is necessary that the role of each method is clearly understood, and that useful assessment criteria are available. SGIMC recommended that a Workshop be held in Aberdeen in week beginning 12 October 2009. The purpose of the workshop is to prepare proposals for a wider suite of assessment criteria for biological effects measurements. In addition to their application in ICES/OSPAR programmes, the work will contribute to a risk based approach to produced water under development of the OSPAR Offshore Industry Committee (OIC). It is likely that such assessment criteria 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).

Objectives

5) The objectives of the workshop are to:

5.1) to complete the development of supporting documents for a coherent set of biological effects measurements linked to planar organic contaminants, particularly PAHs;
5.2) clarify the mutually supporting role of a range of biological effects
measurements in relation to OSPAR objectives and GES under
MSFD;
5.3) develop assessment criteria for EROD, bile metabolites, DNA add-
ducts of PAHs, DR-CALUX and aspects of fish disease related to
contaminants;
5.4) present proposals for the adoption and utilization of assessment cri-
teria for the biological effects measurements indicated above.

Activities
6) The workshop will be organized around the following activities, building
on the activities of WKIMON and SGIMC:

Presentation and discussion of the background information
a) existing background documents on the relevant biological effects
measurements will be presented, reviewed, and updated as necessary.
b) additional new supporting monitoring and research information.

Output from the Workshop
   c) updated background documents
d) reviewed and revised assessment criteria (background conditions)
e) new assessment criteria for thresholds of significant (unacceptable)
   levels of biological effects.

Organisation
7) The workshop will be coordinated by Ian Davies (UK).
8) Participants: the workshop should be attended by experts in the applica-
tion of the methods and in the assessment of data to ensure that the result-
ing advice to OSPAR/ICES is of high quality.
9) Contracting Parties should establish a good link between their experts at-
tending the workshop and those working on the risk based approach in
the OIC framework to ensure that the results are available for consistent
work on hazardous substances across OSPAR thematic work strands.

Preparatory work
10) The following existing documents are relevant to the workshop and will be
collated and distributed prior to the Workshop:
   a) Background documents on the OSPAR Strategy for Hazardous Sub-
      stances, and associated monitoring and assessment programmes (e.g.
      CEMP).
b) OSPAR Background Documents for relevant biological effects (OSPAR
     Publication 2007/333).
c) Reports of meetings of ICES/OSPAR WKIMON and SGIMC groups.
d) CEMP Assessment Manual for contaminants in sediment and biota
   (OSPAR Publication 2008/379).
e) Assessment criteria for biological effects currently adopted by OSPAR,
   for example assessment criteria for imposex in gastropod snails.
Location

The workshop will be held at Marine Scotland Marine Laboratory, Aberdeen, UK.

Time schedule

<table>
<thead>
<tr>
<th>April 2009</th>
<th>ASMO 2009</th>
<th>Agreement of Terms of Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>August–September 2009</td>
<td>Workshop organizer and host organization</td>
<td>Collate and distribute background material and develop programme for the Workshop</td>
</tr>
<tr>
<td>12–16 October 2009</td>
<td>Workshop</td>
<td></td>
</tr>
<tr>
<td>November 2009</td>
<td></td>
<td>Report of the Workshop finalized and made available to ICES and OSPAR Secretariats, and, if agreed, to the ICES/JRC process developing assessment criteria for MSFD Descriptor 8 of Good Environmental Status.</td>
</tr>
<tr>
<td>Early 2010</td>
<td>SGIMC 2010</td>
<td>Output from Workshop reviewed by SGIMC. Proposals for adoption of assessment criteria formulated for adoption by OSPAR through ASMO etc</td>
</tr>
</tbody>
</table>
Annex 4: Assessment criteria for fish diseases

Assessment criteria for fish diseases
T. Lang & W. Wosniok

Introductory remarks
Fish diseases are considered as ecosystem health indicators, reflecting ecologically relevant effects of environmental stressors at the individual and population levels. As such, they differ from other types of indicators that reflect changes at lower levels of biological organization (e.g., molecules, cells) and the ecological relevance of which is considered as low or unclear (e.g., biomarkers of exposure to contaminants.)

The monitoring and assessment of fish diseases are part of the general and the PAH-specific biological effects components of the OSPAR pre-CEMP and the JAMP. Fish diseases monitored belong to three disease categories shown in Table 1 for the common dab (Limanda limanda), the most widely used target species in the North Sea and adjacent waters. The diseases have been recommended for monitoring purposes and relevant guidelines exist (Bucke et al., 1996, Feist et al., 2004).

Table 1: Disease categories included in the general and PAH-specific biological effects component of the pre-CEMP/JAMP.

<table>
<thead>
<tr>
<th>Disease categories</th>
<th>Causes</th>
<th>Link to contaminants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Externally visible diseases</td>
<td>Viral, bacterial, parasitic</td>
<td>Non-specific</td>
</tr>
<tr>
<td>(9 lesions)</td>
<td>infections</td>
<td>Contaminants may contribute (e.g., by</td>
</tr>
<tr>
<td></td>
<td></td>
<td>affecting the immune system), but are not</td>
</tr>
<tr>
<td></td>
<td></td>
<td>the only cause of changes recorded in</td>
</tr>
<tr>
<td></td>
<td>Multifactorial</td>
<td>disease prevalence.</td>
</tr>
<tr>
<td></td>
<td>Unknown</td>
<td></td>
</tr>
<tr>
<td>Liver histopathology</td>
<td>Infections</td>
<td>4 of the 5 lesion types are contaminant-</td>
</tr>
<tr>
<td>(5 lesion types)</td>
<td></td>
<td>specific.</td>
</tr>
<tr>
<td>Macroscopic liver neoplasms (benign and</td>
<td>Contaminants</td>
<td>Contaminant-specific</td>
</tr>
<tr>
<td>malignant tumours)</td>
<td></td>
<td>Contaminants act as initiators and promoters</td>
</tr>
<tr>
<td>(2 lesion types)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Quality assurance of disease studies is achieved through activities of ICES and the BEQUALM programme (http://www.bequalm.org). Guidelines for studies on externally visible diseases, liver histopathology and macroscopic liver neoplasms have been published and became part of the OSPAR JAMP Guidelines for General and for PAH-specific biological effects monitoring (OSPAR 1997, 2008).

At present, fish disease surveys are carried out by OSPAR Contracting Parties on a voluntary basis (pre-CEMP). Countries with regular monitoring programmes designed according to the ICES/BEQUALM guidelines are Germany, The Netherlands and the UK. Other countries such as Belgium, Denmark and Sweden have ceased their monitoring a number of years ago.

Fish disease data generated in the monitoring programmes are submitted to the fish disease database of the ICES Data Centre – Environment Data on a regular basis. Relevant formats and quality assurance procedures have been in place since the mid 1990s.

Since 2006, the ICES Working Group on Pathology and Diseases of Marine Organisms (WGPDMO) has developed a strategy to analyse and assess fish disease data by
applying the Fish Disease Index (FDI) approach. A documentation of the methodology is provided by Lang and Wosniok, 2008 and revisions made since then by ICES, 2009.

In contrast to previous attempts, largely focusing on the analysis and assessment of changes in prevalence of single diseases, the FDI approach was developed with the primary aim to analyse and assess changes in spatial and temporal patterns in the overall disease status of fish, by summarizing information on the prevalence of a variety of common diseases affecting the fish species (see Table 1) as well as their severity grades and effects on the host into a robust numerical value calculated for individual fish and, as mean values, for representative samples from a population.

The common dab (*Limanda limanda*) was selected as a model species for the construction of the FDI approach because most existing data are from fish disease survey with the dab as primary target species. However, the FDI approach is constructed in a way that it can easily be adapted to other fish species for which disease data are available.

Based on an OSPAR request, the WGPDMO carried out an assessment of ICES fish disease data at its 2008 and 2009 meetings (ICES 2008, 2009), the results of which were reviewed by OSPAR ASMO in 2008 and 2009. The results of the assessment will be incorporated in the OSPAR QSR 2010 as a case study.

In 2009, Germany proposed that OSPAR ASMO adopts the FDI formally as an assessment tool and to include externally visible fish diseases, macroscopic liver neoplasms and liver histopathology in the CEMP. However, at the 2009 ASMO meeting, some Contracting Parties expressed reservations with regard to the adoption of the fish disease criteria as formal assessment criteria at the current time and ASMO agreed that it was not possible to adopt the FDI at the ASMO meeting. However, Contracting Parties were encouraged to undertake monitoring of liver neoplasms and liver histopathology and submit the resulting data to the ICES fish disease database in line with pre-CEMP requirements (Note: the reason why only liver neoplasms and histopathology and not externally visible diseases were mentioned is that an apparent lack of liver data was identified in the assessment carried out by WGPDMO, while data on externally visible diseases were considered sufficient for an assessment).

For its 2010 meeting, the WGPDMO was given the Term of Reference to provide an update on the progress achieved in the implementation of the Fish Disease Index (FDI) and the related assessment procedure in marine monitoring and assessment programmes, and review results of the application of the FDI approach on other fish species for which data are available in the ICES Environmental Database (the latter mainly directed to fish disease monitoring in the Baltic Sea as part of the HELCOM activities related to the ecological objective ‘healthy wildlife’ of the HELCOM Baltic Sea Action Plan).

**Basic principles of the Fish Disease Index (FDI) analysis and assessment**

The analysis and assessment of the fish disease data are made on a region-wise basis (as is the monitoring) and calculated regional individual and mean FDI values form the basis of the assignment of assessment criteria.

Components of the FDI calculation are information on:

- presence/absence of diseases;
- severity (mostly 3 grades);
• Effects on host (weighting by expert judgement);
• Adjustment for effects of size, sex, season (based on empirical data).

For an ideal analysis and assessment, the existence of long-term disease data is desirable, from which FDIs are calculated. Based on the long-term data, an observation (= learning) period (ideally 10 observations over a period of 5 years) is defined to be used as a data basis for the assessment done for a subsequent assessment period, e. g., the last 5 years (see scheme in Figure 1). However, the duration of the observation and the assessment period may vary according to the purpose of the monitoring/assessment and the availability of data.

![Figure 1: Schematic illustration of the Fish Disease Index approach (ICES 2009).](image)

With the data from the assessment period, two assessments are made separately: (1) a statistical comparison of mean FDI levels between the observation and the assessment period and (2) a statistical trend analysis in the assessment period. Note that the trend analysis uses only data from the assessment period. The results of the level and trend assessment are visualized using green, yellow or red smiley faces, reflecting an improvement of the disease status, no significant change or a worsening of the disease status. The smiley faces can be placed on a geographical map demonstrating the monitoring regions.

For the time-being, the assessment of fish disease data is done separately for the three CEMP categories externally visible diseases, liver histopathology and macroscopic liver neoplasms. However, it is envisaged that, at a later stage, the categories are combined for a joint assessment reflecting the overall disease conditions.

Some of the advantages of the FDI approach in its original form are:

• The FDI assessment provides information on changes in the overall diseases status of fish and, thus, on general ecosystem health and helps to identify geographical areas of concern.
• The full set of diseases affecting a fish species and, thus, recommended for monitoring is taken into account in the analysis and assessment.
• The FDI approach provides numeric information on the impact of the diseases on the host by applying weighting factors.
• The full range of existing long-term data (implicit for monitoring) is used.
• An assessment can be done without identifying and applying ambiguous numeric background disease levels and assessment criteria.
• Changes in both FDI levels and upward or downward trends are statistically analysed and directly used as basis of the assessment.
• A regional comparison of FDI level change and trend is possible.
• The FDI approach can be applied on a variety of fish species that are monitored for diseases.
• Direct interspecies comparisons as regards FDI level changes and trends are possible, facilitating the application of the FDI approach over larger geographical areas with different fish species and diseases.

Some additions to the FDI approach will be made based on the following aspects:
• If data from a one-off fish diseases survey (e.g., from exploratory monitoring in a geographical region where disease monitoring was not done before) or data from newly implemented monitoring programmes are to be assessed, the FDI assessment strategy can be modified in order to provide background (baseline) data required and to be used as an assessment tool.
• A method that allows a direct comparison of mean FDI levels between regions will be added. However, it has to be taken into account that differences between areas may reflect natural differences in living conditions rather than effects of stressors. Therefore, an assessment only based on regional differences in mean FDI levels is not recommended.
• The assessment of contaminant-specific disease effects should primarily focus on contaminant-specific diseases, i.e., macroscopic liver neoplasms and toxicopathic non-neoplastic and neoplastic liver histopathology.
• The original FDI approach differs from assessment strategies for other indicators (e.g., contaminants, biomarkers) in that it does not apply numeric and universally applicable assessment criteria (e.g., BAC, AC for unacceptable effects) and, therefore, does not fully meet the OSPAR requirements as expressed in the Terms of Reference of the ICES/OSPAR WKIMC 2009. However, such assessment criteria for macroscopic liver neoplasms and for contaminant-specific liver histopathology can be developed (see further below).

Modifications of the FDI approach in order to generate assessment criteria for contaminant-specific effects
In order to meet the OSPAR requirements in relation to the assessment of biological effects of contaminants and the establishment of assessment criteria, some additions were made to the original FDI approach and are detailed below (see Table 2).

A general scheme for the recommended assessment strategy based on the FDI approach is shown in Figure 2. The different disease categories (externally visible diseases, liver histopathology, macroscopic liver neoplasms) are assessed separately. Furthermore, the liver histopathology category is split into (a) ‘non-specific lesions’ (treated in the same way as the externally visible diseases) and (b) ‘contaminant-specific lesions’. This is done in order to separate between disease categories comprising of contaminant-specific from those comprising of non-specific diseases.
The assessment of contaminant effects should be based on contaminant-specific liver histopathology and on macroscopic liver neoplasms (see Figure 2).

Because externally visible diseases and non-specific liver histopathology are considered as non-specific indicators and, thus, contaminants are only one of the contributing environmental factors, the respective FDIs should not directly be used for the assessment of contaminant-specific effects. However, an assessment of the non-specific diseases should be done in addition to the contaminant-specific diseases because (a) the non-specific and the contaminant-specific diseases together characterize the general disease status of the fish and (b) contaminants may also contribute to the non-specific diseases (see Figure 2).

Assessment criteria for externally visible diseases and for non-specific liver histopathology (Table 2): It is unrealistic to expect that all fish in a population sampled are healthy with regard to externally visible or non-specific liver histopathology. Thus, the background level of the FDI cannot be ‘0’. Because of the complex nature of the disease aetiology, it is further unrealistic to assume that the natural background disease level in all geographical regions is identical. Therefore, a universal background level applicable for all geographical regions cannot be provided. The elevated response/above background level concept (equivalent to BAC) is not applicable because all statistically significant increases in FDI level compared with a previous observation period or a significant upward trend in FDI should be regarded as unacceptable.

Assessment criteria for contaminant-specific liver histopathology (Table 2): The definition of background levels and unacceptable levels uses expert judgements on the severity of liver histopathology conditions and is calibrated on the basis of UK data on liver histopathology in dab (DEFRA 2009, in press) because this is the only dataset available so far on histopathological liver lesions generated according to ICES/BEQUALM guidelines. It is recommended to use values of mean FDI < 2 as background level and values of mean FDI ≥ 2 as unacceptable level. A value of mean 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 (reflects the sample size given in the OSPAR Guidelines)). Levels of mean FDI ≥ 2 can be reached if more fish are affected or if combinations of other toxicopathic lesions within this disease category occur.
It is emphasized that the proposed assessment criteria for contaminant-specific liver histopathology should be used in addition to the already developed FDI approach, possibly in a way illustrated in Figure 3. This figure depicts a situation where in the assessment period (a) a significant decrease in FDI level and (b) a significant downward trend in FDI level were recorded and (c) FDI levels recorded were finally below the assessment criterion. All three observations indicate an improvement of the disease status and, therefore, green smiley faces were assigned.

**Assessment criteria for macroscopic liver neoplasms (tumours) (Table 2):** Background levels are recommended on the basis of the (realistic) assumption that all macroscopic liver neoplasms (benign and malignant tumours) occurring in the target fish species are a direct effect of exposure to carcinogenic contaminants. Analogous to assessment criteria for contaminant-specific liver histopathology (see above), it is recommended to set the background level to values of mean FDI < 2 and the unacceptable level to values of mean FDI ≥ 2. A value of mean 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 mean FDI is > 2.

As for the contaminant-specific liver histopathology, it is emphasized that the proposed assessment criteria for macroscopic liver neoplasms should be used in addition to the already developed FDI approach, possibly in a way illustrated in Figure 3, with the appropriate FDI type on the vertical axis. The figure depicts a situation where in the assessment period (a) a significant decrease in mean FDI level and (b) a significant downward trend in mean FDI level were recorded and (c) FDI mean levels recorded were finally below the assessment criterion. All three observations indicate an improvement of the disease status and, therefore, green smiley faces were assigned.
Summary and Conclusions

The concept behind the Fish Disease Index (FDI) approach is presented and its major components and strengths are outlined.

Suggestions are made for additions to the original FDI approach in order to meet the OSPAR requirements as regards the establishment of assessment criteria for biological effects of contaminants. The new assessment procedure can be seen as a combination of the original FDI approach with the OSPAR assessment criteria approach.

For this purpose, the disease category ‘liver histopathology’ was divided into two new categories, namely ‘liver histopathology: non-specific’ and ‘liver histopathology: contaminant-specific’. By this, it is possible to separate between contaminant-specific and non-specific disease categories.
It is recommended to conduct the contaminant-specific assessment on the basis of the FDIs for the categories ‘liver histopathology: contaminant-specific’ and ‘macroscopic liver neoplasms’. Assessment criteria for these categories have been developed and are contained in Table 1. The assessment criteria for ‘liver histopathology: contaminant-specific’ have been derived based on expert judgement on disease severity and on a dataset from the Cefas Weymouth Laboratory (DEFRA 2009, in press) with some modifications.

The assessment criteria developed cannot only be used when long-term data are available, but also in one-off surveys, e.g. as part of exploratory monitoring, or in newly established fish disease monitoring programmes. In particular, the assessment of contaminant-specific liver histopathology and macroscopic liver neoplasms can be done with data of a single sample of at least the recommended size (50 fish).

It is further recommended to do a parallel but separate assessment of the non-specific categories ‘externally visible diseases’ and ‘liver histopathology: non-specific’ in order to provide additional information on significant changes in disease status that are useful for a more integrated assessment of contaminant effects. For this purpose, the original FDI strategy as developed by the ICES Working Group on Pathology and Diseases should be used for the time-being. The possibility of developing numerical assessment criteria will further be investigated. For one-off surveys or newly started monitoring programmes, the FDI approach needs to be modified and ways for this will be explored by the ICES Working Group on Pathology and Diseases of Marine Organisms (WGPDMO).

It is emphasized that the ICES fish disease database is the appropriate data source for the fish disease assessments. However, while the submitted data on externally visible disease are sufficient for a comprehensive assessment, there is still an apparent lack of data on liver histopathology and macroscopic liver neoplasms (repeatedly noted by WGPDMO and, in 2009, also by OSPAR ASMO). ICES Member Countries/OSPAR Contracting Parties are therefore encouraged to generate appropriate data according to established guidelines as part of their national marine monitoring programmes and to submit the data obtained (as well as the historical ones) to ICES.

It is further emphasized that fish diseases should not only be seen in the context of monitoring and assessing biological effects of contaminant, but have much wider applications and perspectives because they are appropriate indicators of general ecosystem health. As such, they are of relevance not only for the MSFD Task Group Descriptor 8 (contaminants) but certainly also for other MSFD Task Groups.

Because the ICES WGPDMO has developed and adopted the FDI approach in its original form, the modifications proposed in the present document should be reviewed by WGPDMO at its next meeting in February 2010 and further action needed should be taken. Possible changes required should subsequently be considered by OSPAR.

**Literature cited**


FEIST, S.W.; LANG, T., STENTIFORD, G.D.; KÖHLER, A. 2004. Biological effects of contaminants: Use of liver pathology of the European flatfish dab (*Limanda limanda* L.) and floun-
der (*Platichthys flesus* L.) for monitoring. ICES Techniques in Marine Environmental Sciences 38, 42 pp.


Table 2: Assessment criteria proposed for the assessment of contaminant-specific effects on fish health (criteria for the elevated response/above background and the significant/unacceptable effects levels are identical because no distinctions can be made, the colour ‘red’ should be used for their graphical representations in maps or similar illustrations; see explanations in the text).

<table>
<thead>
<tr>
<th>Disease category</th>
<th>Background</th>
<th>Elevated response/above background</th>
<th>Significant response/ unacceptable effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Externally visible diseases</td>
<td>Not applicable</td>
<td>Statistically significant increase in mean FDI level in the assessment period compared with a prior observation period</td>
<td>Statistically significant increase in mean FDI level in the assessment period compared with a prior observation period</td>
</tr>
<tr>
<td>(to be used as additional information for the assessment)</td>
<td>or</td>
<td>or</td>
<td></td>
</tr>
<tr>
<td>Liver histopathology: non-specific</td>
<td>Not applicable</td>
<td>Statistically significant upward trend in mean FDI level in the assessment period</td>
<td>Statistically significant upward trend in mean FDI level in the assessment period</td>
</tr>
<tr>
<td>(to be used as additional information for the assessment)</td>
<td>or</td>
<td>or</td>
<td></td>
</tr>
<tr>
<td>Liver histopathology: contaminant-specific</td>
<td>Mean FDI &lt; 2</td>
<td>Mean FDI ≥ 2</td>
<td>Mean FDI ≥ 2</td>
</tr>
<tr>
<td>Mean FDI &lt; 2</td>
<td>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 also be reached if more fish are affected or if combinations of other toxicopathic lesions occur.</td>
<td>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 also be reached if more fish are affected or if combinations of other toxicopathic lesions occur.</td>
<td></td>
</tr>
<tr>
<td>Mean FDI ≥ 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Macroscopic liver neoplasms</td>
<td>Mean FDI &lt; 2</td>
<td>Mean FDI ≥ 2</td>
<td>Mean FDI ≥ 2</td>
</tr>
<tr>
<td>Mean FDI &lt; 2</td>
<td>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 &gt; 2.</td>
<td>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 &gt; 2.</td>
<td></td>
</tr>
<tr>
<td>Mean FDI ≥ 2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Annex 5: Revised Chapter 6 for OSPAR Background Documents on Biological Effects Monitoring Techniques
Background Document on Biological Effects Monitoring Techniques

OSPAR Commission
2007
The Convention for the Protection of the Marine Environment of the North-East Atlantic (the “OSPAR Convention”) was opened for signature at the Ministerial Meeting of the former Oslo and Paris Commissions in Paris on 22 September 1992. The Convention entered into force on 25 March 1998. It has been ratified by Belgium, Denmark, Finland, France, Germany, Iceland, Ireland, Luxembourg, Netherlands, Norway, Portugal, Sweden, Switzerland and the United Kingdom and approved by the European Community and Spain.


© OSPAR Commission, 2007. Permission may be granted by the publishers for the report to be wholly or partly reproduced in publications provided that the source of the extract is clearly indicated.

© Commission OSPAR, 2007. La reproduction de tout ou partie de ce rapport dans une publication peut être autorisée par l'Editeur, sous réserve que l'origine de l'extrait soit clairement mentionnée.
EXECUTIVE SUMMARY ........................................................................................................... 4
RÉCAPITULATIF ......................................................................................................................... 4
CHAPTER 1 ASSESSMENT CRITERIA FOR HEPATIC METALLOTHIONEIN IN FISH ..... 6
CHAPTER 2 ASSESSMENT CRITERIA FOR D-AMINOLEVULINIC ACID
DEHYDRATASE (ALA-D) MEASURED IN FISH BLOOD ....................................................... 10
CHAPTER 3 CYTOCHROME P4501A ACTIVITY (EROD) ...................................................... 14
CHAPTER 4 LYSOSONAL STABILITY AS A GLOBAL HEALTH STATUS INDICATOR
IN BIOMONITORING .............................................................................................................. 20
CHAPTER 5 PAH METABOLITES IN BILE ............................................................................ 27
CHAPTER 6 EXTERNALLY VISIBLE FISH DISEASES, MACROSCOPIC LIVER
NEOPLASMS AND LIVER HISTOPATHOLOGY .................................................................... 37
CHAPTER 7 (WATER) BIOASSAYS INCLUDING PROPOSALS FOR ASSESSMENT
TOOLS AND CRITERIA ......................................................................................................... 52
CHAPTER 8 SEDIMENT SEAWATER ELUTRIATE AND PORE-WATER BIOASSAYS
WITH EARLY DEVELOPMENTAL STAGES OF MARINE
INVERTEBRATES ................................................................................................................... 69
CHAPTER 9 SEDIMENT SEAWATER ELUTRIATE AND PORE-WATER BIOASSAYS
WITH COPEPODS (Tisbe, Acartia), MYSIDS, AND DECAPOD
LARVAE (Palaemon) ............................................................................................................ 75
CHAPTER 10 WHOLE SEDIMENT BIOASSAYS WITH AMPHIPODS
(Corophium sp) ................................................................................................................... 78
CHAPTER 11 DNA ADDUCTS .............................................................................................. 81
CHAPTER 12 REPRODUCTIVE SUCCESS IN EELPOUT (Zoarces viviparus) .......... 88
APPENDIX 1 COMMENTS ON THE REPRODUCTIVE SUCESS TEXT FOLLOWING
THE SECOND ICES/OSPAR WORKSHOP ON INTEGRATED
MONITORING OF CONTAMINANTS AND THEIR EFFECTS IN COASTAL
AND OPEN-SEA AREAS (WKIMON II), ............................................................................. 94
CHAPTER 13 FISH VITELLOGENIN AS A BIOMARKER OF EXPOSURE TO
XENOESTROGENS .............................................................................................................. 97
CHAPTER 14 SCOPE FOR GROWTH .................................................................................... 120
Executive Summary

1. The OSPAR Co-ordinated Environmental Monitoring Programme (CEMP – OSPAR Agreement 2007-01)) is that part of the monitoring within the OSPAR Joint Assessment and Monitoring Programme where the national contributions overlap and are co-ordinated. The aim of the CEMP is to deliver comparable data from across the OSPAR maritime area, which can be used in assessments to address the specific questions raised in the JAMP.

2. CEMP Appendix 7 covers techniques for monitoring of PAH and metal specific biological effects as set out in the JAMP Guidelines for Contaminant Specific Biological Effects Monitoring (OSPAR agreement 2003-10). CEMP Appendix 9 covers techniques for monitoring general biological effects as set out in the JAMP Guidelines for General Biological Effects Monitoring (OSPAR agreement 2003-10).

3. Biological Effect techniques are included under the CEMP to address the following JAMP questions on hazardous substances:
   a. 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 man made substances?
   b. 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?

4. In order to address these questions the techniques can be deployed:
   a. for investigative monitoring or screening;
   b. as an exploratory tool for identifying both the effects of contaminants not included in the monitoring programme and combined effects of contaminants;
   c. as part of an integrated chemical and biological effects monitoring programme;

5. The OSPAR Environmental Assessment and Monitoring Committee (ASMO) recognises that there is still a need for further work on relating concentrations of hazardous substances in the marine environment to biological effects measurements. Furthermore the deployment of biological effects techniques under the CEMP is currently restricted by a lack of assessment tools and in some cases established quality assurance schemes, which are essential for full co-ordination of monitoring between Contracting Parties.

6. This document presents a series of reviews of biological effects monitoring techniques which are either included in the OSPAR Co-ordinated Environmental Monitoring Programme (CEMP) or being considered for inclusion. The reviews were prepared by experts from several OSPAR Contracting Parties to provide background information for use in a review of CEMP monitoring and in the development of proposals for assessment criteria for biological effects monitoring. The reviews cover inter alia the following issues:
   a. an assessment of the applicability of the biological effects techniques across the OSPAR maritime area;
   b. a review of the environmental variables that influence the biological effect;
   c. an assessment of the thresholds when the response of a biological effect technique can be considered of concern and/or require a response;
   d. proposals for assessment tools;
   e. status of quality assurance techniques.

7. The reviews have been developed in collaboration with ICES and examined at two joint ICES/OSPAR workshops on integrated chemical and biological effects monitoring (WKIMON II and WKIMON III) in 2006 and 2007. This background document is intended to be a dynamic document, which is updated in the light of further progress made on assessment criteria for biological effects techniques by both OSPAR and ICES.

Récapitulatif

1. Le Programme coordonné de surveillance continue OSPAR (CEMP – Accord OSPAR 2006-01) correspond à la partie de la surveillance dans le cadre du Programme conjoint d’évaluation et de surveillance OSPAR (JAMP) comportant des contributions nationales qui se chevauchent et sont coordonnées. Le CEMP a pour objectif de produire des données comparables provenant de toute la zone
maritime OSPAR, qui peuvent être utilisées dans les évaluations qui permettent de traiter les questions spécifiques soulevées dans le JAMP.


3. Les techniques utilisées pour les effets biologiques font partie du CEMP pour permettre de traiter les questions suivantes du JAMP sur les substances dangereuses:
   a. quelles sont les teneurs dans le milieu marin ainsi que les effets des substances inscrites sur la liste OSPAR des produits chimiques devant faire l’objet de mesures prioritaires («produits chimiques prioritaires»)? Se situent-elles aux niveaux ambients dans le cas des substances présentes à l’état naturel ou approchent-elles ces niveaux, ou sont-elles proches de zéro dans le cas des substances de synthèse?
   b. y a-t-il des problèmes émergents dus à la présence de substances dangereuses dans le milieu marin? En particulier, l’exposition à des substances dangereuses donne-t-elle lieu à des réactions biologiques imprévues/intolérables, ou à des réactions d’une ampleur imprévue/intolérable?

4. On peut déployer les techniques afin de traiter ces questions:
   a. pour la surveillance ou le filtrage investigatifs;
   b. en tant qu’outil exploratoire pour la détermination des effets des contaminants qui ne figurent pas dans le programme de surveillance et les effets combinés des contaminants;
   c. dans le cadre d’un programme de surveillance intégrée des effets chimiques et biologiques;

5. Le Comité OSPAR évaluation et surveillance de l’environnement (ASMO) reconnaît qu’il est encore nécessaire d’entreprendre des travaux supplémentaires pour établir un lien entre les teneurs des substances dangereuses dans le milieu marin et les analyses des effets biologiques. Le déploiement des techniques utilisées pour les effets biologiques dans le cadre du CEMP est de plus limité actuellement par le manque d’outils d’évaluation et, dans certains cas, de systèmes d’assurance de qualité bien établis, qui sont essentiels pour parvenir à une coordination complète de la surveillance entre les Parties contractantes.

6. Ce document présente une série d’études de techniques de surveillance des effets biologiques qui figurent dans le CEMP ou dont on envisage l’inclusion dans le CEMP. Les études ont été préparées par des experts provenant de plusieurs Parties contractantes afin de fournir des informations contextuelles qui peuvent être utilisées lors de l’examen de la surveillance du CEMP et de l’élaboration de propositions pour des critères d’évaluation pour la surveillance des effets biologiques. Les études couvrent, entre autres, les questions suivantes:
   a. une évaluation de l’applicabilité des techniques pour les effets biologiques dans l’ensemble de la zone maritime OSPAR;
   b. un examen des variables environnementales qui influencent les effets biologiques;
   c. une évaluation des valeurs seuils lorsque la réponse d’une technique pour les effets biologiques peut être considérée préoccupante et ou demande une réponse;
   d. des propositions d’outils d’évaluation;
   e. l’état des techniques d’assurance de qualité.

7. Les études ont été réalisées en collaboration avec le CIEM et examinées lors de deux ateliers conjoints CIEM/OSPAR sur la surveillance intégrée des effets chimiques et biologiques (WKIMON II et WKIMON III) en 2006 et 2007. Ce document de fond devra être un document dynamique qui est actualisé à la lumière des nouveaux progrès réalisés sur les critères d’évaluation des techniques pour les effets biologiques et ce aussi bien par OSPAR que par le CIEM.
Chapter 1  Assessment criteria for hepatic metallothionein in fish

Background

1. This review of hepatic metallothionein in fish was prepared by Norway as the basis for development of advice on the role of the biological effects technique metallothionein (MT) within the CEMP and consideration of what tools should be used for its assessment.

2. The low-molecular-weight protein MT is present in most cells in vertebrates and some invertebrates. It is a peculiar protein in that around a third of the protein is made up by the sulphydryl-containing amino acid cysteine, which contributes directly to the binding of 7-12 metal atoms in each MT molecule (Klaassen et al. 1999). It was early indicated that tissue concentrations of MT would increase following metal exposure (especially to cadmium) and it was then suggested that metal detoxification was the major function of the protein. Later work has shown that MT is involved in the regulation of the intracellular availability of zinc (Zn) and/or copper (Cu) (see e.g. Olsson 1993), which to some implied a possible role in total gene-regulation (through regulating Zn availability). Metallothionein is also induced in cells exposed to free radicals and do indeed appear to protect against damaging effects from e.g. ionising radiation or radical-generating agents (Kagi & Schaffer 1988). The production of MT knock-out mice definitely established that the protein is not essential to survival, at least not in mammals, but will provide protection from toxic metals such as cadmium (Masters et al. 1994).

3. In addition to Cu and Zn, MT binds non-essential metals such as cadmium (Cd), mercury (Hg) and silver (Ag). The synthesis of the protein is induced by elevated intracellular concentration of the above metals. Metallothionein induction is a response to elevated intracellular metal concentrations and the protein has for that reason been suggested as a biomarker for environmental metal contamination (Engel & Roesijadi 1987, Haux & Forlin 1989, George & Olsson 1994).

4. The methods most commonly used to quantify MT in tissues exploit some of the characteristics of this protein: its metal content, the number of sulphhydril groups, its size and/or its heat-stability. In addition, various immunochemical assays using MT-specific antisera have been established in the past 10-15 years (Roesijadi et al. 1988, Hogstrand et al. 1989, Hylland 1999). More recently, metallothionein gene expression (mRNA) has been quantified in fish species using rtPCR (Cheung et al. 2004, Tom et al. 2004). Similar methods have also been developed for mussels (Dondero et al. 2005). Quantification of the protein, and not mRNA, should theoretically be most useful for monitoring purposes for two reasons. Firstly, the physiological and toxicological significance of a response will depend upon the concentration of MT, and not MT mRNA, present in the cell. Secondly, the half-life of MT mRNA is much shorter than the half-life of MT and the time-span of a measurable response thus shorter. On the other hand, recent studies do suggest that expression studies (rtPCR) may also be relevant for monitoring (see e.g. George et al. 2004).

5. There is limited knowledge of how metallothionein in tissues of different fish or invertebrate species respond to metal stress. For the marine species relevant in the North Sea area there appear to be surprisingly little difference in baseline hepatic MT levels. Both design and analytical protocols differ between studies and it is rarely feasible to compare species directly.

Confounding factors

6. There are species differences in how endogenous and exogenous factors other than metal stress affect MT in fish tissues, as indicated in (Ruus et al. 2003). In a comprehensive multiple regression analysis for Atlantic cod (including more than 600 individual fish sampled at 8 sites over 5 years), both organic contaminants and essential metals came up as highly significant, in addition to size and a site-related factor that was not reflected in tissue-levels of metals or organochlorine contaminants. Similar regression models for flatfish included essential metals, but also tissue concentrations of organic contaminants and size. All samples were collected at the same time of the year (September-October) and water temperatures were surprisingly similar (within 2-3°C) at the different locations sampled through a five-year period.

7. The relevance of various factors that affect hepatic MT in flounder was assessed by (Nissen-Lie 1997) in a seasonal study over 1 year. As for cod, hepatic concentrations of the essential metals copper and zinc appeared to drive (or be driven by) hepatic concentrations of MT. Metallothionein in the liver varied through the year in both male and female fish, apparently related to the demand for Zn in gonad development and
maturation. As for cod, results from Lacorn et al. (2001) on dab similar identify season as an important factor for hepatic MT, but also suggest that water temperature is relevant. The latter can not easily be distinguished from season, so no clear conclusions can be drawn. The main conclusions in (Rotchell et al. 2001) were that age and season are important for variations in hepatic MT in flounder.

8. The results in Hylland et al. (1996) indicated that starvation had less influence on MT than on other biomarkers in flounder fed cadmium-contaminated polychaetes. There was also surprisingly little effect of co-exposure of cadmium with BaP or PCB in that study. Corresponding results were found in the study by (Sandvik et al. 1997), but fish treated with BaP prior to Cd appeared to have a lower capacity to synthesise MT than fish who only received Cd or PCB prior to Cd.

Ecological relevance

9. As mentioned above, metallothionein appears to be a household protein in many organisms (at least in mammals) with a range of functions, but none really essential. It is not clear whether this is the case for fish. Fish have higher levels of metals in the liver than most mammals and some fish species have phenomenally high concentrations of metals under normal conditions (Hogstrand et al. 1996, Hogstrand & Haux 1996, Thompson et al. 2003). Data for flounder and cod from unpolluted to moderately polluted areas support a role for metallothionein in trace metal (zinc and possibly copper) metabolism in marine fish, but non-essential metals such as cadmium do not appear to have a strong effect (in contrast to the situation in some freshwater ecosystems). Metallothionein is however one of the genes most clearly expressed in rTaqPCR studies and the protein appears to have a function in a general stress response, as much or more related to organic contaminants as to metals. There are no clear links from metallothionein induction to individual health or reproduction. Some freshwater studies have shown reduced growth at metal-levels which also induce MT (e.g. Roch & McCarter 1984), but such levels are rarely if ever encountered in marine ecosystems.

Quality assurance

10. Metallothionein was included in BEQUALM and intercalibration between 14 laboratories run on liver samples from cod. The results clearly showed some of the inherent problems of the analytical procedures. Metal-saturation assays and polarographic analyses produced remarkably similar results, whereas immunochemical assays (ELISA) and the colorimetric assay produced results that deviated from the overall median.

Background responses and assessment criteria

11. As long as a species for which there is background knowledge is used, sampling is performed during the appropriate time period and individuals selected on the basis of sex and size, absolute assessment criteria may be used for hepatic metallothionein. The main dataset used in this context has been derived from the JAMP monitoring programme (Ruu så et al., 2003), but supplemented with other studies, e.g. Nissen-Lie (1997) for flounder. Baseline values for Atlantic cod are in the range 6.5-16 µg/mg cytosolic protein, for flounder 8.4-15 µg/mg cytosolic protein, for dab 7.2-13 µg/mg cytosolic protein and for plaice 6.5-14 µg/mg cytosolic protein (Ruu så et al. 2003). As will be evident, the upper 90 percentile for baseline values for all four species lies in the range 13-16 µg MT/mg cytosolic protein.

12. Hepatic MT significantly above the values indicated in each species will indicate an external stressor that affects this biomarker. To maintain sufficient power in this analysis (ability to detect effects) a minimum of 25 fish should be sampled and analysed. The criteria are the following: for cod 16, for flounder 15, for dab 13 and for plaice 14 µg/mg cytosolic protein.

13. Hepatic metallothionein in fish can be measured in a reproducible manner by a range of laboratories and there exist sufficient background information for results to be assessed. Measurement of metallothionein protein (the current method) in Atlantic cod does not appear to be a useful marker for metal stress, but reflects general trace metal (Cu, Zn) metabolism by the liver. Hepatic metallothionein has sufficient quality assurance “backing” to be recommended for inclusion in CEMP. It does not appear to be useful for monitoring environmental metal stress with Atlantic cod, but is possibly more so for flatfish species, e.g. dab.
References


Hylland K (1999) Biological effects of contaminants: quantification of metallothionein in fish. ICES Techniques in Marine Environmental Science:1-18


Chapter 2  Assessment criteria for d-aminolevulinic acid dehydratase (ALA-D) measured in fish blood

Background

1. This review of d-aminolevulinic acid dehydratase (ALA-D) measured in fish blood was prepared by Norway as the basis for development of advice on the role of the biological effects technique ALA-D within the CEMP and consideration of what tools should be used for its assessment.

2. Many metals are essential to all organisms, e.g. copper, zinc, iron, manganese and molybdenum, but there are also metals for which no biological function is known. Non-essential metals include mercury, cadmium, lead and silver. They generally have higher toxicity than the essential metals. Uptake, storage and excretion are also less well controlled for the non-essential than for essential metals.

3. One of the most important toxic mechanisms of non-essential metals is interaction with and inhibition of enzymes, especially enzymes with metal co-factors. ALA-D (E.C. 4.2.1.24) is such an enzyme, which has zinc as a co-factor. This enzyme catalyses a step in the synthesis pathway for heme and is found in bacteria, plants and invertebrates as well as in vertebrates. Heme is incorporated in macromolecules such as haemoglobin and cytochromes. In mammals and birds, inhibition of ALA-D may lead to anaemia since it is one of the rate-limiting enzymes in heme (and hence haemoglobin) synthesis. The reason for ecotoxicological interest in ALA-D is its inhibition by lead, even at very low exposure levels (Hodson et al. 1984, Haux & Förlin 1988).

4. ALA-D has been used to investigate effects of metals (mainly lead) in rainbow trout (Hodson 1976, Hodson et al. 1977, Hodson et al. 1978, Addison et al. 1990, Sordyl & Osterland 1990), longear sunfish (Dwyer et al. 1988), perch (Haux, et al., 1985), salmon (Johansson-Sjöbeck & Larsson 1979), flounder (Johansson-Sjöbeck & Larsson 1978, Holth 2004), carp (Nakagawa et al. 1995, Nakagawa et al. 1998), various catostomid species (Schmitt et al. 1984), Atlantic cod (Ruus et al. 2003, Holth 2004) and grey mullet (Krajnovic-Ozretic & Ozretic 1980). As will be apparent from the above, the method has predominantly been used with freshwater species.

5. ALA-D activity can be determined in liver, kidney, spleen and red blood cells in teleosts (see (Hodson et al. 1984). The sensitivity, timing and specificity of ALA-D inhibition in response to lead exposure have been reviewed (Hodson et al. 1984). For rainbow trout, red blood cell ALA-D activity may inhibited following exposure to less than 5 µg/L of lead through water (Haux et al. 1986) and blood lead concentration of 300 µg/L (Hodson et al. 1982), although other studies indicate a lower sensitivity (Burden et al. 1998). In rainbow trout, ALA-D inhibition is evident 1-2 days after exposure to lead and there is a clear dose-response relationship (Hodson et al. 1982). The half-life of the response will depend on the amount of lead accumulated in other tissues (in equilibrium with blood), but will generally be of the order of months. No dose-response studies have been done with marine species. It is however likely that responses will similar to those observed for rainbow trout. It is however important to take into account that lead is generally less bioavailable in marine systems than in freshwater.

6. ALA-D activity in red blood cells of Atlantic cod (Gadus morhua), flounder (Platichthys flesus) and dab (Limanda limanda) has been used in the Norwegian national monitoring programme for 5 years (1997-2001; Ruus et al. 2003). In that programme, ALA-D was generally found to be inhibited in blood of fish from areas with known urban impact and in one area with a known metal impact (Ruus et al. 2003). In fact, results for ALA-D in that study indicated that there could be lead toxicity in individual cod with hepatic lead levels below the detection limit for routine analyses. Despite its specificity, ALA-D has not been widely used in monitoring. It has been included in OSPAR’s list of recommended methods to monitor biological effects of metals (JAMP, 1998). In addition to the above, Holth (2004) investigated ALA-D in Atlantic cod and flounder in southern Norway. He found clear inhibition of the enzyme in both species sampled in the inner Oslofjord compared to individuals collected from the outer Oslofjord. Individuals from both species collected in the outer Oslofjord had values in the range of baseline values (see below).

Confounding factors

7. Lead is generally found in the environment as the inorganic form, but the alkylated form may be present. Alkyllead does not affect ALA-D strongly and there may thus be a situation where there is little inhibition, but high blood lead levels (Hodson et al. 1984). Data on the influence of metals other than lead are
contradictory (Jackim 1973, Dwyer et al. 1988, Sordyl & Osterland 1990). There is some evidence that very high levels of Cd may inhibit the enzyme in flounder (Johansson-Sjöbeck & Larsson 1978). Zinc may ameliorate inhibition by lead, and appears to affect fish red blood cell ALA-D in vivo (Schmitt et al. 1984, Schmitt et al. 1993, Schmitt et al. 2005). The influence of zinc has however been found to be weak and variable in rainbow trout (Hodson et al. 1984).

**Quality control**

8. There has been a tradition to express results for ALA-D in terms of blood cell volume (e.g. hematocrit), but more recent recommendations suggest to use protein (Hylland 2004). It is not possible to compare different studies because values have not been reported in the same format and it is not feasible to estimate blood protein from blood cell volume (or vice versa).

9. There is no commercially available standard reference material for fish ALA-D. Bovine ALA-D is available commercially and may be used to test the assay. The pH of the assay needs to be established for each fish species (and will differ from the pH optimum of bovine ALA-D). Suitable internal reference materials (blood supernatants from relevant fish species) should be prepared by each laboratory undertaking ALA-D analyses. This material should be aliquotted and stored at -80°C for use with each batch of analyses. Control charts should be prepared using internal reference materials.

10. ALA-D was one of the methods included in the EU-funded project BEQUALM (Biological effects quality assurance in marine monitoring programmes; project PL97-3587). The low number of laboratories that participated in the exercise (5) precludes general conclusions on the use of the method. The exercise did show that homogenous test samples for intercalibration purposes can be prepared from red blood cells diluted in buffer.

**Baseline levels and assessment criteria**

11. From the analyses done under Norwegian JAMP, typical values from reference areas are: Atlantic cod 15-21 ng PBG/min/mg protein, plaice 13-21 ng PBG/min/mg protein, flounder 13-21 ng PBG/min/mg protein and dab 10-20 ng PBG/min/mg protein. As will be evident, normal values fall within similar ranges for the species investigated.

12. As indicated above, the normal levels of ALA-D appear to be similar for many species and lie within the range 10-20 ng PBG/min/mg protein. Delta-aminolevulinic acid dehydratase is one biological effects method for which global values can be set. Red blood cell ALA-D significantly below the values indicated in each species will indicate an external stressor that affects this biomarker. To maintain sufficient power in this analysis (ability to detect effects) a minimum of 25 fish should be sampled and analysed. Fish can be considered affected if the red blood cell ALA-D in the population studied has a value significantly below a certain criteria, The proposed criteria are:

<table>
<thead>
<tr>
<th>Fish species</th>
<th>ng PBG/min/mg protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cod</td>
<td>15</td>
</tr>
<tr>
<td>Flounder and plaiice</td>
<td>13</td>
</tr>
<tr>
<td>Dab</td>
<td>10</td>
</tr>
</tbody>
</table>

13. Red blood cell ALA-D can be determined in a reproducible manner by different laboratories, but knowledge of the technique is not widespread. It has been subject to international intercalibration, but with a limited number of participants. The method has been found to be very sensitive to lead exposure in marine ecosystems, which supports its inclusion in the CEMP. Further international intercalibration activities should be pursued to ascertain comparability of results.
References

Addison RF, Fitzpatrick D, Renton KW (1990) Distribution of -aminolevulinic acid synthetase and d-aminolevulinic acid dehydratase in liver and kidney of rainbow trout (Salmo gairdneri). Comparative Biochemistry and Physiology 95B:317-319


Hodson PV, Blunt BR, Spry DJ (1978) Chronic toxicity of water-borne and dietary lead to rainbow trout (Salmo gairdneri) in Lake Ontario water. Water Research 12:869-878


Johansson-Sjöbeck M-L, Larsson Å (1979) Effects of inorganic lead on delta-aminolevulinic acid dehydratase activity and hematomatological variables in the rainbow trout, Salmo gairdnerii. Archives of Environmental Contamination and Toxicology 8:419-431


Larsson Å, Haux C, Sjöbeck M-L (1985) Fish Physiology and Metal Pollution: Results and Experiences from Laboratory and Field Studies. Ecotoxicology and Environmental Safety 9:250-281.


Nakagawa H, Tajima T, Sato T, Kuboyama M (1998) A field study of water lead pollution in fresh water areas of northern Kyushu, based on 5-aminolevulinic acid dehydratase activity and lead concentration in the


Chapter 3  
Cytochrome P4501A activity (EROD)  
Compiled by Anders Ruus & Ketil Hylland (NIVA)

Introduction

1. EROD (7-Ethoxyresorufin-O-deethylase) is a cytochrome P450 catalysed reaction with ethoxyresorufin as the substrate (Burke and Mayer 1974; Stagg & McIntosh, 1998). Cytochrome P450 1A catalyse the deethylation of 7-ethoxyresorufin to resorufin.

2. The cytochrome P450 system is a superfamily of enzymes with several hundred forms comprising more than 250 different families, further divided into subfamilies. The CYP system is highly diversified and is found in bacteria, plants, lower eukaryotes and in animals. Members of the P450 subfamily CYP1A are particularly important in the metabolism of many pollutants. In the case of planar molecules, such as polycyclic aromatic hydrocarbons (PAHs) isoenzymes of CYP1A are responsible for the insertion of oxygen into the molecule, which is the first oxidative step in the biotransformation process (termed ‘phase I’; Williams, 1974).

3. In addition to being substrates for biotransformation, planar compounds, such as PAHs, can also interact with cytochrome P450 1A as inducers, by binding to the cytosolic Ah (aryl hydrocarbon)-receptor. EROD is a tool used 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). It has later gained widespread use (see e.g. Förlin et al., 1990; George et al., 1995; Goksøyr et al., 1991; Whyte et al., 2000).

Dose-response

4. 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 10-fold 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 categorised 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

5. Several factors have been shown to affect hepatic EROD, both endogenous and exogenous. The most important endogenous factors for most fish species are developmental stage (juvenile-mature), gender, reproductive status and age, 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 e.g. rainbow trout (Förlin & Haux 1990), flounder (Hylland et al., 1998), 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).

6. Several species have baseline EROD activities within the same order of magnitude among different studies/measurements and also show greater than 10-fold EROD induction after contaminant exposure (Whyte et al. 2000). These are, however, mostly freshwater species.

Developmental stage of the fish is very important. The main age-related factors are time of exposure/accumulation, food selection and reproductive stage.

7. The mechanism for CYP1A suppression in spawning females is related to 17β-estradiol (E2) (or xenoestrogen) levels. The hormone controls the induction of vitellogenin (VTG; egg yolk protein) production during gonadal recrudescence. Some of the inter-gender differences during spawning can be attributed to increased levels of CYP isoenzymes in males rather than suppression of levels in females.

8. Dietary factors can be 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

9. Baseline levels of EROD in four marine species have been estimated from results derived from the Norwegian monitoring programme (Ruus et al., 2003). The baseline value for Atlantic cod has been suggested to be 9-95 pmol/min/mg protein including fish from the Norwegian west coast and if only fish from the Barents sea had been included, the values would have been 9-25 pmol/min/mg protein. For flounder, baseline values are in the range 10-43 pmol/min/mg protein, for dab 123-529 pmo/min/mg protein and for plaice 33-146 pmol/min/mg protein. The fish were sampled from reference locations (i.e. no known local sources of contamination) in the autumn, the data includes males and females and the water temperature at the sampling locations was 9-11°C.

Assessment criteria

10. As many factors are known to influence EROD and it is not feasible to correct for all in the design, it is advisable always to include an appropriate reference group in studies that include EROD as an endpoint. Experience suggests that an EROD value in most marine species above twice the upper limit of baseline values indicate an ecosystem influenced by planar organic contaminants.

Quality assurance

11. Cytochrome P4501A is possibly the most widely used biomarker. There have been two 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

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

Relevant literature (marine and freshwater fish)


### Table 1. Dose-response, background response and sensitivity in experimental studies with gadoid fish

<table>
<thead>
<tr>
<th>species</th>
<th>substance(s)</th>
<th>lowest-highest concs</th>
<th>exposure time</th>
<th>baseline/control (level/activity)</th>
<th>induction (fold)</th>
<th>reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polar cod Boreogadus saida</td>
<td>Crude oil (Oseberg C)</td>
<td>~200 mg/kg (i.p. inj.)</td>
<td>10 and 21 d post inj.</td>
<td>~30 pmol/min/mg</td>
<td>~8 and ~2.5 (245 and 80 pmol/min/mg)</td>
<td>(George et al. 1995)</td>
</tr>
<tr>
<td>juvenile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polar cod Boreogadus saida</td>
<td>Crude oil (Oseberg C)</td>
<td>~200 mg/kg (oral)</td>
<td>21 d post exposure</td>
<td>28 pmol/min/mg ± 6 (n=12)</td>
<td>~5</td>
<td>(George et al. 1995)</td>
</tr>
<tr>
<td>male</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polar cod Boreogadus saida</td>
<td>Crude oil (Oseberg C)</td>
<td>~200 mg/kg (oral)</td>
<td>21 d post exposure</td>
<td>8 pmol/min/mg ± 2 (n=14)</td>
<td>~5</td>
<td>(George et al. 1995)</td>
</tr>
<tr>
<td>female</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polar cod Boreogadus saida</td>
<td>5β-naphthoflavone</td>
<td>50 mg/kg (i.p. inj.)</td>
<td>21 d post inj.</td>
<td>~30 pmol/min/mg</td>
<td>~12.5 (380 pmol/min/mg)</td>
<td>(George et al. 1995)</td>
</tr>
<tr>
<td>juvenile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod, Gadus morhua</td>
<td>2,3,7,8-TCDD</td>
<td>0.008 mg/kg oral dose twice, d 0 and d 4</td>
<td>9 and 17 d post exposure</td>
<td>55.4 (d 9) and 91.4 (d 17) pmol/min/mg</td>
<td>~4 and ~3 (230 and 277 pmol/min/mg)</td>
<td>(Hektoen et al. 1994)</td>
</tr>
<tr>
<td>juvenile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod, Gadus morhua</td>
<td>PCB-105</td>
<td>10 mg/kg oral dose twice, d 0 and d 4</td>
<td>measure at d 9 and d 17</td>
<td>55.4 (d 9) and 91.4 (d 17) pmol/min/mg</td>
<td>1.5 and 1.2</td>
<td>(Bernhoft et al. 1994)</td>
</tr>
<tr>
<td>juvenile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod, Gadus morhua</td>
<td>5β-naphthoflavone</td>
<td>100 mg/kg (i.p. inj. at d 0 and d 4)</td>
<td>measure at d 7</td>
<td>84 pmol/min/mg ± 8 (n=5)</td>
<td>~13 (1074 ± 340 pmol/min/mg)</td>
<td>(Goksoyr et al. 1987)</td>
</tr>
<tr>
<td>juvenile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod, Gadus morhua</td>
<td>5β-naphthoflavone</td>
<td>100 mg/kg (2 i.p. inj.)</td>
<td>measure 3-4 d after last injection</td>
<td>40 pmol/min/mg</td>
<td>~72 (2870 pmol/min/mg)</td>
<td>(Goksoyr et al. 1991)</td>
</tr>
<tr>
<td>species</td>
<td>substance(s)</td>
<td>lowest-highest concs</td>
<td>exposure time</td>
<td>baseline/control (level/activity)</td>
<td>induction (fold)</td>
<td>reference</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-----------------------------------</td>
<td>----------------------</td>
<td>---------------</td>
<td>-----------------------------------</td>
<td>-----------------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>Cod, Gadus morhua</td>
<td>Crude oil (North Sea)</td>
<td>0.06 – 1 ppm</td>
<td>30 days</td>
<td>∼2 pmol/min/mg</td>
<td>∼ 2-5.5 (∼ 4 – 11 pmol/min/mg)</td>
<td>(Aas et al. 2000)</td>
</tr>
<tr>
<td>juvenile</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rockling, Ciliata mustella</td>
<td>Crude oil (Gullfaks; M.V. Braer spill, Shetland)</td>
<td>85000 tons spill 129 ± 38 ng/g dry wt. of PAHs (selected 2- and 3-ring) detected in muscle.</td>
<td>3 months after spill</td>
<td>∼160 pmol/min/mg ± 50</td>
<td>∼9 (1480 pmol/min/mg)</td>
<td>(George et al. 1995)</td>
</tr>
<tr>
<td>Roundnose grenadier, Coryphaenoide s rupestris</td>
<td>i.a. PAHs and PCBs</td>
<td>260 ± 20 (Male) ∼170 (Female) pmol/min/mg</td>
<td></td>
<td></td>
<td>∼2 (530 ± 70 (male) and ∼350 (female) pmol/min/mg)</td>
<td>(Lindesjoo et al. 1996)</td>
</tr>
<tr>
<td>Hake, Urophycis spp.</td>
<td>Pollution (PAH) from oil platforms (Gulf of Mexico) &lt;100 m from platforms</td>
<td>10.9 ± 6.4 and 11.7 ± 10.5 pmol/min/mg (&gt;3000 m from platforms)</td>
<td></td>
<td></td>
<td>&lt;1 (10.6 ± 3.8 and 10.5 ± 7.1 pmol/min/mg)</td>
<td>(McDonald et al. 1996)</td>
</tr>
</tbody>
</table>

Table 2. Dose-response, background response and sensitivity in field studies with gadoid fish
Chapter 4  Lysosomal stability as a global health status indicator in biomonitoring

Background

1. 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, b; 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., 2006a, b), 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.

2. 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).

3. This biomarker can also be used prognostically to predict liver damage and tumour progression in the liver of various fish species (Broeg et al., 1999 a, b; Köhler et al., 2002; Köhler, 2004), and hepatopancreatic degeneration in molluscs (e.g., blue and green mussels, freshwater 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).

4. 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 destabilisation in mussels (Viarengo et al., 1981, 1985a, b), and if oxyradicals are generated then lipofuscinosis can also occur (1985b).

5. 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).

6. 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).
OSPAR Commission, 2007:  
Background Document on Biological Effects Monitoring Techniques

7. 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.

8. Health status space is analogous 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 multi-dimensional 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

9. 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 anyway 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

10. 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

11. 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.

12. The standards used in this intercalibration involved digestive glands from marine mussels prepared at the University of Genova / 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

13. 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).

14. 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 >20 minutes; stressed but compensating if <20 but >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 >120 minutes; stressed but compensating if <120 but >50 minutes and severely stressed and probably exhibiting pathology if <50 minutes (Moore et al., 2006a).

References


Chapter 5 PAH metabolites in bile
Compiled by Lars-Petter Myhre (IRIS-Akvamiljø) & Ketil Hylland (NIVA)

Background

1. 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.

2. 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).

3. 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).

4. 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 for 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)

5. 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.

6. The method requires that bile is available in the gall bladder. Since 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.

7. 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 normalisation is not standardised 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

8. 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.

Relevance of other factors

9. 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 in 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.

Background responses

10. Baseline levels of PAH metabolites have been established for many of the species relevant for monitoring in Norwegian coastal and offshore waters. From Ruus et al. (2003) values for the relevant species are: (all values standardised 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

11. It is possible to establish global criteria for individual PAH metabolites. Baseline data for individual species may be used to test against to determine whether fish have been exposed to PAHs. As mentioned above, 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

12. 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 QUASIMEME1. 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).

Acknowledgement

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

1 QUASIMEME – organisation that offers quality assurance for chemical endpoints; http://www.quasimeme.org
References


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

<table>
<thead>
<tr>
<th>Species</th>
<th>Substance (lab/field)</th>
<th>Test concentrations/area</th>
<th>Exposure time</th>
<th>Metabolite</th>
<th>Baseline</th>
<th>control reference or exposed.control</th>
<th>reference/comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Feral fish Barents Sea</td>
<td>Baseline</td>
<td></td>
<td>Naph type</td>
<td>5.3 ug/ml</td>
<td></td>
<td>Klungsøyr et al. 2003</td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Feral fish Egersund</td>
<td>Baseline non polluted area</td>
<td></td>
<td>Pyren type</td>
<td>0.8 ug/ml</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Feral fish Sleipner</td>
<td>Baseline polluted area?</td>
<td></td>
<td>BaP type</td>
<td>0.4 ug/ml</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Feral fish Statfjord</td>
<td>Baseline polluted area?</td>
<td></td>
<td>Naph type</td>
<td>6.1 ug/ml</td>
<td></td>
<td>Klungsøyr et al. 2003</td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Feral fish Frøy, ceased installation 10 000 m (ref) 2000 m - 200 m</td>
<td>Baseline polluted area?</td>
<td></td>
<td>Pyren type</td>
<td>1.0 ug/ml</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Feral fish Barents sea</td>
<td>Baseline</td>
<td></td>
<td>BaP type</td>
<td>0.3 ug/ml</td>
<td></td>
<td>Beyer et al. 2003</td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Laboratory 1 ppm crude oil</td>
<td>14 days</td>
<td></td>
<td>Naph type</td>
<td>2,15 ug/g</td>
<td></td>
<td>Sundt, 2002</td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Feral fish Barents sea</td>
<td>Baseline</td>
<td></td>
<td>Pyren type</td>
<td>1.63 ug/g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Feral fish Barents sea</td>
<td>Baseline</td>
<td></td>
<td>BaP type</td>
<td>0.69 ug/g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Laboratory 1 ppm crude oil</td>
<td>14 days</td>
<td></td>
<td>Naph type</td>
<td>5.8 ug/g</td>
<td></td>
<td>Aas &amp; Børseth, 2002</td>
</tr>
<tr>
<td>Species</td>
<td>Substance (lab/field)</td>
<td>Test concentrations/area</td>
<td>Exposure time</td>
<td>Metabolite</td>
<td>Baseline</td>
<td>control reference or exposed/control</td>
<td>reference/comment</td>
</tr>
<tr>
<td>----------------------</td>
<td>-----------------------</td>
<td>--------------------------</td>
<td>---------------</td>
<td>------------</td>
<td>----------</td>
<td>--------------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>(Gadus morhua)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cod</strong> (Gadus morhua)</td>
<td>Laboratory</td>
<td>0.06 - 0.25 - 1 ppm Oil</td>
<td>average 3, 7, 14, 24 days</td>
<td>Naph type</td>
<td>3.9 ug/g</td>
<td>7.5 - 23.7 - 31.4</td>
<td>Skadsheim et al., 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>2.6 ug/g</td>
<td>3.6 - 10.6 - 13</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>1.0 ug/g</td>
<td>1.7 - 2.4 - 2.2</td>
<td></td>
</tr>
<tr>
<td><strong>Cod</strong> (Gadus morhua)</td>
<td>Laboratory</td>
<td>0.06 - 0.25 - 1 ppm Oil</td>
<td>average 3, 17, 31 day</td>
<td>Naph type</td>
<td>53.1 ug/g</td>
<td>0.7 - 2.3 - 2.9</td>
<td>Skadsheim et al., 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>7.0 ug/g</td>
<td>1 - 2.9 - 3.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>1.0 ug/g</td>
<td>1.1 - 1.5 - 1.5</td>
<td></td>
</tr>
<tr>
<td><strong>Cod</strong> (Gadus morhua)</td>
<td>Laboratory</td>
<td>Oil 0.06 - 0.25 - 1 ppm</td>
<td>30 days</td>
<td>Naph type</td>
<td>7.1 fi</td>
<td>5.1 - 9.5 - 227.5</td>
<td>Aas et al. 2000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>2 fi</td>
<td>6.4 - 12.7 - 43.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>0.8 fi</td>
<td>2.3 - 3.6 - 9.6</td>
<td></td>
</tr>
<tr>
<td><strong>Cod</strong> (Gadus morhua)</td>
<td>Laboratory</td>
<td>PW Oseberg, 1:1000 - 1:200 - 0.2 ppm oil + PAHmix</td>
<td>15 days</td>
<td>Naph type</td>
<td>12.6 ug/ml</td>
<td>1.3 - 2.5 - 3.6 - 5.4</td>
<td>Sundt, 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>4 ug/ml</td>
<td>1.7 - 3.7 - 4.1 - 17.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>1.8 ug/ml</td>
<td>1.3 - 1.8 - 1.5 - 2.4</td>
<td></td>
</tr>
<tr>
<td><strong>Cod</strong> (Gadus morhua)</td>
<td>Field, Caged</td>
<td>North Sea - Statfjord, 10000 m - 2000m - 500 m German bight G</td>
<td>5.5 weeks</td>
<td>Naph type</td>
<td>7.5 ug/ml</td>
<td>0.7</td>
<td>1.7 - 1.9 - 2.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>3.1 ug/ml</td>
<td>0.7</td>
<td>1.2 - 1.5 - 1.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>1.2 ug/ml</td>
<td>0.8</td>
<td>1.2 - 1.1 - 1.2</td>
</tr>
<tr>
<td><strong>Cod</strong> (Gadus morhua)</td>
<td>Field, Caged</td>
<td>German bight G4 (Ref) G1 - G2 - G3</td>
<td>5.5 weeks</td>
<td>Naph type</td>
<td>7.5 ug/ml</td>
<td>0.4</td>
<td>0.9 - 0.9 - 1.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>3.1 ug/ml</td>
<td>0.5</td>
<td>0.8 - 0.9 - 1.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>1.2 ug/ml</td>
<td>0.7</td>
<td>0.8 - 1 - 1.3</td>
</tr>
<tr>
<td><strong>Cod</strong> (Gadus morhua)</td>
<td>Field, Caged</td>
<td>North Sea - Troll,</td>
<td>6 weeks</td>
<td>Naph type</td>
<td>4.6 ug/ml</td>
<td>1.4</td>
<td>1.7 - 2.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Species</th>
<th>Substance (lab/field)</th>
<th>Test concentrations/area</th>
<th>Exposure time</th>
<th>Metabolite</th>
<th>Baseline</th>
<th>control reference</th>
<th>or exposed/control</th>
<th>reference/comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Gadus morhua)</td>
<td>Caged</td>
<td>1000 m - 500m</td>
<td></td>
<td>Pyren type</td>
<td>2.4 ug/ml</td>
<td>0.9</td>
<td>1.1 - 1.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>0.9 ug/ml</td>
<td>1.1</td>
<td>1.1 - 1.3</td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Field, Caged</td>
<td>North Sea - Tampen, 10000 - 2500 - 1000 - 500</td>
<td>6 weeks</td>
<td>Naph type</td>
<td>8.8 ug/ml</td>
<td>1.0 - 1.5 - 1.2 - 1.2</td>
<td></td>
<td>Hylland et al., 2005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>1.4 ug/ml</td>
<td>0.9 - 0.7 - 0.8 - 0.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haddock (Melanogrammus aeglefinus)</td>
<td>Feral fish Egersund</td>
<td>Baseline non polluted area</td>
<td></td>
<td>Naph type</td>
<td>5.1 ug/ml</td>
<td></td>
<td></td>
<td>Klungsøyr et al. 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>1.4 ug/ml</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>0.7 ug/ml</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haddock (Melanogrammus aeglefinus)</td>
<td>Feral fish Sleipner</td>
<td>Baseline polluted area?</td>
<td></td>
<td>Naph type</td>
<td>6.8 ug/ml</td>
<td></td>
<td></td>
<td>Klungsøyr et al. 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>1.9 ug/ml</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>0.8 ug/ml</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haddock (Melanogrammus aeglefinus)</td>
<td>Feral fish Statfjord</td>
<td>Baseline polluted area?</td>
<td></td>
<td>Naph type</td>
<td>11.2 ug/ml</td>
<td></td>
<td></td>
<td>Klungsøyr et al. 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>2.5 ug/ml</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>0.7 ug/ml</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haddock (Melanogrammus aeglefinus)</td>
<td>Feral fish Barents sea</td>
<td></td>
<td></td>
<td>Naph type</td>
<td>2.52 ug/g</td>
<td></td>
<td></td>
<td>Sundt, 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>1.69 ug/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>0.77 ug/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haddock (Melanogrammus aeglefinus)</td>
<td>Feral fish Barents sea</td>
<td></td>
<td></td>
<td>Naph type</td>
<td>2.0 ug/g</td>
<td></td>
<td></td>
<td>Aas &amp; Børseth, 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>1.3 ug/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>0.6 ug/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Species</td>
<td>Substance (lab/field)</td>
<td>Test concentrations/area</td>
<td>Exposure time</td>
<td>Metabolite</td>
<td>Baseline</td>
<td>control reference or exposed/control</td>
<td>reference/comment</td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------------------</td>
<td>---------------------------</td>
<td>---------------</td>
<td>------------</td>
<td>----------</td>
<td>--------------------------------------</td>
<td>-------------------</td>
<td></td>
</tr>
<tr>
<td>Haddock (Melanogrammus aeglefinus)</td>
<td>Feral fish</td>
<td>Frøy, ceased installation 10 000 m (ref) 2000 m - 200 m</td>
<td>Baseline polluted area?</td>
<td>Naph type</td>
<td>5.6 ug/ml</td>
<td>1.3 - 2.2</td>
<td>Beyer et al., 2003</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>1.4 ug/ml</td>
<td>1.4 - 0.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>0.75 ug/ml</td>
<td>1.8 - 0.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sheepshead minnow</td>
<td>Laboratory</td>
<td>North sea oil A 0.1 - 0.4 - 0.7 ppm</td>
<td>5 weeks</td>
<td>Naph type</td>
<td>6916</td>
<td>2.3 - 6.2 - 9.3</td>
<td>Bechmann et al. 2004</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>569</td>
<td>2.5 - 5 - 6.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>107</td>
<td>4 - 13.1 - 19.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sheepshead minnow</td>
<td>Laboratory</td>
<td>North sea oil B 0.1 - 0.9 - 5.6 ppm</td>
<td>6 weeks</td>
<td>Naph type</td>
<td>18164</td>
<td>1.8 - 4.3 - 12.5</td>
<td>Bechmann et al. 2004</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>438</td>
<td>5.6 - 12.6 - 30.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>110</td>
<td>12.6 - 42.7 - 123.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sheepshead minnow</td>
<td>Laboratory</td>
<td>2 - 14 - 214 ppb</td>
<td>5 weeks</td>
<td>Naph type</td>
<td>267280</td>
<td>0.9 - 2.2 - 18.6</td>
<td>Bechmann et al. 2004</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>9926</td>
<td>0.9 - 1.5 - 9.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>5152.7</td>
<td>3 - 17.4 - 207</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polar cod (Boreogadus saida)</td>
<td>Laboratory, feral fish 2001, 2002</td>
<td>1.5 ppm StatfjA oil baseline, control</td>
<td>14 days</td>
<td>Naph type</td>
<td>16.0 ug/g</td>
<td>2</td>
<td>16.9</td>
<td>Sundt &amp; Bechmann, 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>0.9 ug/g</td>
<td>5.5</td>
<td>74.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>0 ug/g</td>
<td>0</td>
<td>1.8</td>
<td></td>
</tr>
</tbody>
</table>
## Table 2. PAH-metabolites in marine fish – measured by GC-MS.

<table>
<thead>
<tr>
<th>Species</th>
<th>Substance (lab/field)</th>
<th>Test concentrations</th>
<th>Exposure time</th>
<th>Metabolite</th>
<th>Baseline</th>
<th>control or exposed/control</th>
<th>reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Feral fish</td>
<td>Barents sea</td>
<td>baseline</td>
<td>Naph sum</td>
<td>150.6 ng/g</td>
<td></td>
<td>Aas &amp; Børseth, 2002</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>61.2 ng/g</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren</td>
<td>4.6 ng/g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Feral fish</td>
<td>Barents sea</td>
<td>baseline</td>
<td>Naph sum</td>
<td>1285 ng/g</td>
<td></td>
<td>Sundt, 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>220 ng/g</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren</td>
<td>3.5 ng/g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Feral fish</td>
<td>Egersund</td>
<td>Baseline non polluted area</td>
<td>Naph sum</td>
<td>2005.1 ng/g</td>
<td></td>
<td>Klungsøy et al., 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>230.2 ng/g</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren</td>
<td>3.9 ng/g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Feral fish</td>
<td>Sleipner</td>
<td>Baseline polluted area?</td>
<td>Naph sum</td>
<td>1296.1 ng/g</td>
<td></td>
<td>Klungsøy et al., 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>197.8 ng/g</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren</td>
<td>0 ng/g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Feral fish</td>
<td>Statfjord</td>
<td>Baseline polluted area?</td>
<td>Naph sum</td>
<td>1361.7 ng/g</td>
<td></td>
<td>Klungsøy et al., 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>351.1 ng/g</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren</td>
<td>4.0 ng/g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Laboratory</td>
<td>0.06 - 0.25 - 1 ppm Oil</td>
<td>average 3, 7, 14, 24 days</td>
<td>Naph sum</td>
<td>2549 ng/g</td>
<td>4.6 - 13.4 - 23.6</td>
<td>Skadsheim et al., 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>691 ng/g</td>
<td>7.7 - 22.9 - 34.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren</td>
<td>27 ng/g</td>
<td>7.3 - 16.2 - 25.1</td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Laboratory</td>
<td>0.06 - 0.25 - 1</td>
<td>average 3, 17, 31</td>
<td>Naph sum</td>
<td>5702 ng/g</td>
<td>4 - 13.3 - 12.7</td>
<td>Skadsheim et al.,</td>
</tr>
<tr>
<td>Species</td>
<td>Substance (lab/field)</td>
<td>Test concentrations</td>
<td>Exposure time</td>
<td>Metabolite</td>
<td>Baseline control or reference</td>
<td>exposed/control</td>
<td>reference</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-----------------------</td>
<td>---------------------</td>
<td>---------------</td>
<td>------------</td>
<td>-----------------------------</td>
<td>----------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Cod (Gadus morhua) Field, Caged</td>
<td>North Sea - Statfjord, 500 - 2000 - 10000 m</td>
<td>ppm Oil day</td>
<td>Phen sum</td>
<td>377 ng/g</td>
<td>10.5 - 40.3 - 48.7</td>
<td>2004</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren</td>
<td>5 ng/g</td>
<td>8.6 - 63 - 88.4</td>
<td>Aas et al., in press</td>
</tr>
<tr>
<td>Cod (Gadus morhua) Field, Caged</td>
<td>North Sea - Troll, 1000 m - 500m</td>
<td>6 weeks</td>
<td>Naph sum</td>
<td>1150 ng/g</td>
<td>3.0 - 2.0 - 1.3</td>
<td></td>
<td>Børseth et al., 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>340 ng/g</td>
<td>3.5 - 2.7 - 2.5</td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua) Field, Caged</td>
<td>North Sea - Tampen, 10000 - 2500 - 1000 - 500</td>
<td>6 weeks</td>
<td>Naph sum</td>
<td>1515.1 ng/g</td>
<td>1.1</td>
<td>1.1 - 1.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>327.2 ng/g</td>
<td>1.6</td>
<td>2.1 - 2.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren</td>
<td>173.2 ng/g</td>
<td>1.2</td>
<td>0.9 - 1.2</td>
</tr>
<tr>
<td>Cod (Gadus morhua) Field, Caged</td>
<td>North Sea - Statfjord, 10000 m – 2000m - 500m</td>
<td>5.5 weeks</td>
<td>Naph sum</td>
<td>965.3 ng/g</td>
<td>0.9 - 1.7 - 0.9 - 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>934.5 ng/g</td>
<td>1.4 - 3 - 1.8 - 1.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren</td>
<td>3.7 ng/g</td>
<td>0 - 0 - 0.5 - 0.0</td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua) Field, Caged</td>
<td>German bight G4 (Ref) G1 - G2 - G3</td>
<td>5.5 weeks</td>
<td>Naph sum</td>
<td>228 ng/g</td>
<td>0.2</td>
<td>0.9 - 1.1 - 0.9</td>
<td>Aas et al., in press</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>482 ng/g</td>
<td>2.0</td>
<td>3 - 4.5 - 6.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren</td>
<td>28 ng/g</td>
<td>10.2</td>
<td>29.5 - 31.1 - 41.5</td>
</tr>
<tr>
<td>Cod (Gadus morhua) Field, Caged</td>
<td>Egersund non polluted area</td>
<td>Baseline</td>
<td>Naph sum</td>
<td>1346.9 ng/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>526.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Species</td>
<td>Substance (lab/field)</td>
<td>Test concentrations</td>
<td>Exposure time</td>
<td>Metabolite</td>
<td>Baseline</td>
<td>control or exposed/control</td>
<td>reference</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------</td>
<td>---------------------</td>
<td>---------------</td>
<td>------------</td>
<td>-----------</td>
<td>----------------------------</td>
<td>-----------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ng/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pyren</td>
<td>5.7 ng/g</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ng/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haddock (Melanogrammus aeglefinus)</td>
<td>Feral fish</td>
<td>Sleipner</td>
<td>Baseline polluted area?</td>
<td>Naph sum</td>
<td>1111.5 ng/g</td>
<td></td>
<td>Klungsøyr et al. 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>331.5 ng/g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pyren</td>
<td>10.4 ng/g</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ng/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haddock (Melanogrammus aeglefinus)</td>
<td>Feral fish</td>
<td>Statfjord</td>
<td>Baseline polluted area?</td>
<td>Naph sum</td>
<td>1279.7 ng/g</td>
<td></td>
<td>Klungsøyr et al. 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>331.9 ng/g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pyren</td>
<td>3.1 ng/g</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ng/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haddock (Melanogrammus aeglefinus)</td>
<td>Feral fish</td>
<td>Barents sea</td>
<td></td>
<td>Naph sum</td>
<td>1474 ng/g</td>
<td></td>
<td>Sundt, 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>165 ng/g</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pyren</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ng/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polar cod (Boreogadus saida)</td>
<td>Laboratory, feral fish 2001, 2002</td>
<td>1.5 ppm StatfA oil, baseline, control</td>
<td>14 days</td>
<td>Naph sum</td>
<td>1330 ng/g</td>
<td>1.3</td>
<td>114</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>538 ng/g</td>
<td>0.9</td>
<td>90</td>
</tr>
<tr>
<td>Pyren</td>
<td>52 ng/g</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14.6</td>
<td>60</td>
</tr>
</tbody>
</table>
Chapter 6 Externally visible fish diseases, macroscopic liver neoplasms and liver histopathology

Compiled by S W Feist and T Lang

Summary

1. 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.

2. 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).

3. 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 organisation. 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.

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

5. 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.

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

---

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

3 Federal Research Centre for Fisheries, Institute of Fishery Ecology, Deichstrasse 12, D-27472 Cuxhaven, Germany
Assessment of the applicability of fish disease and liver pathology techniques across the OSPAR maritime area

7. 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).

8. 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 Table 1 and 4).

9. 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 is available from monitoring programmes that were terminated in the 1990s or early 2000s (e.g. carried out by Belgium, Denmark and Sweden).

10. 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.

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

12. 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*).

13. 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

14. 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 standardised 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.

15. A number of practical ICES sea-going workshops on board research vessels were organised 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 standardise 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 flatfish livers for the occurrence of neoplastic lesions at a later stage. Further intercalibration and standardisation 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).

16. 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 formalised 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.

17. The ICES fish disease database is extended on an annual basis to include data from other species and areas within the OSPAR 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).

18. Current ICES WGPDMO activities have focussed 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 about 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.

19. 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

20. The multifactorial aetiology of diseases, in this context in particular of externally visible diseases, is generally accepted. Therefore, externally visible disease 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 organisation (molecule, sub-cellular units, cells, tissues, organs) leading to disease without involving pathogens.

21. 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 since 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.

22. 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 form 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 found to be 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 among the explaining quantities (Lang and Wosniok, 2000; Wosniok et al., 2000).

23. 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.

24. 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.

25. As indicated above, ICES has developed requirements for the international reporting of fish diseases over many years in order to minimise 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 2 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.

26. 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 minimising the possibility of mis-diagnosis. Ring tests and other intercalibration exercises are regularly undertaken in order to minimise 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.

27. 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**

28. 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).

29. 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 summarise 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.

30. 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, categorised 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 3 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

<table>
<thead>
<tr>
<th>Lymphocystis</th>
<th>Benign neoplasms</th>
<th>Non-specific lesions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epidermal hyperplasia/papilloma</td>
<td>Malignant neoplasms</td>
<td>Early non-neoplastic toxicopathic lesions</td>
</tr>
<tr>
<td>Acute/healing skin ulceration</td>
<td></td>
<td>Pre-neoplastic lesions (FCA)</td>
</tr>
<tr>
<td>X-cell gill disease</td>
<td></td>
<td>Benign neoplasms</td>
</tr>
<tr>
<td>Hyperpigmentation</td>
<td></td>
<td>Malignant neoplasms</td>
</tr>
<tr>
<td>Acute/healing fin rot/erosion</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Stephanostomum baccatum</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Acanthochondria cornuta</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Lepeophtheirus pectoralis</em></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Liver histopathology**
  - Non-specific lesions
  - Early non-neoplastic toxicopathic lesions
  - Pre-neoplastic lesions (FCA)
  - Benign neoplasms
  - Malignant neoplasms

31. 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.

32. The assessment of the FDI data generated is done on a region-wise 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 10 observations, e.g. in the case of biannual monitoring over a period of 5 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.

33. 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 5 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, 4).

34. 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 data set, in 2009 (ICES 2008, 2009). The results will be included in the OSPAR QSR 2010 as a case study.

35. 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**

35. 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’ since 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’ since this term more correctly describes the technique to be applied.

**Literature cited**


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

<table>
<thead>
<tr>
<th>Histopathology Categories</th>
<th>Histopathological Lesions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-specific lesions</td>
<td>Coagulative necrosis</td>
</tr>
<tr>
<td></td>
<td>Apoptosis</td>
</tr>
<tr>
<td></td>
<td>Lipoidosis</td>
</tr>
<tr>
<td></td>
<td>Haemosiderosis</td>
</tr>
<tr>
<td></td>
<td>Variable glycogen content</td>
</tr>
<tr>
<td></td>
<td>Increased numbers and size of macrophage aggregates</td>
</tr>
<tr>
<td></td>
<td>Lymphocytic/monocytic infiltration</td>
</tr>
<tr>
<td></td>
<td>Granuloma</td>
</tr>
<tr>
<td></td>
<td>Fibrosis</td>
</tr>
<tr>
<td></td>
<td>Regeneration</td>
</tr>
<tr>
<td>Early toxicopathic non-neoplastic lesions</td>
<td>Phospholipidosis</td>
</tr>
<tr>
<td></td>
<td>Fibrillar inclusion</td>
</tr>
<tr>
<td></td>
<td>Hepatocellular and nuclear polymorphism</td>
</tr>
<tr>
<td></td>
<td>Hydropic degeneration</td>
</tr>
<tr>
<td></td>
<td>Spongiosis hepatitis</td>
</tr>
<tr>
<td>Foci of cellular alteration</td>
<td>Clear cell foci</td>
</tr>
<tr>
<td></td>
<td>Vacuolated foci</td>
</tr>
<tr>
<td></td>
<td>Eosinophilic foci</td>
</tr>
<tr>
<td></td>
<td>Basophilic foci</td>
</tr>
<tr>
<td></td>
<td>Mixed cell foci</td>
</tr>
<tr>
<td>Benign neoplasms</td>
<td>Hepatocellular adenoma</td>
</tr>
<tr>
<td></td>
<td>Cholangioma</td>
</tr>
<tr>
<td></td>
<td>Haemangioma</td>
</tr>
<tr>
<td></td>
<td>Pancreatic acinar cell adenoma</td>
</tr>
<tr>
<td>Malignant neoplasms</td>
<td>Hepatocellular carcinoma</td>
</tr>
<tr>
<td></td>
<td>Cholangiocarcinoma</td>
</tr>
<tr>
<td></td>
<td>Pancreatic acinar cell carcinoma</td>
</tr>
<tr>
<td></td>
<td>Mixed hepatobiliary carcinoma</td>
</tr>
<tr>
<td></td>
<td>Haemangiosarcoma</td>
</tr>
<tr>
<td></td>
<td>Haemangiopericytic sarcoma</td>
</tr>
</tbody>
</table>
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)

<table>
<thead>
<tr>
<th>Disease category</th>
<th>Background</th>
<th>Elevated response/ above background</th>
<th>Significant response/ unacceptable effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Externally visible diseases (to be used as additional information for the assessment)</td>
<td>Not applicable</td>
<td>Statistically significant increase in mean FDI level in the assessment period compared to a prior observation period</td>
<td>Statistically significant increase in mean FDI level in the assessment period compared to a prior observation period</td>
</tr>
<tr>
<td></td>
<td></td>
<td>or</td>
<td>or</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Statistically significant upward trend in mean FDI level in the assessment period</td>
<td>Statistically significant upward trend in mean FDI level in the assessment period</td>
</tr>
<tr>
<td>Liver histopathology: non-specific (to be used as additional information for the assessment)</td>
<td>Not applicable</td>
<td>Statistically significant increase in mean FDI level in the assessment period compared to a prior observation period</td>
<td>Statistically significant increase in mean FDI level in the assessment period compared to a prior observation period</td>
</tr>
<tr>
<td></td>
<td></td>
<td>or</td>
<td>or</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Statistically significant upward trend in mean FDI level in the assessment period</td>
<td>Statistically significant upward trend in mean FDI level in the assessment period</td>
</tr>
<tr>
<td>Liver histopathology: contaminant-specific Mean FDI &lt; 2</td>
<td>Mean FDI ≥ 2</td>
<td>A value of FDI = 2 is 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.</td>
<td>Mean FDI ≥ 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>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.</td>
<td></td>
</tr>
<tr>
<td>Macroscopic liver neoplasms Mean FDI &lt; 2</td>
<td>Mean FDI ≥ 2</td>
<td>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 &gt; 2.</td>
<td>Mean FDI ≥ 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>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 &gt; 2.</td>
<td></td>
</tr>
</tbody>
</table>
Table 4: Fish disease monitoring in the OSPAR Coordinated Environmental Monitoring Programme (CEMP) reflecting ICES advice (ICES 2005)

Table 2a: PAH-specific biological effects monitoring

<table>
<thead>
<tr>
<th>Species</th>
<th>Diseases</th>
<th>Numbers</th>
<th>Guidelines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flounder (Platichthys flesus)</td>
<td></td>
<td>Size group ≥ 30 cm: 50 (if not available in sufficient numbers, include size group 25-29 cm)</td>
<td>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</td>
</tr>
<tr>
<td>Dragonet (Callionymus spp.)</td>
<td></td>
<td>Size group 10-15 cm: 50</td>
<td></td>
</tr>
<tr>
<td>Species</td>
<td>Diseases</td>
<td>Numbers</td>
<td>Guidelines</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>--------------------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Dab (1st priority) (Limanda limanda)</td>
<td>Lymphocystis</td>
<td>Size group 15-19 cm: 100</td>
<td>JAMP Guidelines based on:</td>
</tr>
<tr>
<td></td>
<td>Acute/healing skin ulcers</td>
<td>Size group ≥ 25 cm : 50</td>
<td>Relevant in addition:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>BEQUALM</td>
</tr>
<tr>
<td>Flounder (Platichthys flesus)</td>
<td>Lymphocystis</td>
<td>Size group 20-24 cm: 100</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Acute/healing skin ulcers</td>
<td>Size group 25-29 cm: 100</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Size group ≥ 30 cm: 50</td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Acute/healing skin ulcers</td>
<td>Size group &lt; 29 cm: 100</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Skeletal deformities</td>
<td>Size group 30-44 cm: 100</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pseudobranchial swelling</td>
<td>Size group ≥ 45 cm: 50</td>
<td></td>
</tr>
<tr>
<td>Whiting (Merlangius merlangus)</td>
<td>Epidermal hyperplasia/papilloma</td>
<td>Size group 15-19: 100</td>
<td>No JAMP guidelines so far</td>
</tr>
<tr>
<td></td>
<td>Lernaeocera branchialis</td>
<td>Size group 20-29: 100</td>
<td>Relevant:</td>
</tr>
<tr>
<td></td>
<td>Clavella adunca</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Macrophscopic liver neoplasms | Dab (1st priority)  
(Limanda limanda) | Size group ≥ 25 cm: 50  
(if not available in sufficient numbers, include size group 20-24 cm) | JAMP Guidelines based on:  
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 ≥ 30 cm: 50  
(if not available in sufficient numbers, include size group 25-29 cm) | |
| Liver histopathology | Dab (1st priority)  
(Limanda limanda) | Size group 20-24 cm: 50 | JAMP Guidelines based on:  
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 for Dragonet  
(*Limanda limanda* L.) and flounder (*Platichthys flesus* L.) for monitoring. ICES TIMES 38, 42 pp. |
Chapter 7 (Water) Bioassays including proposals for assessment tools and criteria

Executive summary

1. Applicability across the OSPAR maritime area. In vivo and in vitro 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 specific in vitro bioassays, which can be applied to saltwater, brackish water and freshwater, 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.

2. 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.

3. 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 data base for subsequent assessment as appropriate by ICES / OSPAR. A standardised 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 is currently compiled by the Netherlands and the UK and will be submitted for publication to ICES in 2007.

4. 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.

5. 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.

6. 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.

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

8. 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 (whole sediment/pore water) in the OSPAR maritime area.

Assessment of the applicability of water bioassays across the OSPAR maritime area

9. 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 freshwater whole effluents (e.g. Oris and Klaine, 2000; Power, 2004), and for dredged material (e.g. Stronkhorst et al., 2003).

10. 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 vitro and in vivo bioassays to coastal and offshore water column and micro surface layer monitoring (Stebbing et al. 1992; Hylland et al. 2002, 2006).

11. 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 recognised 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 vitro tests and in vivo bioassays with microorganisms 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 bioassays to the CEMP and status of quality assurance

12. 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.

13. The OSPAR JAMP CEMP lists water bioassays, sediment elutriate bioassays and pore 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-hr growth).

14. 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, standardised and “transcribed” into OSPAR guidelines. A standardised 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 is currently compiled by the Netherlands and the UK. Following review and approval by the ICES Working Group on the Biological Effects of Contaminants
Synergism between CEMP and WFD

15. 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 recently 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 and WFD in the case of bioassay applications in coastal and estuarine areas, but that further work and agreement are needed.

Thresholds and assessment tools

General

16. Thresholds are currently developed and will be available soon (ICESWGBEC 2007 meeting). 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\textsubscript{xx}, LC\textsubscript{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.

17. Water, sediment bioassays, in vivo and in vitro 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. For method to derive the background (assessment) levels of whole sediment bioassays, see background document on sediment bioassays. Pore water and elutriate bioassay can be considered with water bioassays insofar that these techniques involve a sediment manipulation procedure followed by the use of small volume water in vivo bioassays. Assessment criteria for in vitro bioassays are currently being developed and are mentioned shortly below.

Background responses and assessment criteria for (water) bioassays currently in JAMP

Water bioassays

18. The species recommended for water bioassays are:

- Copepod (\textit{Tisbe battagii} and \textit{Acartia} sp); 48hr exposure using mortality as the end point.
- Bivalves (\textit{Crassostrea gigas}, \textit{Mytilus} spp) embryos: 24 hr exposure using Percent Net Response as the end point.
- Sea urchin (\textit{Paracentrotus lividus}): 24 hr embryo exposure using percent normal development and larval length as the end points.

19. 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 \textit{Tisbe} assay.

20. In all water bioassays a control and positive control is used. The control is a “pristine water” of known water quality and characteristic ie no contamination, full salinity, appropriate pH and dissolved oxygen eg
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 to 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.

Assessing the data

21. 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.

22. 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 20% for the bivalve embryo bioassay. These values however need to be defined and further established (see also Table 3 below). Above these values would be the warning level and at 100% this would be categorised as a level of serious concern. Responses at the warning level would prompt further sampling and assay 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.

Sediment pore water and elutriates

23. Any of the four water bioassays listed above can be used to assess the toxicity of pore water and elutriates. The procedures used to prepare the pore and elutriates are well documented and described elsewhere. For the assessment of elutriate data, confounding factors that must be considered are: volume of sediment-water ratio, ammonia, sulphide, sediment quality – (i.e. sand or mud). For the assessment pore water data confounding factors that must be considered are: salinity, pH, DO, ammonia and sulphide. Data is produced and assessed in the same manner as for the water bioassays above and the results may be expressed in term of EC50 values and/or toxic units, depending on the purpose and the objectives of the study. Toxic elutriates and pore water can be diluted for testing if appropriate.

In vitro assays

24. In vitro bioassays are being developed for use with water, and sediment bioassays. In general this requires sample manipulation and/or concentration techniques, and clean-up using extraction procedures in analog to chemical compounds. 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 needs to be progressed within the current ICES OSPAR framework.

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

25. A preliminary derivation of background response levels was attempted at the meeting for the water bioassays using Tisbe batagillii, 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 was entered into a template (Table 2) and the following calculations made. Data from controls were collected for several test 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)

<table>
<thead>
<tr>
<th>Test</th>
<th>Test reference</th>
<th>Reference to the origin of the data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>year</td>
<td>Year of production</td>
</tr>
<tr>
<td></td>
<td>Country</td>
<td></td>
</tr>
<tr>
<td></td>
<td>lab</td>
<td>Laboratory that performed the analyses</td>
</tr>
<tr>
<td></td>
<td>type</td>
<td>Is it a control or other type of sample</td>
</tr>
<tr>
<td></td>
<td>Endpoint</td>
<td>Type of measurement</td>
</tr>
<tr>
<td></td>
<td>unit</td>
<td></td>
</tr>
<tr>
<td></td>
<td>idnr</td>
<td>Sample number within a data set</td>
</tr>
<tr>
<td></td>
<td>Replicates</td>
<td>Number of replicates</td>
</tr>
<tr>
<td></td>
<td>Result</td>
<td>Average value of the control</td>
</tr>
</tbody>
</table>

**Median**

**Minimum of the individual data**

**Maximum of the individual data**

**Standard deviation of the individual exposures**

**Information about sediment properties**

### Table 3: Preliminary results of background response levels for water bioassays

<table>
<thead>
<tr>
<th>Test</th>
<th>lab</th>
<th>Average</th>
<th>0.1 perc</th>
<th>0.5 perc</th>
<th>0.9 perc</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mussel</td>
<td>IEOV</td>
<td>Control</td>
<td>14.1</td>
<td>12.0</td>
<td>13.5</td>
<td>16.5</td>
</tr>
<tr>
<td>embryo</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test</th>
<th>lab</th>
<th>Average</th>
<th>0.1 perc</th>
<th>0.5 perc</th>
<th>0.9 perc</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Copepods</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tisbe</td>
<td>CEFAS</td>
<td>Control</td>
<td>1.3</td>
<td>0.0</td>
<td>0.0</td>
<td>5.0</td>
</tr>
</tbody>
</table>

### Ecotoxicological assessment criteria for bioassays

**Introduction and definitions**

26. 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.

27. The following definitions and terminology are used.
28. 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.

29. 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.

30. 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.

31. 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.

32. 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.

33. 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 saltwater and freshwater 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).

34. 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.

35. 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.

36. 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 in the case of 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**

37. 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.

38. 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**

39. 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

40. 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

41. 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 covers 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

42. 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 in the end points 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

43. 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 standardised method laid down in a protocol with validity criteria and control for modifying factors is essential for a stable bioassay. All bioassay tests now use positive controls; this consists of a standardised 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

44. 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.

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

financial

46. 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
realised 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

47. 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 micro-scale in operation, which by definition require less space and are more cost effective.

use of test animals

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

49. 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.

50. 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

51. The proposed normative 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.

52. 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 sea water 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:

53. 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 – freshwater, saltwater or brackish water – can be tested using the same (freshwater or saltwater) 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.

54. Bioassays conducted on surface water samples generally use a small sample volume, typically 20 – 100ml taken from a discrete water sample of say 2 litres. Water extraction procedures require a larger sample volume (egg 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.

55. 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.

56. 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.

57. 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 fertilisation and embryo development tests, can be a quick, cheap alternative, as can in vitro tests.

Assessment framework: metric and criteria

Experience in the Netherlands

58. 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)

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

Proposed metric and criteria for use of in vivo bioassays

60. 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):
   - 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 & Jaworska (2000). The criterion for the MPR (or MPE in this case) is the 95% protection level, or PAF\(_5\) (PAF = potentially affected fraction);
   - 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.

61. 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 freshwater, 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 & 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).

62. The boxes below detail the metrics for surface water.

---

**Standard for in vivo bioassays for surface water**

**Method 1: Standard with ‘preliminary effect assessment’ (Cf = concentration factor compared to the untreated sample (original water sample); this can be seen as the ‘assessment factor’ applied in the case of 3 acute or chronic tests from different taxonomic groups)**

**Acute tests**

<table>
<thead>
<tr>
<th>NE (negligible effect):</th>
<th>in 3 acute tests effect = 0 (in practice &lt; EC(_{50})), Cf = 100</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPE (maximum permissible effect):</td>
<td>in 3 acute tests effect = 0 (in practice &lt; EC(_{50})), Cf = 10</td>
</tr>
<tr>
<td>SE (serious effect):</td>
<td>in 1 acute test effect ≥ EC(<em>{50}), Cf = 10 or in 2 acute tests EC(</em>{20})&lt; effect &lt; EC(_{50}), Cf = 10</td>
</tr>
</tbody>
</table>

**Chronic tests**

<table>
<thead>
<tr>
<th>NE (negligible effect):</th>
<th>in 3 chronic tests effect = 0, Cf = 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPE (maximum permissible effect):</td>
<td>in 3 chronic tests effect = 0, Cf = 1</td>
</tr>
<tr>
<td>SE (serious effect):</td>
<td>in 1 chronic test effect ≥ EC(<em>{50}), Cf = 1 or in 2 chronic tests NOEC &lt; effect &lt; EC(</em>{50}), Cf = 1</td>
</tr>
</tbody>
</table>

EC\(_{50}\) = 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 PAF\(_{5}\) (for the MPE) and PAF\(_{50}\) (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 (PAF\(_{5}\))) is defined and divided by 10. This gives the concentration factor at which the NE acts. This result is compared to the results of the undiluted sample in order to determine whether this conforms to the MPE or the NE.
62. The standard for in vivo bioassays for sediment (whole sediment/pore water) 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 3 chronic tests must be available, including at least 2 ‘whole sediment’ tests.
- As for surface water, the MPE is: in 3 chronic tests effect \(= 0\). If a negative effect is measured in at least 1 of 3 chronic tests, the MPE is exceeded.
- The SE level is reached when an effect \(\geq EC_{50}\) is measured in 1 chronic test (on the ‘as is’ sample), or an effect between NOEC and EC\(50\) in 2 chronic tests.

63. 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 box). The SE has been derived only for surface water and not for sediment, as it is not possible to concentrate the sediment sample.

64. 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 proved 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).

65. 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 & 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 EC\(50\) values were extrapolated to chronic NOEC values with a factor 10. Although De Zwart & 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.

66. 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 3 bioassays.
Table 1: Indication of toxicity in surface water of the Western Scheldt estuary on basis of 3 different bioassay responses allowing a preliminary effect assessment as proposed in Maas et al., 2003

<table>
<thead>
<tr>
<th>location</th>
<th>date</th>
<th>Cf (EC50)*</th>
<th>Cf (MTE)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Daphnia</td>
<td>Algae</td>
</tr>
<tr>
<td>SvOD-1</td>
<td>12-2-00</td>
<td>42</td>
<td>20</td>
</tr>
<tr>
<td>SvOD-2</td>
<td>9-4-00</td>
<td>28</td>
<td>16</td>
</tr>
<tr>
<td>SvOD-3</td>
<td>11-6-00</td>
<td>54</td>
<td>2.4</td>
</tr>
<tr>
<td>SvOD-4</td>
<td>2-8-00</td>
<td>56</td>
<td>3.5</td>
</tr>
<tr>
<td>SvOD-5</td>
<td>17-10-00</td>
<td>96</td>
<td>4.5</td>
</tr>
<tr>
<td>SvOD-6</td>
<td>15-12-00</td>
<td>87</td>
<td>9</td>
</tr>
<tr>
<td>SvOD-1</td>
<td>13-01-05</td>
<td>95</td>
<td>20</td>
</tr>
<tr>
<td>SvOD-2</td>
<td>9-03-05</td>
<td>87</td>
<td>30</td>
</tr>
<tr>
<td>SvOD-3</td>
<td>2-05-05</td>
<td>127</td>
<td>17</td>
</tr>
<tr>
<td>SvOD-4</td>
<td>27-6-05</td>
<td>197</td>
<td>14</td>
</tr>
<tr>
<td>SvOD-5</td>
<td>23-8-05</td>
<td>251</td>
<td>10</td>
</tr>
<tr>
<td>SvOD-6</td>
<td>19-10-05</td>
<td>94</td>
<td>12</td>
</tr>
<tr>
<td>W Scheldt Vlissingen</td>
<td>4-6-03</td>
<td>416</td>
<td>52</td>
</tr>
<tr>
<td>W Scheldt Honte</td>
<td>4-6-03</td>
<td>180</td>
<td>56</td>
</tr>
<tr>
<td>W Scheldt Terneuzen</td>
<td>4-6-03</td>
<td>403</td>
<td>28</td>
</tr>
<tr>
<td>W Scheldt Hansweert</td>
<td>2-6-03</td>
<td>243</td>
<td>16</td>
</tr>
<tr>
<td>W Scheldt Boei s.v WO3</td>
<td>2-6-03</td>
<td>271</td>
<td>15</td>
</tr>
<tr>
<td>Scheldt Bath</td>
<td>3-6-03</td>
<td>271</td>
<td>9</td>
</tr>
<tr>
<td>Schaar vo Doel (SvoD)</td>
<td>3-6-03</td>
<td>92</td>
<td>9</td>
</tr>
<tr>
<td>Scheldt Antwerpen</td>
<td>18-6-03</td>
<td>144</td>
<td>2</td>
</tr>
</tbody>
</table>

- corrected for recovery

expected chronic effect in surface water:

- green = neglicable effect (NE)
- yellow = NE<effect< maximum permissible effect (MPE)
- red = serious effect (SE)
Proposed metric and criteria for use of in vitro bioassays

67. 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.

68. 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.

69. 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

70. 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 to 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.

71. 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, 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

72. 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

- *In vivo* and *in vitro* 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 is provided for by BEQUALM, therefore bioassay data can be submitted to the ICES data base 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 (sediment/pore 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 calculated and/or proposed but further work in this area is required.

References


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


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.


Figure 1

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)
**Chapter 8  Sediment seawater elutriate and pore-water bioassays with early developmental stages of marine invertebrates.**

**Background**

1. Early developmental stages are more sensitive than adults and the weakest link in the life cycle. The embryo-larval bioassays detect a broad spectrum of toxicants at comparatively low concentrations, in the order of 1 µg/L for TBT and other antifouling products, 10 µg/L for Hg, Cu and Zn, 100 µg/L for Pb, Cd and other metals, 1 mg/L for organochlorine pesticides, detergents and refined oil, and 10 mg/L for crude oil (Kobayashi 1995, His et al. 1999).

2. The embryogenesis and early larval development of marine invertebrates have been frequently used as a rapid, sensitive, cost-effective biological tool for the assessment of seawater, sediment elutriates and pore water quality. Detailed descriptions of methods and applications are available for bivalves (Woelke 1961, Thain 1991, His et al. 1999) and sea-urchins (ASTM 1995, Carr 1998). Gametes are obtained from mature adults either by stripping or thermally induced spawning, fertilized in vitro in a measuring cylinder and delivered into the experimental samples. After 24 to 48 h incubation at 18 to 24°C (depending on the species), samples are fixed and microscopically observed to record the percentage of normally developed larvae and, in the case of sea-urchins, larval length.

3. Sensitivity of embryos of different species to the main pollutants of concern in the marine environment is very similar, particularly within bivalves. This allows comparison of results of embryo-larval bioassays conducted with different species. A review on the EC50 values of 13 priority pollutants to bivalve vs. sea-urchin embryos reflected a correlation coefficient r=0.95 and a slope b=0.999. Due to their abundance and broad geographical distribution or availability from commercial sources the following species are recommended: *Crassostrea gigas*, *Mytilus edulis/galloprovincialis*, *Paracentrotus lividus*. In the case of sea-urchins, other species like *Strongylocentrotus droebachiensis*, extend the applicability of the assay with indigenous species to Northern countries (see figure).

4. Within bivalves, *Crassostrea gigas* and *C. virginica* oysters have been most often used for embryo-larval ecotoxicological bioassays because, unlike the mussel or the native flat oyster (*Ostrea edulis*), in *Crassostrea* fertile gametes can be obtained straight from the gonad by stripping, although this method requires high percentages of embryogenesis success in the controls to guarantee comparability of the results (His et al. 2000). The marine mussels of the *Mytilus* genus occurring in European waters (*M. edulis* and *M. galloprovincialis*) are nearly ubiquitous, easy to collect and to maintain in aquaria. Also these species show the advantage that the adults are commonly used in marine pollution monitoring programmes, and OSPAR encourages the use of the same species for different biological tools of pollution assessment, spanning molecular, cellular and individual responses. Another advantage of the mussel embryogenesis bioassay is that this species is tolerant to a broader range of salinities, including estuarine waters down to 20 ppt (His & Beiras, 1995). The *Paracentrotus lividus* sea-urchin has a somewhat more restricted distribution, but it is easier than bivalves to feed and maintain in captivity avoiding accidental spawning. Another advantage of the sea-urchin embryogenesis bioassay is to provide a quantitative, more gradual, observer-independent and statistically treatable response: larval length (Fernández & Beiras, 2001).
5. Currently, the main limitation of the embryo-larval bioassays is the availability of reliable, good quality biological material all year round, particularly outside the natural spawning season of the different species, which changes among different European countries. The maintenance of fertile adult stocks in aquaria is feasible, particularly for sea-urchins, and conditioned bivalves should be available from aquaculture facilities, but even commercial hatcheries are unable to provide 100% reliable adult broodstocks all year round. Cryopreservation of gametes of bivalves and sea-urchins is a promising solution to provide homogeneous biological material at any time, but up to date these techniques are still on development and standard methods are not available. Combination of different species with different spawning seasons seems to be still necessary.

6. Sediment toxicity can be tested with water column organisms by either obtaining an elutriate from the sediment (mixed with control sea-water) or by directly obtaining the interstitial pore-water from the sediment. The advantages of the first method are: smaller amounts of sediment and simpler equipment are necessary, the environmental parameters of the elutriate (dissolved oxygen, pH, salinity, ammonia, sulphides) are closer to those of the natural water column than in the case of pore water, in particular when dealing with anoxic or hypoxic sediments. These parameters are the most common source of false positives (see confounding factors), and pore water requires adjusting their values within the optimum range for the test species prior to testing. In reverse, pore-water has the advantage that no control sea-water is needed and the dilution of the potential toxicants present is lower, enhancing sensitivity. The choice of the method can depend on sampling constrictions and sample availability, since when the confounding factors are taken into account both methods yield comparable results (Beiras 2001).

7. The embryo-larval bioassay generally showed higher sensitivity to polluted sediments than the amphipod bioassay (Becker et al. 1990, Long et al. 1990, Carr & Chapman 1992), although similar sensitivities have also been reported (Williams et al. 1986). However not rarely the concordance in estimates of toxicity using different organisms is small, and different tests indicate different patterns in toxicity (Long et al. 1996). Therefore, comparisons among different sediment toxicity tests must be conducted using samples representing a broad range of pollution in order to evaluate the comparability of the different tests.

Confounding factors

8. In order to avoid false positives, water quality parameters in the elutriate (or pore water) must be checked prior testing and they must fall within optimum ranges for the embryo development of the test species or otherwise adjusted. In the case of molluscs His et al. (1999) provide a broad review on this topic. Generally speaking, full salinity, a pH higher than 7.5 and a dissolved oxygen concentration above 2 mg/L are required. This is particularly important in the case of pore waters from highly reduced sediments, which broadly depart from those values.

9. More often the presence of the toxic reduced compounds un-ionized ammonia and H₂S has been identified as the main sources of false positives in sediment elutriate toxicity testing (Cardwel et al. 1976; Matthiesen et al. 1998). Some threshold toxicity values for sea-urchin and bivalve embryos are available in the literature (Knezovic et al. 1996), but further research is strongly needed on this topic.
Regarding temperature, elutriates and pore waters are microbially rich and exposure to high temperatures during manipulation should be avoided. This includes centrifugation, when necessary. For incubation 18ºC (48 h) is recommended for mussels, 20ºC (48 h) for Paracentrotus lividus urchins and 24ºC (24 h) for Crassostrea gigas oysters.

**Ecological relevance**

10. The ecological relevance is one of the strong points of the embryo-larval bioassay. Any impairment of embryo development would lead to reduced recruitment and decrease population size.

**Assessment criteria**

11. Marine invertebrate embryo-larval bioassays have resorted to different species and a suit of endpoints. This issues need to be discussed prior to the implementation of assessment criteria.

**End-points measured.**

12. The end-point recorded in the standard embryo-larval bioassays is the percentage of morphologically normal larvae. The definition of morphological abnormalities change among authors and, obviously, among test species. For routine applicability’s sake it is advised that only very conspicuous abnormalities were taken into account. This would reduce the time necessary to record the endpoint, and facilitate automatization and observer-independence. In bivalves normal D-shape is advised as normality criteria. This excludes larvae with protruding mantle and convex hinge. Illustrations of these abnormalities can be found in Quiniou et al. (2005). However, more detailed abnormalities such as the presence of indentations in the larval shell would complicate observation and in our view should not be taken into account.

13. In sea-urchins normal larvae should exhibit four fully-formed arms (two longer post-oral arms and two shorter oral arms) and a normal outer contour of the body. Pre-pluteus stages where oral arms were not yet fully separated, or larvae with missing arms, should be considered as abnormal. However more detailed abnormalities such as those related to the internal anatomy of the larvae (skeletal rods, gut) would greatly complicate observation. Their identification even depends on the position of the larva under the microscope.

**Assessment criteria**

14. **Qualitative approach:** An elutriate can be classified as toxic when it induces a statistically significant reduction in the percentage of normal larvae compared to the elutriate from the reference site, for a confidence level of 95%. In the practice, undiluted elutriates from polluted sediments cause a very marked reduction in normal embryogenesis and the response is close to yes or not, so the identification of polluted sites is normally easy.

15. Mortality data must be arcsine transformed prior to analysis using ANOVA and a posteriori Dunnett’s test, comparing each sampling site with the reference site. The difficulty here is to establish a reference site we were sure from comprehensive analytical data that it is not polluted but was otherwise similar to the problem sites (see confounding factors). Control seawater may not be appropriate as reference because it lacks the physicochemical and microbiological properties of an elutriate, some of which may affect the response.

16. **Quantitative approach:** Once identified as polluted, the toxicity of any sediment elutriate that cause a marked inhibition in normal development can be quantified by serial dilution with reference seawater, and calculation of the toxic units (TU). TU = 100 / ED50, where ED50 is the theoretical dilution, expressed in percentage, that causes 50% abnormal larvae. This parameter can be obtained by fitting the data for the serial dilutions to standard toxicity curves (logit, probit, etc.). When data from different campaigns were pooled together for statistical analysis, they must be previously corrected by the respective controls by using Abbott’s formula: P’=(P-Pc)x100/(100-Pc); where P and P’ are the raw and corrected abnormality percentages, and Pc is the control abnormality. Once corrected, percentages must be arcsine transformed for subsequent analysis. When using this quantitative approach with sea-urchins, larval length after 48 h, or even better, size increase from fertilized egg after 24 h, is preferred to percentage of normal larvae. This is because size increase is a more sensitive -and thus more discriminant- response than morphologically normal development. In this case the effect level used for the calculation of the toxicity units must be lower than 50% (e.g. ED20, dilution of the elutriate that causes a 20% reduction in larval size), because the range of larval growth inhibition in 48h rarely attains 50%. For comparative purposes, the choice of the effect level is not important as long as it was always the same for all bioassays compared, obviously.
Quality assurance

17. Sediment manipulations during sampling, storage and testing, and quality of the test organisms have been often identified as the main sources of variability in sediment toxicity bioassays. Concerning the first point, sediments intended to toxicity testing should not be frozen but stored under refrigeration in the dark inside airtight containers, and tested within one week. Some authors argue that testing can be delayed by freezing the liquid phase (elutriate or pore-water) after elimination of particles. However it must be taken into account that glass fibre filters adsorb metals and some organic filters might retain organic compounds, so refrigerated centrifugation may be preferred. After thawing, samples should be shaken and salinity checked and adjusted, if necessary.

18. Concerning the effect of homogeneous biological material, interlaboratory comparisons carried out following strict protocols are necessary. In these intercalibrations it would be desirable that not only different populations of a certain species, but also different species (oysters, mussels, clams, sea-urchins) were included.

19. The control treatment in an embryo-larval bioassay gives essential information regarding biological quality of the test organisms. Acceptability criteria must be developed concerning minimum embryogenesis success and larval length in the control for a test to be considered reliable. Those criteria must take into account both the normal seasonal variability within a certain population and interpopulation variability. In the case of bivalves, His et al. (1997) reported mean values in controls ranging from 75.8 to 97.0, thus suggesting a minimum of 75% normality, whereas Quiniou et al. (2005) arbitrarily recommend a minimum of 80% normal D-larvae in the control as acceptability criterion. Preliminary results of background response levels for Mytilus embryo bioassays are shown in Table 1 below.

20. In the case of the P. lividus normal larval development, the distribution of the endpoints measured (percentage of normal larvae and larval length) in 155 controls throughout several years of tests conducted 48h at 20ºC was the following (Fernández, 2003):

![Graphs showing normal larvae distribution and larval length](image)

21. From these data, and taking the 5% percentile as the acceptability criteria, a test is correct when mean response in the control exceeds 91% embryogenesis success and 340 µm larval length.

22. Percentage fertilization prior to testing must always be recorded. To run a reference toxicant test may be further useful to check the biological quality of the test organisms.
Table 1. Preliminary results of background response levels for water and whole sediment bioassays

<table>
<thead>
<tr>
<th>Test</th>
<th>lab</th>
<th>Average</th>
<th>0.1 perc</th>
<th>0.5 perc</th>
<th>0.9 perc</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Corophium</strong></td>
<td>RIKZ</td>
<td>Control</td>
<td>12.3</td>
<td>6.6</td>
<td>10.5</td>
<td>19.3</td>
</tr>
<tr>
<td>Corophium</td>
<td>Cefas</td>
<td>Control</td>
<td>9.5</td>
<td>0.0</td>
<td>6.7</td>
<td>20.0</td>
</tr>
<tr>
<td><strong>Arenicola</strong></td>
<td>Cefas</td>
<td>Control</td>
<td>4.7</td>
<td>0.0</td>
<td>0.0</td>
<td>13.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test</th>
<th>lab</th>
<th>Average</th>
<th>0.1 perc</th>
<th>0.5 perc</th>
<th>0.9 perc</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mussel embryo</strong></td>
<td>IEOV</td>
<td>Control</td>
<td>14.1</td>
<td>12.0</td>
<td>13.5</td>
<td>16.5</td>
</tr>
<tr>
<td><strong>Copepods</strong></td>
<td>Tisbe</td>
<td>Control</td>
<td>1.3</td>
<td>0.0</td>
<td>0.0</td>
<td>5.0</td>
</tr>
</tbody>
</table>

References


Chapter 9  Sediment seawater elutriate and pore-water bioassays with copepods (*Tisbe*, *Acartia*), mysids, and decapod larvae (*Palaemon*).

**Background**

1. Crustaceans, and particularly early life stages, are several orders of magnitude more sensitive to insecticides than echinoderms and mollusca (Ramamoorthy & Baddaloo, 1995; Bellas et al. 2005). Crustaceans are also particularly sensitive to cadmium (Mariño-Balsa et al. 2000) compared to other marine invertebrates. Therefore when these contaminants were suspected the inclusion of a crustacean test within the battery of bioassays is strongly recommended.

2. Acute static survival tests with benthic (*Tisbe battagliai*) and planktonic (*Acartia tonsa*) copepods have been proposed to assess the biological quality of sediment elutriates (Matthiessen et al. 1998). Detailed methods are available (Hatchinson & Williams 1989, UNEP 1989). The endpoint recorded may be mortality or motility after 48 to 96h incubation in the test samples at 20°C and 16 h light 8 h dark photoperiod. *Tisbe battagliai* is an abundant component of meiobenthic fauna, whereas *Acartia* and other calanoid copepods are components of the holoplankton in Atlantic waters. Both are easy to feed on microalgae. Ovigerous females can be isolated and age-controlled cultures can be obtained from the eggs. A water bioassay programme is running within BEQUALM which includes the 48h *Tisbe battagliai* acute test.

3. Mysids, particularly the American species *Mysidopsis bahia*, are recommended test organisms by US-EPA for estuarine and marine water toxicity tests (US-EPA 2002). The maintenance of fertile adult stocks in aquaria, fed on *Artemia*, is feasible. Since these organisms undergo direct development in short time periods they are suitable for life cycle assessments. Some European mysids such as *Neomysis*, *Praunus* and *Siriella* have been proposed, but sensitivity intercomparisons are lacking. Also, the salinity range of tolerance for each species must be determined before recommendation for routine toxicity testing.

4. The use of decapods early life stages is less frequent (Cheung et al. 1997, Mariño-Balsa et al. 2000). The main advantages are the economic value of some species (shrimps, crabs), and the possibility to obtain ovigerous females from commercial stocks. The main restriction is to find broadly distributed species across all Europe. The *Palaemon* genus may be a potential candidate since it shows a broad geographical distribution, from Mediterranean sea to North Sea, they are easy to feed, the maintenance of fertile adult stocks in aquaria is feasible, and larval development is well known.

**Confounding factors**

5. In order to avoid false positives, water quality parameters in the elutriate (or pore water), specifically salinity, pH and dissolved oxygen, must be checked prior testing and they must fall within optimum ranges for the survival and motility of the test species or otherwise adjusted. This is particularly important in the case of pore waters from highly reduced sediments, which broadly depart from those values.

6. More often the presence of toxic reduced compounds, un-ionized ammonia and H₂S, have been identified as the main sources of false positives in sediment elutriate toxicity testing (Cheung et al. 1997). Further research is strongly needed on this topic.
Ecological relevance

7. Copepods and mysids are dominant components of holoplankton in marine ecosystems. They are primary consumers and an important food source for fish. Therefore any toxicant affecting them is a threat to the whole food web in coastal and oceanic ecosystems.

Assessment criteria

8. The issues about assessment criteria for these bioassays are very similar to those discussed in Annex 1.

Quality assurance

9. Sediment manipulations during sampling, storage and testing, and quality of the test organisms have been often identified as the main sources of variability in sediment toxicity bioassays. Concerning the first point, sediments intended to toxicity testing should not be frozen but stored under refrigeration in the dark inside airtight containers, and tested within one week. Some authors argue that testing can be delayed by freezing the liquid phase (elutriate or pore-water) after elimination of particles. However it must be taken into account that glass fibre filters adsorb metals and some organic filters might retain organic compounds, so refrigerated centrifugation may be preferred. After thawing, samples should be shaken and salinity checked and adjusted if necessary.

10. Concerning the effect of homogeneous biological material, interlaboratory comparisons carried out following strict protocols are necessary. In these intercalibrations it would be desirable that not only different populations of a certain species, but also different species (Tisbe, Tigriopus, Acartia, mysids, shrimp larvae...) were included.

11. Acceptability criteria must be developed concerning minimum survival/motility in the control for a test to be considered reliable. Those criteria must take into account both the normal seasonal variability within a certain population and interpopulation variability. A stringent acceptability criteria is essential to guarantee reliable toxicity data, particularly when test organisms come from wild populations and experience a sharp change in environmental conditions in the laboratory. Preliminary results of background response levels for Tisbe bioassays are shown in Table 1 (Annex 1).

12. To run a reference toxicant test may be further useful to check the biological quality of the test organisms. The reference toxicant, ideally, should be stable in aqueous solution and not dangerous for human beings.
References


Hutchinson TH, Williams TD (1989). The use of sheepshead minnow (Cyprinodon variegatus) and a benthic copepod (Tisbe battagliai) in short-term tests for estimating the chronic toxicity of industrial effluents.


Chapter 10 Whole sediment bioassays with amphipods (Corophium sp).

Background

1. The Rhepoxynius abronius amphipod test is commonly used in North America to evaluate the quality of sediments intended for dredging or dumping, and very detailed protocols are available (Swartz et al. 1985, ASTM 1992). The endpoint is survival after 10 days incubation in the whole sediment at 20°C. These protocols can be easily adapted to the European species (Corophium spp). Some efforts have already been made to compare methods and sensitivity for different amphipod species (van den Hurk et al. 1992, Casado-Martínez et al. 2006).

2. The Corophium genus is broadly distributed across Europe. An internationally agreed protocol for toxicity testing of offshore chemicals with C. volutator has been published (OSPAR 1995). ICES has also provided detailed methods (Roddie & Thain 2001). Those protocols are also suitable for other macroscopically indistinguishable Corophium species more abundant in Southern Europe, C. multisetosum. In fact ICES claims that the procedure can be used not only with any Corophium species but with any infaunal amphipod (Roddie & Thain 2001).

3. Other sediment dwelling species from different taxa (polychaetes, echinoderms, bivalves) may be also suitable after methodological standardisation and sensitivity comparisons with amphipods. Furthermore, Corophium is not tolerant to coarse grain sediments. Should sandy sediments be tested alternative species such as Arenicola, Echinocardium or Cerastoderma will be needed.

4. Some sublethal responses have been proposed as additional endpoints in order to enhance sensitivity, including reburial after the 10 day exposure (Bat & Raffaelli 1998), and 28-days growth (Nipper & Roper 1995). The later considerably delays the outcome of the test and may be a limitation for routine application. The use of fast growing juvenile stages might overcome this limitation.

Confounding factors

5. The presence of toxic reduced compounds such as un-ionized ammonia and H₂S in interstitial and overlying water has been identified as confounding factors in whole sediment toxicity testing (Phillips et al. 1997). The studies have been carried out with North America species. Further research on this topic with Corophium spp. is strongly needed.

6. Grain size also affects amphipod survival (De Witt et al. 1988). The studies have been carried out with North America species. Further research on this topic with Corophium spp. is strongly needed.
**Assessment criteria**

7. **Qualitative approach:** According to USEPA (1998) a sediment sample is classified as toxic when it induces a mortality 20% higher than control and the difference is statistically significant. Mortality data must be arcsine transformed prior to analysis.

8. **Quantitative approach:** Once identified as polluted, the toxicity of a sediment can be quantified by serial dilution with a reference sediment, and calculation of the toxic units (TU) of the sediment. \( \text{TU} = \frac{100}{\text{LD}_{50}} \), where \( \text{LD}_{50} \) is the theoretical dilution with reference sediment (expressed in percentage) that causes a 50% mortality to the amphipods. This parameter can be obtained from the mortalities of the serial dilutions by fitting those data to standard toxicity curves (logit, probit, etc.). When data from different tests were pooled together for statistical analysis, mortalities must be previously corrected by the respective controls by using Abbott's formula: \( P' = \frac{(P-Pc)\times100}{(100-Pc)} \); where \( P \) and \( P' \) are the raw and corrected mortality percentages, and \( Pc \) is the control mortality. Again, mortality data must be arcsine transformed prior to analysis.

**Quality assurance**

9. Sediment manipulations during sampling, storage and testing, and quality of the test organisms have been often identified as the main sources of variability in sediment toxicity bioassays. Concerning the first point, sediments intended to toxicity testing should not be frozen but stored under refrigeration in the dark inside airtight containers, and tested within one week.

10. Concerning the effect of homogeneous biological material, interlaboratory comparisons carried out following strict protocols are necessary. The following issues have been identified as relevant for the success of the intercalibration round. Sediment samples should be homogeneous in grain size and organic content but spanning from pristine to highly polluted. Preservation of the sediment from sampling to testing should be similar for all participants, including time and temperature. Since for this species with no commercial value the test individuals must be collected from the field, they should be acclimated and maintained in laboratory long enough to assess the population health prior to testing.

11. Acceptability criteria must be developed concerning minimum survival/reburial in the control for a test to be considered reliable. Those criteria must take into account both the normal seasonal variability within a certain population and interpopulation variability. A stringent acceptability criteria is essential to guarantee reliable toxicity data, particularly when test organisms come from wild populations and experience a sharp change in environmental conditions in the laboratory. In an intercalibration round in Spain, Casado-Martínez et al. (2006) set acceptable maximum control mortality at 10%, following USEPA (1994). Roddie & Thain (2001) raise this threshold to 15%. Preliminary results of background response levels for *Corophium* bioassays are shown in Table 1 (Annex 1).

12. The third year of a bioassay programme is running within BEQUALM from December 2006 to June 2007, and includes the 10-d *Corophium volutator* survival bioassay.

**References**


OSPAR 1995. PARCOM Protocols on methods for the testing of chemicals used in the offshore oil industry.


Chapter 11 DNA adducts

Background

1. 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 environmental carcinogens and DNA, termed DNA adducts, can be utilised as a biological marker of exposure to genotoxic compounds. DNA adducts can be removed by cellular repair processes or by cell death, but during chronic exposures they often reach steady state concentrations in carcinogen target tissues such as the liver. As a consequence, DNA adducts have several important features which make them suitable as biomarkers of carcinogen 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.

2. 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 (Ictalurus 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.

3. 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 characterisation is not required. It is this last feature that makes the assay particularly appropriate for 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 carcinogens (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

4. 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 freshwater 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 toxicopathic lesions including foci of cellular alteration and neoplasia (for review see Reichert et al., 1998).

5. Studies from North America and Europe suggest that DNA adduct levels 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 insure 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 effect 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

6. 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 32P-postlabelling assay is that it is not species specific and therefore can be utilised 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.

7. 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 of 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[α]pyrene, which bind to DNA to form adducts (Sikka et al., 1991). As mentioned previously, a major strength of the 32P-postlabelling 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 glad (Ericson et al., 1999; French et al., 1996; Lyons et al., 1997; Harvey and Parry, 1998).
Methodology and technical considerations

32P-postlabelling

8. In the 32P-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 32P-ATP. For hydrophobic aromatic DNA adducts, such as PAH-DNA adducts, the enrichment steps can enhance 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 radio-labelled at the 5'-hydroxyl using 32P-ATP and T4-polynucleotide kinase to form 3', (32P)5'-bisphosphates. Separation of the 32P-labeled adducts is accomplished by multidimensional high-resolution anion exchange thin-layer chromatography. Autoradiography is then used to locate the radiolabelled 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

9. The 32P-postlabelling assay uses large amounts of 32P, 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 32P and must be frequently monitored for exposure to 32P. In the UK the use of 32P 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 32P-postlabelling outside the UK must contact their own national licensing organisation to clarify the legislative procedures required.

10. Main considerations to help minimise and monitor 32P exposure:

- All researchers who handle 32P 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 minimise handling of tubes and vials is recommended.
- Double latex gloves are worn while handling 32P and they should be regularly checked for radioactivity by passing them under a radiation monitor. Gloves should immediately be changed and discarded if found to be contaminated.
- Laboratory working surfaces are checked frequently with the radiation monitor when handling 32P. 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 32P

11. All 32P 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 insure 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 standardised assays

12. There are presently no active QA programmes running for the detection of DNA adducts using the 32P-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 standardise 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 standardised protocol for the detection of bulky aromatic DNA adducts by the $^{32}$P-postlabelling assay (IARC, 1993; Phillips & Castegnaro, 1999). The standardised protocol has now been adopted by the International Programme on Chemical Safety (IPCS) 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**

13. It is recognised 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-postlabelling generated DNA adduct profiles either exhibited no detectable adducts or very faint diagonal radioactive zones (DRZs) (Fig 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.

**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 Prof. David Phillips and Dr Alan Hewer, Cancer Research Institute, Sutton UK). Figure adapted from Lyons et al., 2004b).

14. Using such studies to define reference locations it should be possible to gain an international consensus (via ICES Working Group for the Biological Effect of Contaminants, WGBEC) on what level of DNA adducts should be considered background for a particular location and species. However, the following issues will 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).

---

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

**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 $^{32}$p-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 viviparus*) and cod (*Gadus morhua*). A particular advantage of the $^{32}$P-postlabelling assay is that it is not species specific and therefore can be utilised on any organism deemed fit for purpose.

• **Status of quality assurance.** There are presently 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 standardised protocols are available in the form of an ICES Times technical document and IARC publications.

• **Assessment thresholds.** These have not been agreed, but are likely to be determined from comparisons of DNA adduct levels detected at pristine reference locations. It is recommended that further work to define baseline or background levels of DNA adducts using the 32-postlabelling assay is taken forward 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.


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.


Chapter 12 Reproductive success in eelpout (*Zoarces viviparus*)

**Background**

The eelpout (*Zoarces viviparus*), also called viviparous blenny, can be used as a bioindicator of the impact of hazardous substances on reproductive success of fish in the marine environment.

![Image of eelpout](image)

**Figure 1.** The eelpout is a viviparous fish and the pregnant female bears 20 – 300 living embryo and larvae in the ovarian cavity.

Studies of reproductive success in eelpout are recommend by ICES, OSPAR and HELCOM for marine monitoring programmes of biological effects (OSPAR 1997, HELCOM 2003, 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. As method quality assurance, some international and national workshops have been held in relation to the monitoring programmes (e.g. BEQUALM 2000, Strand 2005a).

Among other, the presence of abnormal development of embryo and larvae in eelpout broods has been suggested to be useful as a biomarker of impaired fish reproduction in the marine environment, as chronic exposure to various contaminants has the potential to induce adverse developmental effects in fish embryos and larvae. 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).

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, Ådvers et al. 2001, Sjölin et al. 2003, Strand et al. 2004, Kalmarweb 2005, Gercken et al. in press), 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.
According to the technical guidelines used in monitoring programmes (Neuman et al. 1999, Strand & Dahllöf 2004), various 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 be seen as supplementary parameters and how they can be integrated should be further evaluated.

Abnormal development of embryo and larvae in eelpout broods (Neuman et al., 1999, Strand et al 2004) can be characterised 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 & 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 to 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 can not 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.

The lower level, which can be associated with significant effects (class III), can be determined as the value significant different from the upper level of the background response tested by G-test and 2x2 contingency tables.

The interval in between (class II) is regarded as a kind of zone of uncertainty, where effects induced by environmental factors such as contaminants cannot be excluded.

The statistical analyses imply a sample size of 40 broods containing 40 larvae, i.e. in total 1600 larvae per station.

**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.

Based on the statistical analysis it can be argued that a frequency of >2%, >3 and >6% of malformed, late dead and growth retarded larvae, respectively, are significantly different, when the sample size consists of 40 pregnant female eelpouts with minimum 40 larvae in each brood (G > 3.84, p < 0.05, 2 × 2 contingency table).

The range between 90% percentiles and the level significant different from this can be regarded as a zone of uncertainty where effects induced by environmental factors like contaminants cannot be excluded (Table 1).

<table>
<thead>
<tr>
<th>Assessment class</th>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean frequency of malformed larvae</td>
<td>0 – 1%</td>
<td>&gt;1% - 2%</td>
<td>&gt;2%</td>
</tr>
<tr>
<td>Mean frequency of late dead larvae</td>
<td>0 – 2%</td>
<td>&gt;2% - 3%</td>
<td>&gt;3%</td>
</tr>
<tr>
<td>Mean frequency of growth retarded larvae</td>
<td>0 – 4%</td>
<td>&gt;4% - 6%</td>
<td>&gt;6%</td>
</tr>
<tr>
<td>Background response. The upper limit is the 90% percentile of response at reference sites.</td>
<td>Effects cannot be excluded.</td>
<td>Significant effect level compared to background response.</td>
<td></td>
</tr>
</tbody>
</table>

**Table 1. Proposal for assessment criteria for the mean frequencies of malformed larvae, late dead larvae and growth retarded larvae per station**

Comparisons of datasets shows that class III, i.e. significantly 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.
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 small 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>
<thead>
<tr>
<th>Assessment class</th>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency of broods with &gt;5% malformed larvae</td>
<td>0 – 5%</td>
<td>&gt;5% - 20%</td>
<td>&gt;20%</td>
</tr>
<tr>
<td>Frequency of broods with &gt;5% late dead larvae</td>
<td>0 – 5%</td>
<td>&gt;5% - 20%</td>
<td>&gt;20%</td>
</tr>
</tbody>
</table>

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.

Based on the statistical analysis it can be argued that a frequency of 20% of the broods containing >5% malformed as well as late dead larvae is significantly different from 5% when the sample size consists of 40 pregnant female eelpouts (G > 3.84, p < 0.05, 2 × 2 contingency table).

The range between 5% and 20% can be regarded as a kind of zone of uncertainty, where effects induced by environmental factors like contaminants cannot be excluded (Table 2).

Comparisons of the datasets show that class III, i.e. significantly 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).
**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, since differences have been shown between areas regarded as reference sites and not.

Proposals for three assessment classes of effect levels (I, II and III) 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**


---

**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 to the 90% percentile of the reference sites.


Appendix 1 Comments on the Reproductive Success text following the Second ICES/OSPAR Workshop on Integrated Monitoring of Contaminants and their Effects in Coastal and Open-Sea Areas (WKIMON II)\(^5\),

Underlying science described (including a review of environmental variables that influence the biological effect): The underlying science has been adequate described and with reference to relevant literature. There is short overview of relevant environmental variables and their effects.

Measurement described: The measurements of impaired larvae development are briefly described with reference to more detailed descriptions in the Swedish and Danish guidelines.

Supporting parameters detailed: Supporting parameters are briefly mentioned, but there should be referred to the Swedish guideline for further details.

Availability of Technical Annexes for method and supporting parameters: The technical annex 10 in the JAMP guidelines for general biological effects monitoring is not sufficiently detailed. It is suggested that Swedish guideline is accepted as a basis for the official JAMP guideline.

External QA: No proficiency testing schemes are going on at present. One Bequalm workshop has been held in 1999. National workshops have also been held in Denmark.

It is possible and necessary to organise an external QA, for instance by practical workshops or by sample exchange.

Assessment Criteria (including assessment of thresholds when the response can be considered to be of concern and/or require a response): The approach used for deriving the assessment criteria is acceptable, but the exact statistical analyses need further evaluation.

The basic approach is also valid for other biological effect techniques.

Applicability across the OSPAR maritime area: The eelpout is inhabiting coastal waters from the white Sea to southern north sea. However it is not equally abundant in all areas and it may therefore be difficult to sample adequate numbers in all areas. Regional assessment is more warranted.

It should be noted that eelpouts are protected in the pregnancy period in some areas and official sampling licence for monitoring activities should be obtained.

List of gaps:
- Reference to supporting parameters
- Ongoing external QA
- Further evaluation of the statistical analyses used for deriving the assessment criteria
- Applicability across the entire OSPAR maritime area

WKIMON II recommends:
To ICES: that a TIMES document be prepared on the determination of fish reproductive success.

\(^5\) To be addressed by ICES WGBEC
### WKI-MON II Review of the coverage the existing technical annex on reprod. success in eelpout for CEMP related monitoring

<table>
<thead>
<tr>
<th>Title</th>
<th>Status</th>
<th>Applicability across OSPAR area</th>
<th>Confounding variables</th>
<th>Measurement described</th>
<th>Supporting parameters</th>
<th>External QA Assessment criteria</th>
<th>Changes required?</th>
</tr>
</thead>
<tbody>
<tr>
<td>JAMP guidelines for general biological effects monitoring</td>
<td>Cat II</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technical Annex 10 – reproductive success in fish</td>
<td>Cat II</td>
<td>It is dependant on a single species, which is not endemic across the region. Therefore its use across the region is questionable. No Times document exists.</td>
<td>Confounding variables are not adequately discussed in the Annex or the background document.</td>
<td>Not adequately described in either the Annex or the background document.</td>
<td>Supporting parameters are not mentioned</td>
<td>No current external QA</td>
<td>The Annex requires some updating and improvement. A Times document should be produced.</td>
</tr>
</tbody>
</table>
OSPAR BACKGROUND DOCUMENT FOR BIOLOGICAL EFFECTS MONITORING TECHNIQUES: FISH VITELLOGENIN AS A BIOMARKER OF EXPOSURE TO XENOESTROGENS

Craig D Robinson and Alexander P Scott

June 2006
Chapter 13  Fish Vitellogenin As A Biomarker Of Exposure To Xenoestrogens

Compiled by Craig D Robinson⁶ and Alexander P Scott⁷

Executive Summary

The Need for Determining Vtg

1. Within the OSPAR Convention, the Joint Assessment and Monitoring Programme (JAMP) requires Contracting Parties to monitor the quality of the marine environment and the effects of activities or natural and anthropogenic inputs to it. The Hazardous Substances Strategy requires monitoring of certain Priority Chemicals and a “research effort on endocrine disruptors leading to the development of testing and assessment tools for identifying substances of concern and their occurrence and distribution and effect in the marine environment”. The identification of potential endocrine disruptors and the quantification of their environmental concentrations is potentially a complicated and expensive process, involving tiered toxicological screening and testing, and the development of sensitive chemical analysis techniques. An alternative approach is to assess whether effects on hormonal systems are occurring in the marine environment, and then address the question as to what the causative compounds are. At the SIME 2005 meeting, the UK recommended this approach for assessing the occurrence, distribution and effect of compounds that act as xenoestrogens (SIME 05/2/16). Xenoestrogens are exogenous compounds that act upon exposed organisms to produce responses usually associated with the action of natural oestrogen hormones. The recommended method is the measurement of the egg yolk protein vitellogenin in the blood plasma of male fish. This protein is naturally produced by female fish in response to endogenous oestrogens and incorporated into the developing eggs as yolk. Male fish exposed to oestrogens will also produce Vtg in the liver, but with no ovaries to sequester the protein, it accumulates in the blood from extremely low basal concentrations (<10 ng/ml), and can reach very high concentrations (>50 x 10⁶ ng/ml). This range of response helps makes it a very sensitive, as well as a highly specific, marker of oestrogenic exposure.

Field Study Design

2. The recommendations of the existing JAMP Guidelines on Contaminant Monitoring in Biota and on Contaminant-specific Biological Effects Monitoring also apply, with minor modification, to designing a sampling strategy to assess xenoestrogen contamination. The existing Guidelines recommend using cod (Gadus morhua) and dab (Limanda limanda) in monitoring surveys, with flounder (Platichthys flesus) specified as a suitable substitute. The Guidelines suggest collecting at least 12 fish of one sex of a constant size range, that sampling is undertaken outside the breeding season and that it always takes place at the same time of year. To allow an assessment of oestrogenic endocrine disruptors, cod (offshore) and flounder (estuaries) are the recommended species. Dab may potentially also be used as the offshore species. Additional requirements needed to allow assessment of xenoestrogen contamination and effects are that male fish should be used and that any surveys take place in January/February, prior to the offshore migration of flounder.

The Recommended Method of Determining Vtg

3. It is recommended that the Vtg in male fish plasma be determined by Enzyme-Linked ImmunoSorbent Assay (ELISA). There are many different types of ELISA assay, each with advantages and disadvantages,
and therefore a standard operating protocol should be utilised for each species. The Vtg ELISA should be sensitive enough to quantify basal Vtg concentrations (i.e. have detection limits <10 ng/ml). It is also strongly recommended that, for each species, all laboratories utilise a common supplier of Vtg antibody and antigen; this is to remove analytical differences resulting from the purification of slightly different mixes of the different Vtg proteins produced by each fish species. For cod, an ELISA using lipovitellin is recommended, due to the instability of Vtg from this species.

**Supporting Parameters**

4. As with the other biological effects techniques within the CEMP, data should be collected for a range of supporting parameters and co-factors. These include the location (latitude and longitude) and date of collection, water temperature at collection site, unique fish identifier, gender, notes of any grossly visible anomalies and lesions, fish weight and length, gonadosomatic index and the age of the fish. Sex should be confirmed and the occurrence and severity of intersex assessed by histological examination of a gonad. Bile samples can be used in bioassay-directed fractionation and analysis in order to identify any oestrogenic metabolites. If mRNA transcripts are to be quantitated, a liver sample should be preserved in liquid nitrogen. To enable integrated monitoring of contaminants and effects, further samples of bile, liver and muscle would be required for chemical analysis and for the determination of other biological effects measurements, such as EROD and metallothionein.

**Applicability across the OSPAR Maritime Region**

5. Cod, dab and flounder can be caught across the OSPAR area. A common source of Vtg antigen and antibody (Lv for cod) for use in a defined standard protocol is required to allow between-laboratory data comparison. The development of the antibody, antigen and protocol should be undertaken by an experienced laboratory. The establishment and validation of a standard ELISA protocol then ought to be straightforward for most laboratories.

**Quality Assurance**

6. In order to have confidence in the determination of Vtg in male fish, suitable analytical control and proficiency must be demonstrated. This can be achieved by charting the results from the repeated analysis of internal laboratory reference materials and by the participation in the BEQUALM laboratory proficiency scheme. BEQUALM has undertaken one ring-test for cod Vtg assays, with the five participants producing comparable results. However, if Vtg is to be included as a CEMP determinant, flounder (and dab) Vtg must be added to the BEQUALM scheme and more laboratories encouraged to participate.

**Assessment Criteria**

7. Environmental Assessment Criteria (EACs) cannot currently be set, due to a lack of scientific agreement as to whether Vtg is a marker of exposure or effect, and due to an absence of data correlating Vtg concentrations with higher-level effects in OSPAR sentinel fish species. Sufficient UK data are available to be able to propose a provisional Background Concentration of 0.13 µg/ml for Vtg in male flounder. Less data are available for cod, and a more tentative BC of 0.23 µg/ml is proposed, based upon a single study. Data for dab are not yet available, although one study is nearing completion. These figures should be reviewed once data conforming to the recommended sampling and QA regime are available for the wider OSPAR area.

**Introduction**

8. Under the Convention for the Protection of the Marine Environment of the North East Atlantic (the “OSPAR Convention”), which came into force in 1998, the 16 Contracting Parties have agreed six Strategies (OSPAR, 2003) by which they will “take all possible steps to prevent and eliminate pollution and to take the necessary measures to protect the maritime area against adverse effects of human activities so as to safeguard human health and to conserve marine ecosystems and, when practicable, restore marine areas which have been adversely affected”. This chapter addresses the use of vitellogenin assays as measures of the biological effects of endocrine disrupting chemicals, specifically xenoestrogens, in OSPAR-wide monitoring. As such, it is relevant to issues raised under the OSPAR Joint Assessment and Monitoring Programme (JAMP) and Hazardous Substances Strategy (HSS).
9. The JAMP assessments include evaluating the effectiveness of measures taken and planned for the protection of the marine environment, through the repeated measurement of:
   a) **the quality of the marine environment** and each of its compartments;
   b) **activities or natural and anthropogenic inputs** which may affect the quality of the marine environment;
   c) **the effects** of such activities and inputs.

10. Participation in a Co-ordinated Environmental Monitoring Programme (CEMP) ensures that environmental monitoring by the Contracting Parties addresses the aims of the JAMP.

11. The OSPAR Hazardous Substances Strategy (HSS) aims to achieve concentrations in the marine environment near background values for naturally occurring substances and close to zero for man-made synthetic substances (OSPAR, 2003). Included within the HSS are lists of “Chemicals for Priority Action” and “Substances of Possible Concern”. These include some compounds selected due to their effects on endocrine systems. Under the HSS, the OSPAR Commission will “collaborate with various international forums with a view to optimising international research effort on endocrine disruptors leading to the development of testing and assessment tools for identifying substances of concern and their occurrence and distribution and effect in the marine environment”.

12. Testing compounds for hormonal effects and then deciding which should be monitored in the environment is very costly as it would require thousand of compounds to be screened and tested for their hormonal effects, and then the development of analytical techniques to determine the environmental concentrations of selected compounds. An alternative approach is to determine whether biological effects are occurring, and then to identify the compounds responsible; this involves less toxicological testing and has the added benefits of indicating contaminant bioavailability and of measuring effects of exposure to chemical mixtures. The use of biological effects monitoring to investigate whether ED compounds are present in the environment is also consistent with the European Union Water Framework Directive (EU, 2000).

13. Monitoring of Priority Chemicals is required under the CEMP. Several Priority Chemicals, or groups of chemicals are endocrine disruptors. The List of Chemicals of Possible Concern includes synthetic hormones (e.g. diethylstilbestrol and ethynylestradiol), natural hormones (oestradiol and oestrone) and other compounds that have ED action (e.g. bisphenol A). Most of these compounds are xenoestrogens, and cause biological effects through their interaction with the oestrogen receptor. The use of a biological effects measurement (vitellogenin) can allow an assessment of the occurrence and distribution and effect of xenoestrogenic endocrine disruptors in the OSPAR maritime environment; the individual problem compound(s) could subsequently be identified and concentrations determined at targeted sites. Such an approach was recommended in a report from the UK (SIME 05/2/16) to the March 2005 meeting of the OSPAR Working Group on Concentrations, Trends and Effects of Substances in the Marine Environment (SIME).

14. The 2005 SIME meeting also produced terms of reference (SIME 05/10/1, para. 2.17) for a review of the biological effects component of the CEMP, requiring the development of background documents for biological effects monitoring techniques currently included in the CEMP. The meeting additionally asked the UK to elaborate on how to apply techniques for measuring Vtg in OSPAR monitoring and invited parties with experience of using Vtg to inform the SIME 2006 meeting of their experiences. Several Contracting Parties indicated that they already undertake monitoring of xenoestrogens, or are considering so doing, by the determination of the egg yolk protein vitellogenin in male fish and there is existing recognition within ICES of the usefulness of vitellogenin measurements in biological effects monitoring (Scott and Hylland, 2002; WGBEC, 2005). Building upon experience from its Endocrine Disruptors in the Marine Environment (EDMAR) programme, the UK volunteered to develop an OSPAR background document on the use of vitellogenin as a tool for assessing xenoestrogens in the marine environment.

**Background Scientific Information**

15. There are a number of publications reviewing the hormonal control of vitellogenesis and the role of vitellogenin in fishes and that have been used in preparing this paper. The reader is referred to the following for the primary sources of information: Specker and Sullivan (1994), LaFluer (1999), Jalabert *et al.* (2000),

Vitellogenin – Definition and Properties

16. Vitellogenin (Vtg) is a large, calcium-rich, phospholipoglycoprotein. It is produced by the liver of female oviparous vertebrates in response to increases in circulating oestrogens. It is then transported in the bloodstream to the ovary, where it is sequestered by developing oocytes and processed to form egg yolk. The main elements involved in the production and utilisation of vitellogenin during oocyte maturation are shown in Figure 1 and described below. Concentrations of Vtg in the blood plasma of vitellogenic female fish can be six orders-of-magnitude higher than in immature or male fish. The production and utilisation of the optimal quality and quantity of Vtg is thus of critical importance to the reproductive success of most fish species.

17. As well as being the major source of amino acids and lipids for the developing embryo and a nutrient reserve for hatched larvae, Vtg also binds cations such as Ca, Mg, Zn and Fe and transports them into the oocyte for use during future embryogenesis. It binds steroid and thyroid hormones (Cyr and Eales, 1996; Scott et al., 1995; Tagawa et al., 2000) and may transport these into the egg to control embryo development.

18. In blood plasma, Vtg is transported as a dimer with a molecular mass of 300-600 kDa that is readily cleaved into the monomer protein (150-200 kDa). It consists of approximately 80% protein and 20% lipid (of which 70% is phospholipids) and includes three main component yolk proteins, two lipid-rich (lipovitellin I, Lv-I; lipovitellin II, Lv-II) and one phosphorous-rich (phosvitin, Pv). Sequence homology indicates that Lv-I and Lv-II are more conserved between taxa than is the Pv domain.

19. When examined by polyacrylamide gel electrophoresis (PAGE), several smaller bands are also present, often due to breakdown of Vtg that may occur during collection, purification and/or storage. In some species, the breakdown is increased dramatically by the process of freeze-drying (Arukwe and Goksøyr, 2003). In some species, there are two major bands between 150-200 kDa due to the presence of more than one type of Vtg. Species, such as white perch (Hiramatsu et al., 2002) and mosquito fish (Sawaguchi et al., 2005) possess three types of Vtg (denoted VtgA, VtgB and VtgC), while others such as Japanese goby have two types (Ohkubo et al., 2004). VtgC lacks phosvitin.

Reasons for Natural Occurrence

20. Vtg belongs to an ancient gene family, being found in many invertebrates as well as all oviparous vertebrates. Following changes in environmental stimuli (Fig. 1), such as photoperiod, temperature or food availability, the hypothalamus secretes gonadotrophin releasing hormone (GnRH) to stimulate the pituitary to synthesise and release gonadotrophin hormones (GtH I and GtH II). These cause the immature ovary to release the hormone oestradiol (E2) into the bloodstream. Oviparous females produce Vtg in response to the rise in circulating E2 titres. Within liver cells, E2 binds to the nuclear oestrogen receptor (ER), which dimerises and binds to the oestrogen responsive element (ERE) of the promoter region of E2-inducible genes, including Vtg and the ER itself. The binding of the ER to the ERE thus results in increased mRNA transcription and consequent production of E2-inducible proteins. Hepatically produced Vtg is transported in the bloodstream to the ovary, and is sequestered by the developing oocytes. The incorporation of large amounts of Vtg into the oocytes results in the characteristic increase in the size of the ovaries during sexual maturation.

21. During oocyte maturation, proteolysis of Lv derived from VtgA is used to generate free amino acids that, by increasing the osmotic pressure within the oocytes, draw in water. Selective proteolysis of yolk proteins gives rise to distinct differences in the free amino acid profiles of pelagic and demersal eggs (Finn et al., 2002) and accounts for their differing buoyancy as proteolysis is especially important in fish with pelagic eggs (Reith et al., 2001; Matsubara et al., 2003). The Lv derived from VtgB is mainly laid down in the form of yolk within membrane-bound platelets and there is evidence that the developing embryo utilises Lv-derived lipid whilst the protein component is reserved for larval use (Hartling and Kunkel, 1999).
Figure 1  Diagram of the brain-pituitary-gonad-liver axis in female teleosts, showing the main elements involved in production and utilisation of vitellogenin during oocyte maturation. E2, 17β-oestradiol; ER, oestrogen receptor; ERE, oestrogen responsive element; GnRH, gonadotrophin releasing hormone; GtH, gonadotrophin hormone; HSP, heat shock proteins; Lv, lipovitellin; MIS, maturation inducing steroid; Pv, phosvitin; Vtg, vitellogenin; Vtg-R, vitellogenin receptor.
Vitellogenic Response of Fish to Contaminant Exposure

22. Hepatic Vtg production in female fish is controlled by 17β-oestradiol (E2). However, Vtg genes are also present in male fish and the artificial administration of E2 to immature or male fish results in dose-responsive increases in Vtg mRNA expression, and plasma Vtg titres. In most species, Vtg concentrations in plasma can rise by a million-fold in response to oestrogen stimulation. No other recorded biological effect has such a high range of response. Even very low doses of oestrogen can result in large increases in plasma Vtg, making the determination of Vtg in male fish an extremely sensitive marker of oestrogen exposure (Sumpter and Jobling, 1995).

23. Many natural and artificial compounds are able to act as oestrogens. The degree to which an oestrogenic compound is able to induce Vtg is governed by its ability to bind to the hepatic oestrogen receptor alpha (ERα). Many compounds (e.g. alkylphenols, many halogenated organic compounds, certain pesticides, some phthalate plasticizers, paraben preservatives and phytosterols) are weakly oestrogenic, whilst a few pharmaceutical compounds (e.g. EE2, DES) may be more potent that E2 itself. Some organochlorine compounds, including \( o,p'\)-DDT, methoxychlor and certain PCB congeners are themselves very weak ER agonists, but their hydrolysed metabolites are much stronger ER agonists (Bulger et al., 1978; Korach et al., 1988; Soontornchat et al., 1994), although as metabolites they can be expected to be rapidly eliminated from the body.

24. ER agonists have the same mechanism of action and thus exposure to mixtures of exogenous oestrogen mimics results in increased Vtg response. The effects of exposure to multiple oestrogenic compounds are additive, based upon the relative concentrations and potencies of the individual components of the mixture (Payne et al., 2000; Thorpe et al., 2003; Brian et al., 2005). Consequently, even when the individual compounds are each below the threshold concentration that causes Vtg induction, the overall potency of the mixture may be sufficient to induce Vtg. This allows the potencies of oestrogenic mixtures to be determined (in vitro or in vivo) in terms of E2-equivalents (e.g. ng E2-equivalents l\(^{-1}\)), in an analogous manner to the determination of toxic equivalents for planar organic compounds. Importantly, it also allows an assessment of whether fish are being exposed to xenoestrogens in the environment, even when the concentrations of individual compounds may be very low and difficult to detect through chemical analysis. This is one of the key reasons for recommending the determination of Vtg in male fish as a method allowing the monitoring and assessment of the occurrence and effects of oestrogenic endocrine disruptors in the marine environment.

25. Co-exposure with certain other compounds can result, however, in perturbations to this model. For example, some compounds block the binding of E2 to the ER and thus are anti-oestrogenic. Others (e.g. some planar organic compounds, certain metals) have indirect anti-oestrogenic effects. Many polycyclic aromatic hydrocarbons (PAH), and other planar organic compounds that interact with the arylhydrocarbon receptor (AhR) to cause induction of the cytochrome P450 1A (CYP1A) detoxification enzyme system, are also believed to be indirectly anti-oestrogenic. Female exposure to these compounds can lead to increased E2 metabolism and reduced plasma E2 titres, thus inhibiting Vtg production, impairing oocyte development and reducing reproductive output (reviewed in Nicolas, 1999). The inhibition of oestrogenic effects by AhR agonists may also be due to a reduction in ERα expression (Bermanian et al., 2004). The effects of simultaneous exposure to ER-agonist(s) and AhR-agonist(s) on Vtg expression in fish are complex and not well understood. Responses appear to depend upon the relative ratios of ER-agonist to AhR-agonist, the sex and maturation state of the fish, and the sequence of exposures (Anderson et al., 1996; Arukwe and Goksøyr, 2003). Despite these complications, male flounder from UK estuaries contaminated by AhR-inducers and ER-agonists demonstrate both elevated plasma Vtg (Kirby et al., 2004a) and induced CYP1A enzyme activity (Kirby et al., 2004b). Thus Vtg can be used as a marker of oestrogenic exposure in environments that are also contaminated with AhR-inducers. Recent unpublished laboratory studies by Kirby et al. have additionally shown that AhR-inducers had no influence on the ability of E2 to induce Vtg in immature flounders. However, E2 was very effective at inhibiting the ability of AhRs to induce CYP1A.

Existing Evidence for Effect of Oestrogenic EDs on Marine Fish

26. There is now a large body of evidence for the existence of oestrogenic endocrine disruption in the marine environment. Male flounder (Platichthys flesus) caught in industrialised estuaries of the UK and the Netherlands have been found with elevated concentrations of Vtg in their plasma (Lye et al., 1997, 1998; Matthiessen et al., 1998; Allen et al., 1999a, 1999b; Vethaak et al., 2002, 2005; Kirby et al., 2004a; Kleinkauf et al., 2004). The range of Vtg concentrations that has been measured is great (from <10 ng/ml to >50 x 10\(^6\) ng/ml). Some male flounder with elevated Vtg concentrations have also been caught in the open sea (Allen et al., 1999a), but these were hypothesised to be fish that had recently emigrated from a contaminated
estuary. In estuarine and coastal areas of the USA (Mills et al., 2003; Roy et al., 2003) and Japan (Hashimoto et al., 2000; Ohkubo et al., 2003; Hara et al., 2001), many fish have also been found with high concentrations of Vtg in their plasma. There is evidence that cod from some inshore areas of Norway have elevated Vtg (Scott et al., 2006).

27. In the open seas, oestrogenic effects have been observed in swordfish (*Xiphias gladius*) from the Mediterranean (Fossi et al., 2001; 2002; 2004; Desantis et al., 2005) and off the coast of South Africa (Desantis et al., 2005), but not in the Pacific Ocean (Desantis et al., 2005). Similarly, many male tuna (*Thunnus thynnus*) caught in the Mediterranean had Vtg in their plasma (Fossi et al., 2002) while tuna (*Thunnus obesus*) caught in the Pacific Ocean (Hashimoto et al., 2003) did not. Plasma Vtg concentrations in male cod from the NE Atlantic show a positive correlation with size, possibly due to changes in diet as they grow or to long-term accumulation of contaminants (Scott et al., 2006).

28. In addition to direct evidence of oestrogenic effects (i.e. Vtg protein or mRNA in males), there is also indirect evidence. The presence in males of intersex gonads (Lye et al., 1997; Allen et al., 1999a; Cho et al., 2003; De Metrio et al., 2003) or feminised secondary sexual characteristics (Kirby et al., 2003) also suggests oestrogenic exposure. However, sexual differentiation in fish is very plastic. Other types of compounds (e.g. aromatase inhibitors, androgens and anti-androgens), or even temperature change, are able to alter fish gender (Baroiller et al., 1999; Al-Abtani and Phelps, 2002). In contrast, only oestrogenic compounds are so far known to be involved in the induction of Vtg (Sumpter and Jobling, 1995). Field studies of flounder in the UK have indicated a poor association between Vtg induction and male intersex (Allen et al., 1999b). For example, at a site with a prolonged history of high Vtg induction (the Tees estuary in the UK), hermaphrodites have yet to be recorded (Kirby et al., 2004a).

**Methodology**

**Design of Field Surveys**

29. The existing JAMP Guidelines for Monitoring Contaminants in Biota, or Contaminant-specific Biological Effects (OSPAR agreements 1999-02 and 2003-10) can be readily adapted for use in the assessment of oestrogenic exposure through the determination of Vtg in male fish. The similarity between the Guidelines and the benefits of integrating chemical analysis with biological effects determinations has already been noted. For this reason, the 2005 and the 2006 meetings of the Joint ICES/OSPAR Workshop on Integrated Monitoring (WIKIMON) produced draft JAMP Guidelines for the Integrated Monitoring and Assessment of Contaminants and their Effects ("The MERGED Guidelines"; WIKIMON 2006, draft report). Further advice on the use of Vtg in biological effects monitoring, including field survey design, is available in the form of an ICES Techniques in Marine Science (TIMES) paper (Scott and Hylland, 2002).

30. The existing JAMP Guidelines recommend using cod (*Gadus morhua*) and dab (*Limanda limanda*) in monitoring surveys, with flounder (*Platichthys flesus*) specified as a suitable substitute for dab. The Guidelines suggest collecting at least 12 fish of one sex and of a constant size range; that sampling is undertaken outside the breeding season and that it always takes place at the same time of year. These recommendations (for reasons outlined below) would also apply when designing a sampling strategy to assess xenestrogen contamination. Additional requirements are for male fish to be used and for any survey to be conducted in late winter/early spring.

31. There are sufficient data from the UK, and elsewhere, to demonstrate that male flounder are suitable for the assessment of xenestrogenic contamination of estuarine environments (e.g. Allen et al., 1999a;b; Vethaak et al., 2002, 2005; Kirby et al., 2004a; Kleinkauf et al., 2004). Preliminary data from caged fish in the BECP and AG project (Scott et al., in press) and a further survey of wild fish (Scott et al., 2006) suggest that male cod may be suitable for use in offshore areas. The UK has recently developed an ELISA for dab and initial findings suggest that this species may also be suitable as a sentinel in monitoring oestrogenic contamination, although the development of a more sensitive ELISA will be necessary since Vtg concentrations are on average lower than those found in cod (Scott et al., in prep.).

32. UK data for flounder indicates a marked seasonality in male plasma Vtg concentrations, with concentrations being highest in February/March and lowest in June/July (Kirby et al., 2004a; Kleinkauf et al., 2004). This reflects migration of the fish to and from the estuary and indicates that annual surveys of flounder Vtg must take place just before the fish are expected to migrate offshore to breed (i.e. in late winter/early spring).
33. There is evidence that plasma Vtg concentrations are correlated with fish size for species caught away from point sources, for example in dab (Scott et al., in prep), cod (Scott et al., 2006), tuna and other species from high trophic levels (Barucca et al. 2006). To reduce biological variability, it is recommended that only fish from a narrow size range be analysed for Vtg. To assess recent xenoestrogen contamination, smaller cod (30-45 cm) should be sampled, as recommended in the existing JAMP Guidelines.

34. Spermiating male plaice can show slight Vtg induction (Scott et al., 1999) that is thought to be related to either endogenous production of E2 (Wingfield and Grimm, 1977) and/or to androgen-ER interactions (Kim et al., 2003) and, for this reason, Vtg surveys should avoid the breeding season. At contaminated sites, there are often large (several orders of magnitude) inter-individual variation in Vtg concentrations, due to some fish responding and others not (e.g. Allen et al., 1999a). This causes difficulties in statistical analysis due to the large heterogeneity of variances between sites, which could be due to inter-individual differences in genotype, size, migration, or prey selection. This variability and lack of response in some fish emphasises the requirement to reduce the biological variability as much as possible through careful design of the survey and sampling strategies and the need for the analytical technique to be able determine basal Vtg concentrations. Many of the requirements of a Vtg monitoring survey are the same as are included in the JAMP Guidelines for other CEMP determinants, as can be seen in Table 1.

Sample Collection

35. As noted above, Vtg is present in plasma as a dimer and is naturally cleaved in the oocytes to form individual yolk proteins. Because of this, Vtg is relatively unstable and blood samples must be collected and handled in a rigorously standardised manner to avoid damage to the protein. Blood should be collected into heparinised tubes, stored immediately on ice and centrifuged within 30 minutes. Protease inhibitors must be added to the plasma, which should be divided into two sub-samples, snap frozen in liquid nitrogen and the two sub-samples stored in separate cryogenic freezers or liquid nitrogen cryostores. Samples must not be allowed to thaw prior to assay.

Determination of Vtg

36. The expression of Vtg genes can be measured either as mRNA transcripts or as protein, and either endpoint can be estimated semi-quantitatively, or accurately quantified. The induction of Vtg mRNA transcription is measurable within a few hours of exposure, and decays rapidly (t½ = 3-4 days) once the exposure ceases. In comparison, plasma Vtg titres take several days to reach measurable concentrations and decay with a half-life of days to weeks (Craft et al., 2004), such that males may take several months to clear Vtg from their bloodstream. A low plasma Vtg concentration may thus indicate a recent small exposure or a larger exposure some time in the past.

Determination of hepatic Vtg mRNA

37. The induction of hepatic Vtg mRNA can be measured either semi-quantitatively through reverse transcription polymerase chain reaction (RT-PCR) or quantitatively with real-time PCR (Q-PCR). In both methods, a cDNA probe is hybridised with mRNA from samples and standards and then amplified through PCR. Measurement of mRNA requires specialised equipment and access to Vtg cDNA probes for each species, and although internal QC methods maybe used, there is currently no external QA scheme available. Therefore, Vtg mRNA cannot be currently recommended for use in OSPAR integrated monitoring.

Determination of plasma Vtg concentrations

38. Vtg protein titres in blood plasma can be indirectly estimated by determining total protein, Ca or P concentrations. However, these are not specific measures of Vtg, are not sufficiently sensitive to quantify basal Vtg concentrations and are thus not suitable for use in environmental monitoring of oestrogenic exposure.

39. Recent advances in proteomic and chemical analysis have led to instrumental means of assessing Vtg. However, these techniques involve specialised instrumentation (e.g. liquid chromatography combined with tandem mass spectrometry) that is not readily available to most monitoring laboratories and therefore immunoassays are a much more commonly used technique for Vtg determination.

40. Since Vtg is highly antigenic, quantitative methods of determination based upon immunoassay are relatively straightforward to develop, and many are described in the scientific literature (e.g. Mañanos et al., 1994; Allen et al., 1999a; Fujiwara et al., 2005). Specker and Anderson (1994) and Hiramatsu et al. (2005)
detail how to develop each of the different types of Enzyme Linked ImmunoSorbent Assay (ELISA) procedures that are available. Some types are more sensitive assays than others are, and ICES have published guidelines on the use of Vtg determinations in biological effects monitoring programmes (Scott and Hyland, 2002). Immunoassays are very sensitive and may have a high specificity. The specificity is ultimately controlled by the purity of the Vtg used for radiolabelling (in radioimmunoassays), for plate-coating (in the case of competitive ELISAs), or for raising the antibodies (in the case of sandwich ELISAs).

41. To increase the stability of Vtg, it is usual to add protease inhibitors (e.g. aprotinin or PMSF) and to keep it cold during all stages of handling. Apparent success in the purification of Vtg (e.g. as a single band on an electrophoresis gel) may be due to these precautionary measures, but can also be due to some species having more robust Vtg than others (Silversand et al., 1993; Henries et al., 2003). Cod Vtg appears to be particularly difficult to prepare in an intact form (Silversand et al., 1993; Arukwe and Goksøyr, 2003), leading to the use of Lv as the standard antigen in immunoassays (Meier et al., 2002; Scott et al., 2006). Vtg from some other species (e.g. flounder) have proved to be relatively robust proteins during the processes of purification and storage and thus much easier to develop into immunoassays for monitoring purposes (e.g. Matthiessen et al., 1998). Validated Vtg RIA and ELISA assays tend to have inter- and intra-assay variations (5-20%) that are insignificant in comparison with the concentrations in exposed fish (e.g. Matthiessen et al., 1998; Parks et al., 1999; Mosconi et al., 2002; Eidem et al., 2005).

42. To date, it has been unclear in most fish Vtg studies whether the assay has been measuring VtgA, VtgB, VtgC (see above), or a mixture of all three. In only three species (tilapia, Japanese goby and medaka) have assays been developed for the different forms (Takemura and Kim, 2001; Ohkubo et al., 2003; Fujiwara et al., 2005). In the tilapia and Japanese goby, VtgA responds to lower doses of E2 and reaches higher levels than VtgC. Similarly, in medaka, VtgA has been shown to reach ten times higher levels in the plasma in response to E2 than VtgB. These differences mean that if two laboratories use antibodies and antigens that have slightly different VtgA:VtgB:VtgC ratios then they could produce different results for a common sample. It is difficult and expensive to purify a specific Vtg and therefore it is recommended to use a common source of Vtg antibody and antigen in Vtg surveys, rather than specify that the assay to be used should be directed against, for example, VtgA. In this case, it will not matter what the precise mix of Vtg proteins present is as it will be the same in all testing laboratories. The detection limit, inter-assay and intra-assay variation should be quantified and inter-laboratory comparisons made through ring-tests, e.g. external quality assurance schemes such as BEQUALM (see below).

43. Analytical ELISA kits (including antibodies, standards and reagents) for the determination of Vtg are available for several fish species, including salmon, medaka, zebrafish, fathead minnow, carp, rainbow trout, and cod (Biosense, Norway; Amersham Biosciences, UK) and are straightforward for a suitably equipped laboratory to employ. The commercially available cod assay attempts to measure intact Vtg. However, this is particularly unstable in cod and for this reason an ELISA based upon Lv is recommended. Unfortunately, ELISA kits are not currently commercially available for flatfish species used in routine OSPAR monitoring programmes. If Vtg is added to the list of determinants required by the OSPAR CEMP, then this may create sufficient demand for commercial suppliers to provide a flatfish Vtg ELISA. Additionally, for the technical reasons given above, all of the participants in a monitoring programme should use a common source of Vtg antibody and antigen.

**Supporting Parameters**

44. In order to aid interpretation of the data produced, the collection of data for a number of supporting parameters and co-factors is recommended. Each fish must be given a unique identifier and its sampling location (site name, site code, latitude and longitude) recorded; any grossly visible anomalies and lesions should be noted. Biological variability will be reduced by sampling plasma from only male fish of a specified size range and at a specified time of year. However, the length, weight and gonad weight of each individual fish should be recorded. Recording gonad weight allows the calculation of gonadosomatic index (gonad weight as a percentage of body weight), an indication of the sexual maturity of the fish. A gonad should also be preserved to allow histological examination, both to confirm sex in the event of a high Vtg titre, and in order to determine the incidence and severity of intersex (presence of oocytes in the testes). Systems to describe the severity of intersex have been developed for flounder (Bateman et al., 2004) and roach (Bjerregaard et al., 2005), and can be adapted for use with other species, such as dab (Stentiford and Feist, 2005). Otolith samples should also be taken in order to age each fish. Bile samples could be collected to allow bioassay-directed fractionation and identification of any oestrogenic metabolites present. If mRNA transcripts are to be quantitated, a liver sample should be preserved in liquid nitrogen. As part of an
integrated ED monitoring programme, additional samples of liver and muscle would be required for chemical analysis.

45. Bioassay-directed fractionation and identification of oestrogenic metabolites involves chemical extraction and chromatographic fractionation of the extracts coupled to the use of in vitro assays, such as the yeast oestrogen screen (YES), to identify any oestrogenically active fractions of the extracts. The active compounds can be further separated and subsequently identified using mass spectrometry. Following the findings of male Vtg and intersex in male freshwater fish in UK rivers (e.g. Purdom et al., 1994; Harries et al., 1996, 1997), this technique was used to identify the major oestrogenic compounds in sewage treatment work effluents (Desbrow et al., 1998). It can also be applied to sediment extracts (Thomas et al., 2004) in order to identify oestrogenic compounds present in the environment, or to bile samples to identify bioavailable oestrogenic compounds.

**Applicability across the Ospar Maritime Area**

**Geographical Considerations**

46. The species recommended for use in assessing the occurrence and effects of xenoestrogenic endocrine disruptors (flounder and cod) are already recognised as being suitable for OSPAR-wide contaminant and biological effects monitoring. This is due to their widespread distribution and life habits. Flounder are recommended as the sentinel species for estuarine environments as they are known to occur throughout the OSPAR region and are responsive to oestrogenic exposure. However, they are not continually resident in estuaries as they migrate offshore to breed and, in order to sample following the maximum period of potential exposure, any survey should take place prior to the emigration. It is not known, however, whether all the adult fish of a population migrate, nor is the exact timing of the emigration known for all populations. Cod are widely distributed within the OSPAR region and are the recommended species for monitoring xenoestrogen contamination in offshore areas. Dab may also be suitable for offshore monitoring.

**Technical Considerations**

47. Validation of a recognised Vtg ELISA, for which the protocol, antigen and antibodies are available, is relatively straightforward for a laboratory with ecotoxicological experience. The only equipment required is a centrifuge, 96-well plate reader and cryogenic storage facilities. An automated 96-well plate washer is advantageous as it makes the procedure less laborious. The laboratory will need to demonstrate parallel dilution curves for plasma samples and standards, determine robust limits of detection and limits of quantification, and develop internal quality assurance procedures, using control charts to demonstrate statistical control of the assay. This may involve the analysis of a laboratory reference material on each 96-well plate, or the reanalysis of a sample from the previously analysed batch. To develop a laboratory reference material, one or more fish are induced to produce Vtg by repeated weekly injection with oestradiol (e.g. three injections of 5 mg/kg E2, dissolved in a carrier such as corn oil or squalene). After a further week, blood is collected into heparinised tubes, protease inhibitors added and the samples centrifuged to obtain plasma. The plasma samples are thoroughly mixed together and homogenous aliquots frozen. Long-term storage requires an alarmed cryogenic freezer as a minimum; storage in liquid nitrogen is preferred since it is colder and less liable to failure. Each aliquot of LRM plasma is assayed once only.

48. Should a common source of antibody and antigen not be available, each laboratory will need to develop its own antibody and standard antigen. As already noted, this is not recommended, as it reduces the reliability of simple comparison between results from different laboratories. However, many Vtg assays are described in the literature (e.g. Mañanós et al., 1994; Allen et al., 1999a; Fujiwara et al., 2005). Briefly, the procedure requires purified Vtg to be obtained for the species of interest, the raising of an antibody to the Vtg, demonstration of its specificity for Vtg and full validation of the assay (including external QA) as described above and detailed elsewhere (e.g. Specker and Anderson, 1994; Hiramatsu et al., 2005). Purification of Vtg requires specialised chromatographic instrumentation and a degree of experience in its use, which may restrict the numbers of laboratories able to develop their own Vtg assay.

**External QA**

49. There are currently no Certified Reference Materials available for Vtg in male fish plasma, although an external QA scheme is available to allow inter-laboratory comparisons to be made. This developed from an
EU-funded project (Biological Effects Quality Assurance In Monitoring Programmes - BEQUALM) that ran from 1998 to 2002 and has since become a self-funding Quality Assurance scheme for biological effects monitoring, similar to the QUASIMEME scheme for chemical analysis. The determination of plasma Vtg concentrations is included in this scheme, with the lead laboratory being the Norwegian Institute for Water Research (NIVA). The first Vtg ring-test under the self-funding programme took place in 2004-2005, with five laboratories returning data to NIVA. Each laboratory received five plasma samples from Atlantic cod that were pooled from E2-treated fish, control fish, or a combination of induced and control. One of the five laboratories returned semi-quantitative data. All five laboratories were able to distinguish the three samples with induced Vtg from the two with low Vtg. The quantitative data showed good agreement in the concentrations determined for each of the three induced samples, with greater inter-laboratory variability for the two uninduced samples, in which the Vtg concentrations were close to the limit of quantification. The five laboratories produced comparable results, and all would have been able to detect the effects of xeno-estrogens on Vtg in Atlantic cod (NIVA, 2005). In order for Vtg determinations to be compatible with JAMP principals, the BEQUALM scheme should be expanded to include other fish species that are used in CEMP monitoring (e.g. flounder, dab) and a larger number of laboratories should be encouraged to participate.

Assessment Criteria

50. At the 2004 Joint ICES/OSPAR Workshop on the Evaluation and Update of Background Reference Concentrations (B/RCs) and Ecotoxicological Assessment Criteria (EACs; Moffat et al., 2004), it was agreed to replace the use of Background/Reference Concentrations (BRCs) and Ecotoxicological Assessment Criteria (EACs) with Background Concentrations (BCs) and Environmental Assessment Criteria (EACs). BCs are concentrations typical of remote areas, or (for contaminant data) from within sediment cores at depths deposited prior to industrialisation. The Workshop also proposed that testing of whether mean observed concentrations are near background concentrations should be carried out using “Background Assessment Concentrations” (BACs). These are statistical tools defined in relation to the background concentrations (BCs) on the basis of the variability within the monitoring dataset and which enable testing of whether observed mean concentrations can be considered to be near to background concentrations. EACs are the concentrations below which there should be no deleterious effects on the exposed individual, population or community.

51. Applied to Vtg, these definitions imply that the induction of Vtg in male fish is either a marker of oestrogenic exposure (BC) or a marker of harmful effect (EAC). Male Vtg induction is recognised as an indicator of oestrogenic exposure, but there is currently no scientific consensus that it is also an indicator of harmful effect. In laboratory studies on some fish species, Vtg tends to be more sensitive to oestrogenic exposure than gross reproductive endpoints such as fecundity and fertility (e.g. Kang et al., 2002; Brion et al., 2004; Tilton et al., 2005), whilst in other reports these reproducive endpoints have been shown to be more sensitive than Vtg (e.g. Gronen et al., 1999; Cheek et al., 2001). Few reproductive studies have used marine fish, and none have used OSPAR sentinel species.

52. Field studies linking Vtg with reproductive effects are very rare. However, very recent unpublished work from North America (Johnson et al., presented at the SETAC North America 26th Annual Meeting in Baltimore, MD, USA, November 2005) showed that at sites where males had elevated plasma Vtg, female English sole began vitellogenesis early and matured more slowly compared to fish from clean sites. These results would appear to confirm earlier studies on flounders (Janssen et al., 1997) that showed that female fish experimentally exposed to contaminated harbour sediment for three years showed premature vitellogenesis and elevated plasma Vtg levels (although males in this case did not have elevated Vtg levels).

Environmental Assessment Criteria

53. To set an EAC it needs to be shown that plasma Vtg above a certain concentration has a deleterious effect on individuals, populations or communities. In individual fish, very high plasma Vtg concentrations are associated with kidney failure and poor Ca homeostasis (Herman and Kincaid, 1988), reduced reproductive success and histologically abnormal gonads (intersex). The correlation between plasma Vtg concentration and intersex is weak, since the timing of exposure is critical. Intersex is believed to be an irreversible condition and can be consequent to male fish having been exposed to oestrogens at early life stages, during gonadal differentiation, whereas Vtg production can be induced at any age and concentrations in the plasma decrease following cessation of exposure. A fish may have been exposed during gonadogenesis and be intersex, but it may not show Vtg induction if the exposure ceased several weeks before sampling.
54. A number of studies have examined the relationship between Vtg and reproductive success. However, model aquarium species used for life-cycle testing (e.g. Japanese medaka, zebrafish, fathead minnow) are small in size and blood plasma samples are difficult to obtain from them. Consequently, there are few studies correlating absolute plasma Vtg concentrations with reproductive output. Those that do have tended to use fathead minnow (FHM, *Pimephales promelas*), which is larger than the other species (e.g. Kramer *et al.*, 1998; Harries *et al.*, 2000; Länge *et al.*, 2001). Reproductive endpoints were affected when male FHM plasma Vtg concentrations exceeded approximately 100 μg/l (Harries *et al.*, 2000) or 10 μg/l (Länge *et al.*, 2001). However, the relative sensitivities of the Vtg and reproductive endpoints, and the strength of the correlation between Vtg and reproductive effects, depend upon many factors. These include the particular compound(s) tested, the timing and duration of the exposure in relation to the age of the fish and its reproductive status, the time of sampling in relation to both the time of exposure and the reproductive status, and upon the species being investigated (e.g. Segner *et al.*, 2003; Brion *et al.*, 2004).

55. Because there are no data available linking plasma Vtg and reproduction in CEMP-recommended sentinel organisms, and because several of the factors listed above are variable in the marine environment, it is unlikely that it will be possible to set a reliable EAC for plasma Vtg concentration.

**Background Concentrations**

56. Due to the lack of agreement over whether Vtg is a marker of effect or not, it should be, in theory, easier to establish a BC for plasma Vtg concentrations, than to set an EAC. There are a limited amount of environmental and laboratory data available for plasma Vtg concentrations in flounder and cod (Tables 2 and 3), suggesting that setting a BC should be possible. Vtg data for dab have not yet been submitted for peer review, although a study has recently been completed on dab from offshore UK waters (Scott *et al.*, in prep). A variety of factors including inter-specific, seasonal and size differences in Vtg concentrations (Hiramatsu *et al.*, 2005), plus the presence of phytoestrogens in commercial diets often used in laboratory studies (Pelissero *et al.*, 1989; Matsumoto *et al.*, 2004) and a paucity of data in marine fish, currently make it difficult to set a firm BC for the entire OSPAR-region. However, sufficient data exist from studies in the UK to set a provisional BC for Vtg in male flounder, although this figure should be reviewed once additional data are obtained from field surveys that fulfil the requirements noted above. There is one field experiment (BECPELAG) from which data can be used to suggest a provisional BC for cod.

57. UK flounder data demonstrates variability within sites due to non-standardised sampling months, fish sizes, etc. Nevertheless, a provisional BC of 0.13 μg/ml can be set, based upon the 90th percentile of all of the male Vtg concentrations (range = <0.01 – 0.17 μg/ml, n=95) in fish collected at a UK reference estuary (River Alde) between 1996 and 2001. In the BECPELAG study (Scott *et al.*, in press), Vtg concentrations in caged male cod from reference sites in the North Sea were determined to be in the range <0.01-1.35 μg/ml (n=69). Based upon the 90th percentile of these data, a provisional BC of 0.22 g/ml is proposed for cod.

**Concluding Comments**

- In order to assess the occurrence, distribution, bioavailability and effects of oestrogenic endocrine disrupting chemicals in the OSPAR region, it is recommended that the determination of the egg yolk protein vitellogenin (Vtg) in the blood plasma of male fish be added to the list of Contaminant-specific Biological Effects included in the OSPAR CEMP.
- The recommended sentinel species are flounder (estuaries) and cod or dab (offshore), which should be sampled between January and March, depending on location.
- In addition to plasma samples, gonad samples should be sampled and histologically examined to allow the presence and severity of intersex to be determined.
- Other sampling requirements are covered by existing JAMP Guidelines.
- The recommended analytical method is Enzyme-Linked ImmunoSorbent Assay (ELISA), using purified Vtg as standard for flounder and dab, and lipovitellin (Lv) for cod.
- A common source of Vtg antibody and antigen should be made available to the different Contracting Parties for the flounder and dab assays, and a common source of Lv antibody and antigen for the cod assay.
- The existing BEQUALM Quality Assurance scheme includes the determination of Vtg in cod, and should be expanded to included flounder and dab.
• Existing data from the UK allows a provisional Background Concentration for Vtg in male flounder plasma to be set at 0.13 µg/ml. The limited Vtg data for cod suggest a provisional BC of 0.22 µg/ml for this species. These values should be reviewed as more data becomes available. Background Assessment Concentrations should be developed and used as the criteria against which to assess whether oestrogenic exposure is occurring.

• Data in the scientific literature indicates that the suggested provisional BC is exceeded in flounder from several UK estuaries and some areas of the southern North Sea. This indicates that biologically significant concentrations of oestrogenic endocrine disruptors are present in these areas.

• In areas where male fish exhibit elevated Vtg, the identity of the compounds responsible should be established by bioassay-directed fractionation and mass spectrometry of bile and/or sediment extracts.

References


Madsen, L.L., Korsgaard, B. and Bjerregaard, P. 2003. Estrogenic effects in flounder Platichthys flesus orally exposed to 4-tert-octylphenol. Aquatic Toxicology, 64, 393-405.


Moffat, C.F., Pijnenburg, J. and Trass, T. (eds.). 2004. OSiPAR/ICES Workshop on the Evaluation and Update of Background Reference Concentrations (B/RCs) and Ecotoxicological Assessment Criteria (EACs) and How These Assessment Tools should be Used in Assessing Contaminants in Water,


Scott, A.P., Nagahama, Y., van den Kraak, G. and Nagler, J.J. 1995. Sulfation and uptake of the maturation-inducing steroid, 17α,20β-dihydroxy-4-pregnen-3-one by rainbow trout ovarian follicles. Fish Physiology and Biochemistry, 14, 301-311.


Scott, A.P. and Hylland, K. 2002. Biological effects of contaminants: Radioimmunoassay (RIA) and enzyme-linked immunosorbent assay (ELISA) techniques for the measurement of marine fish vitellogenins. ICES Techniques in Marine Environmental Sciences, no 31. 21pp.

Scott, A.P., Kristiansen, S.I., Katsiadaki, I., Thain, J., Tollefsen, K.E., Gokseyr, A. and Barry, J.  In Press. Assessment of estrogen exposure in cod (Gadus morhua) and saithe (Pollachius virens) in relation to their proximity to an oil field. In *Biological Effects of Contaminants in the Pelagic ecosystem (BECPELAG)* (Hylland, K., eds), Society of Environmental Toxicology and Chemistry.


### Table 1: Requirements for Vtg monitoring, in comparison with existing JAMP recommendations

<table>
<thead>
<tr>
<th></th>
<th>Existing JAMP Recommendation</th>
<th>Recommendation for xenoestrogen assessment</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Species and size</strong></td>
<td>Cod, 30-45 cm, Flounder, 15-35 cm, dab, 20-25 cm</td>
<td>Cod, 30-45 cm Flounder 15-35 cm</td>
<td>Cod are suitable for offshore areas, but Vtg is related to size of fish so sampled animals must be of a proscribed size range. Flounder are suitable for estuarine areas Dab may be suitable, but no data are currently available.</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td>Single, either</td>
<td>Male</td>
<td>Females have naturally high Vtg concentrations that are too variable for them to be useful in monitoring (anti-) oestrogens in the environment.</td>
</tr>
<tr>
<td><strong>Time of year to sample</strong></td>
<td>Out-with the breeding season</td>
<td>Flounder: January/February Cod: out-with the breeding season.</td>
<td>Prior to flounder emigration from estuaries and after they are resident for maximum time period.</td>
</tr>
<tr>
<td><strong>Numbers required</strong></td>
<td>At least 12</td>
<td>15-20</td>
<td>Data are usually variable, lacking normality and have uneven variance, making parametric statistical analysis of means problematic.</td>
</tr>
<tr>
<td><strong>Supporting parameters required</strong></td>
<td>Fish identifier; Site code and name; Taxonomic identification; Temperature at collection site; Date of sample collection; Gender; Location and description of grossly visible anomalies and lesions; Hepatosomatic index; Gonadosomatic index; Fish weight and length;</td>
<td>Fish identifier; Site code and name; Taxonomic identification; Temperature at collection site; Date of sample collection; Gender; Location and description of grossly visible anomalies and lesions; Gonadosomatic index; Fish weight and length; Age of the fish; Histological examination of gonad</td>
<td>Histological examination of the gonad is useful to confirm the sex, if high Vtg is found. Intersex incidence and severity can provide information on degree of exposure to xenoestrogens in early life stages. Bile sample can be used in bioassay-directed chemical fractionation and identification of xenoestrogen metabolites. Chemical analysis will give information on known endocrine disruptors, e.g. POPs.</td>
</tr>
</tbody>
</table>
**Age of the fish**
- confirm sex and determine intersex occurrence and severity;
- Bile sample for chemical analysis;
- Liver and flesh samples for chemical analysis
- Liver sample for mRNA analysis

**Sampling requirements**
- Staff must be trained to collect blood samples
- Blood sampling must be carried out as rapidly as possible
- Blood must be stored immediately on ice and centrifuged within 30 min
- Plasma should be snap frozen in liquid nitrogen
- Samples must not thaw prior to assay.

**Preferred analytical method**
- **Flounder – Vtg ELISA**
- **Cod – lipovitellin (Lv) ELISA**

**Assay requirements**
- The antibody must be specific and not cross-react with other plasma proteins;
- The detection limit must be below 10 ng/ml

Collection of liver samples for transcript analysis can provide additional information on the status of the animal.

Vtg is a fragile protein, and erroneous results can potentially be generated by the way that the samples are handled. It is important that there is a standard operating procedure (SOP) for collection of blood samples. Repeated freezing and thawing will cause degradation of the sample.

Plasma samples can be split prior to freezing and then stored in separate freezers to reduce the risk of damage.

Flounder Vtg is relatively stable and thus the standard relatively reliable and easy to handle.

Cod Vtg is very unstable, Lv is a constituent protein of Vtg and can be used as a surrogate for it.

For each species, a standard ELISA protocol and a common source of antigen and antibody are required.

Plasma concentrations in uninduced fish are very low. To aid subsequent statistical analysis it is important that these concentrations are quantifiable.
Table 2: Plasma vitellogenin concentrations in male flounder from UK “reference” sites between 1996 and 2001, including the overall 90th percentile concentration from all “reference” sites used to set a Background Concentration. Also included are concentrations from the Mersey estuary for comparison.

<table>
<thead>
<tr>
<th>Area</th>
<th>Concentration (µg/ml)</th>
<th>n</th>
<th>Range of values</th>
</tr>
</thead>
<tbody>
<tr>
<td>North Sea</td>
<td>0.05 ± 0.01</td>
<td>18</td>
<td>&lt;0.02 – 0.12</td>
</tr>
<tr>
<td>River Tyne</td>
<td>0.04 ± 0.01</td>
<td>18</td>
<td>&lt;0.01 – 0.05</td>
</tr>
<tr>
<td>River Thames site1</td>
<td>0.03 ± 0.003</td>
<td>15</td>
<td>&lt;0.01 – 0.48</td>
</tr>
<tr>
<td>River Thames site2</td>
<td>0.03 ± 0.004</td>
<td>15</td>
<td>&lt;0.01 – 0.59</td>
</tr>
<tr>
<td>Laboratory fish1</td>
<td>0.02 ± 0.007</td>
<td>15</td>
<td>&lt;0.01 – 0.13</td>
</tr>
<tr>
<td>Pillar Bank – Clyde2</td>
<td>0.02 ± 0.002</td>
<td>19</td>
<td>&lt;0.01 – 0.04</td>
</tr>
<tr>
<td>Bowling – Clyde2</td>
<td>0.08 ± 0.02</td>
<td>10</td>
<td>&lt;0.01 – 0.24</td>
</tr>
<tr>
<td>Petty Roy – Clyde2</td>
<td>0.26 ± 0.22</td>
<td>20</td>
<td>&lt;0.01 – 4.69</td>
</tr>
<tr>
<td>Alde reference site 1996-2001</td>
<td>0.26 ± 0.22</td>
<td>95</td>
<td>&lt;0.01 – 0.17</td>
</tr>
<tr>
<td>Mersey estuary Dec 19961</td>
<td>42,623 ± 15,607</td>
<td>25</td>
<td></td>
</tr>
</tbody>
</table>

References: 1Allen et al., 1999a; 2Allen et al., 1999b; 3Kirby et al., 2004a
Table 3: Plasma vitellogenin concentrations in male cod between 2001 and 2003: concentrations in fish from reference sites and from induced fish

<table>
<thead>
<tr>
<th>Area</th>
<th>Concentration (µg/ml)</th>
<th>n</th>
<th>Range of values</th>
</tr>
</thead>
<tbody>
<tr>
<td>**BECPELAG study (North Sea)**³</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males caged away from oil rig</td>
<td>0.04 ± 0.01</td>
<td>11</td>
<td>&lt;0.01 - 0.13</td>
</tr>
<tr>
<td>Males caged near oil rig</td>
<td><strong>0.24 ± 0.08</strong></td>
<td>13</td>
<td>0.02 - 0.85</td>
</tr>
<tr>
<td>**Norwegian cod⁴</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males from two northern fjords</td>
<td>0.09 ± 0.01</td>
<td>22</td>
<td>0.04 – 0.25</td>
</tr>
<tr>
<td>Males from Oslofjord</td>
<td><strong>3.45 ± 1.15</strong></td>
<td>13</td>
<td>0.04 – 11.5</td>
</tr>
<tr>
<td>Captive males (Matre)</td>
<td><strong>5.56 ± 2.67</strong></td>
<td>22</td>
<td>0.06 – 46.5</td>
</tr>
<tr>
<td><strong>Wild-caught cod in NE Atlantic⁴</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fish &lt; 7 kg</td>
<td>0.17 ± 0.03</td>
<td>233</td>
<td>&lt;0.01 – 4.04</td>
</tr>
<tr>
<td>Fish &lt; 7 kg</td>
<td></td>
<td></td>
<td>(15% &lt; assay detection limit of 0.01)</td>
</tr>
<tr>
<td>Fish &gt; 7 kg</td>
<td><strong>13.14 ± 3.18</strong></td>
<td>84</td>
<td>&lt;0.01 – 159.00</td>
</tr>
<tr>
<td>Fish &gt; 7 kg</td>
<td></td>
<td></td>
<td>(2% &lt; assay detection limit of 0.01)</td>
</tr>
<tr>
<td><strong>Experimentally induced⁵</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E₂-implanted (after 28 days)</td>
<td>ca 40,000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Values in bold are statistically significantly higher than the other values within the same sub-set.

References: ³Scott et al., in press; ⁴Scott et al., 2006; ⁵Scott et al., 2002
Chapter 14 Scope For Growth

Background

1. Growth provides one of the most sensitive measures of stress in an organism because growth integrates major physiological responses, specifically the balance between processes of energy acquisition (feeding and digestion) and energy expenditure (metabolism and excretion). Each physiological response can be readily determined in bivalves, converted into energy equivalents and alterations in the energy available for growth and reproduction (Scope For Growth - SFG).

2. SFG when used as an integrated approach to monitoring is based on the combined measurement of SFG and chemical contaminants in mussels, and has been used successfully to detect, quantify, and identify the potential causes of pollution in estuaries and bays (typically over small spatial scales of ca. 10 km; reviewed by Widdows and Donkin, 1992), as well as over larger spatial scales of >1000 km of North Sea coastline (Widdows et al., 1995a) and the Irish Sea coastline (Widdows et al., 2002). Furthermore, SFG has been applied over a wide range of latitudes from subtropical (Bermuda; Widdows et al., 1990) to Subarctic (Iceland; Halldórsson et al., 2005).

3. Not only has SFG been correlated with concentrations of toxic contaminants in the tissues of mussels, but recent studies have also demonstrated that SFG correlates with measures of biodiversity in the benthic community (Crowe et al., 2004). Therefore SFG can provide an effective indicator of pollution effects at the individual and at the community levels.

4. The methodology for determining the SFG of mussels (and other bivalve species) has been applied routinely to both toxicant-exposed mussels in laboratory studies and mussels collected from polluted environments. Laboratory studies are used to establish relationships between toxicant tissue concentrations and SFG, which can then be used to provide a quantitative toxicological interpretation of SFG and tissue contaminant levels in field monitoring programmes (Widdows and Donkin, 1992; Widdows et al., 1995a, b, 1997, 2002).

5. The underlying objective of both field and laboratory studies is to maintain and measure the SFG of individual mussels under "near optimal" conditions, so that the SFG will be maximized at a given ration level, and any reduction in SFG will reflect the stress induced by the toxicants accumulated in their body tissues.

6. The methodology for measuring SFG is well documented in the ICES TIMES series (no 40). By standardizing the ration level in SFG measurements, the food absorption efficiency remains relatively constant, and food is removed as a key variable, so allowing SFG to reflect the underlying impact of the total toxicant load accumulated within the body tissues (Widdows and Johnson, 1988; Widdows et al., 1990, 1995a, b, 1997, 2002). Although this SFG measurement does not predict the actual growth in the field, because food availability in the coastal environment is temporally and spatially variable and difficult to measure routinely, it does reflect the overall growth potential for individuals and mussel populations. For example, mussels from the Liverpool Bay and Morecombe Bay region of the Irish Sea had the lowest SFG values, and this was consistent with very low growth rates (an order of magnitude lower than unpolluted areas; Widdows et al., 2002). Subsequent studies by Crowe et al. (2004) have demonstrated a lower biodiversity within the mussel bed community at the Irish Sea study sites with low mussel SFG.

7. In addition, more detailed chemical analyses of the mussel tissues have confirmed that the lowered SFG and biodiversity values were correlated with increased concentrations of hydrocarbons accumulated in the mussel tissues, particularly those associated with the "unresolved complex mixture" (Crowe et al., 2004). Recent studies by Donkin et al. (2003) and A. Booth et al. (unpublished data) have begun to identify these previously unresolved compounds and shown them to be toxic to mussels.

8. This demonstrates that SFG is able to detect and quantify pollution impact and that subsequent independent studies were able to analyse and identify the nature of the toxicants in more detail and show changes at the population and community level.
Confounding factors

9. The preferred season for measuring SFG of field-collected mussels is during the period of maximum growth potential (i.e. from early summer to early autumn). It is important to avoid measurement of SFG or any other cellular/biochemical response during the spawning period, the timing of which is variable depending on latitude and seasonal temperature regime (but generally late winter/spring). The process of sampling and transportation of mussels at this time increases the probability of inducing the release of gametes (particularly following a prolonged period of air exposure or a temperature/physical shock), and this naturally stresses the animal. The release of gametes makes it very difficult to perform SFG measurements successfully. In addition, it is advisable not to measure SFG in autumn during a period of natural quiescence before the onset of gametogenesis.

10. Mussels chosen to represent clean reference sites in field studies, should be collected from a location that is free from significant chemical contamination (i.e. removed from local sewage inputs, urban development, and industry). Mussels collected from the mouth of most estuaries are not representative of a clean reference site. It is advisable to analyse body tissues for contaminants, particularly organics such as hydrocarbons, to confirm that the site is not significantly contaminated. (Visual assessment of the site is not sufficient.)

11. Mussels used for SFG can be collected either from native populations or from specific sites where mussels from a clean reference site have been transplanted and exposed in cages for a period of >4 w.

12. The basic physiological responses of mussels (such as feeding and respiration rate) remain relatively independent of short-term changes in natural environmental variables over a wide range of conditions; for example food/seston concentration (0.1-20 mg seston l¹; Widdows et al., 1979; Kjørboe et al., 1980), temperature (6-20°C; Widdows, 1976), and salinity (20-33; Widdows, 1985b). In addition, transplantation experiments over >1000 km have shown that any measurable differences in physiological responses and growth rates of different populations reflect environmental factors rather than genetic differences (Kautsky et al., 1990; Widdows et al., 1995a), permitting the direct comparison of mussels over a wide geographical area. This does not imply that genetically determined population differences in physiological responses do not exist, but that they are only apparent under extreme environmental conditions (e.g. elevated temperatures and reduced salinities).

Ecological Relevance

13. Scope for growth (SFG) provides an instantaneous measure of the energy status of an animal, which can range from maximum positive values under optimal conditions, declining to negative values when an animal is severely stressed and utilizing body reserves. Although direct measurements of total production and growth rate are often difficult to quantify and interpret in relation to pollution (Widdows and Donkin, 1992), SFG is rapidly determined, providing a sensitive, quantitative, and integrated response that can be related to the contaminant levels in the body tissues. Scope for growth has been applied in laboratory and mesocosm experiments to assess the toxic effects (from sublethal to lethal) of a range of environmentally important chemical contaminants, including aromatic and aliphatic hydrocarbons (Widdows et al., 1982; Donkin et al., 1989, 1991), sewage sludge (Butler et al., 1990), tri- and dibutyltin (Widdows and Page, 1993), nonylphenol (Granmo et al., 1989), pentachlorophenol (Widdows and Donkin, 1991), and organochlorine, organophosphate, and pyrethroid pesticides (Donkin et al., 1997). These laboratory studies have been particularly important in 1) establishing concentration-response relationships between the concentration of contaminants in the body tissues and the physiological responses of mussels, including SFG; 2) utilizing a quantitative structure-activity relationship (QSAR) approach to study the sublethal toxicity of organic contaminants (Donkin and Widdows, 1990). Such laboratory-derived concentration-response relationships have been used subsequently to provide a quantitative toxicological interpretation of field-derived SFG measurements and tissue residue chemistry.

14. There are many examples of the field application of SFG measurements combined with tissue residue chemistry as a means of assessing environmental pollution. These include studies of pollution gradients in Maine (Giffilan et al., 1977), Narragansett Bay (Widdows et al., 1981), San Francisco Bay
(Martin and Severeid, 1984), at the North Sea oil terminal in Sullom Voe, Shetlands (Widdows et al., 1987, 1995b), Venice Lagoon (Widdows et al., 1997), and two IOC GEEP Workshops concerned with contaminant gradients in a Norwegian fjord (Widdows and Johnson, 1988) and Bermuda (Widdows et al., 1990). In addition, two field studies by Nelson (1990) and Anderlini (1992) have used SFG to assess the impact of sewage inputs to Narragansett Bay (USA) and Wellington Harbour (New Zealand), respectively.

15. More recently, the combined measurement of SFG and chemical contaminants in mussels has been successfully extended and applied over a larger spatial scale of >1000 km of North Sea coastline (Widdows et al., 1995a) and Irish Sea coastline (Widdows et al., 2002). The main features of the approach in the North Sea and Irish Sea studies were to: 1) identify regions as well as specific sites that were significantly stressed by pollutants; 2) quantify the degree of sublethal stress and how near the animals were to the lethal limit; 3) provide a quantitative toxicological interpretation of much of the contaminant data.

16. Therefore, these various field studies have demonstrated that this approach is able to detect and quantify changes in environmental quality, as well as identify some of the cause(s) of these changes through use of QSAR relationships and established cause-effect relationships (i.e. between contaminant concentrations in mussel tissues and the SFG response).

17. SFG has also been shown to be a sensitive and ecologically meaningful biological response that can provide a powerful, rapid (i.e. results can be obtained within days of sampling), and cost-effective method for monitoring changes in environmental quality (Widdows et al., 1995a, 2002; Crowe et al., 2004).

<table>
<thead>
<tr>
<th>TOXICANT</th>
<th>REFERENCE</th>
<th>REDUCTION IN FEEDING RATE (CR) OR SFG (µg/g dry weight)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HMW Aromatics</td>
<td>Widdows and Donkin (1992)</td>
<td>&gt;2000 (CR)</td>
</tr>
<tr>
<td>Lindane</td>
<td>Donkin et al. (1997)</td>
<td>1400 (CR)</td>
</tr>
<tr>
<td>Cd</td>
<td>Poulson et al. (1982)</td>
<td>&gt;150 (SFG)</td>
</tr>
<tr>
<td>Di(ethyhexyl)phthalate</td>
<td>Donkin et al. (1996)</td>
<td>330 (CR)</td>
</tr>
<tr>
<td>LMW Alkanes</td>
<td>Donkin et al. (1989)</td>
<td>125 (CR)</td>
</tr>
<tr>
<td>LMW Aromatics</td>
<td>Donkin et al. (1989)</td>
<td>125 (CR)</td>
</tr>
<tr>
<td></td>
<td>Widdows et al. (1995a)</td>
<td>20 (SFG)</td>
</tr>
<tr>
<td>Carbaryl</td>
<td>Donkin et al. (1997)</td>
<td>50 (CR)</td>
</tr>
<tr>
<td>Pentachlorophenol</td>
<td>Widdows and Donkin (1991)</td>
<td>45 (SFG)</td>
</tr>
<tr>
<td>Cu</td>
<td>Widdows and Johnson (1988)</td>
<td>30 (SFG)</td>
</tr>
<tr>
<td>Tributyltin</td>
<td>Widdows and Page (1993)</td>
<td>4 (SFG)</td>
</tr>
<tr>
<td>Dichlorvos</td>
<td>Donkin et al. (1996)</td>
<td>2.2 (CR)</td>
</tr>
</tbody>
</table>

**Quality Assurance**

18. The method has been successfully tested nationally in a range of UK monitoring programmes and internationally as part of IOC Biological Effects Workshops to evaluate and compare pollution effects measurements at different levels of biological organization.
19. A successful intercalibration has been carried out across Europe to establish AQC procedures. This exercise provided the basis and mechanism for undertaking further intercalibration exercises. AQC is able to be provided through BEQUALM.

**Background Responses and Assessment Criteria**

20. Health status thresholds have yet to be defined for SFG but from the extensive data sets that exist the following values can be estimated:

- Animals may be considered to be healthy if the SFG is greater than +5 (J/h/g)
- Animals may be considered to be stressed but compensating if the SFG is between +5 and -2 (J/h/g)
- Animals may be considered as severely stressed if the SFG is less than -2(J/h/g).

It should be noted that these values are provisional and require further validation.

Background SFG responses may be as high as +20 (J/h/g).

21. The added value of SFG in mussels is that the response measures the overall impact of multiple contaminants 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).

22. Furthermore, the recent work by Crowe et al., 2004, shows that SFG correlates with measures of biodiversity in the benthic community.

**Acknowledgement**

23. The majority of the above text was provided from the ICES TIMES (2006) guideline on the measurement of Scope For Growth in blue mussels provide by John Widdows and Fred Staff (Plymouth Marine Laboratory, UK).

**References**


Bayne, B. L., Clarke, K. R., and Moore, M. N. 1981. Some practical considerations in the measurement of pollution effects on bivalve molluscs, and some possible ecological consequences. Aquatic Toxicology, 1: 159-174.


ICES TIMES 2006. International Council for the Exploration of the Sea, Techniques In Marine Environmental Sciences. Copenhagen, Denmark; Biological Effects In Contaminants, Measurement of Scope For Growth in Blue Mussels, No 40, 1-34.


Annex 6: Background document (revised) on DNA adducts

Amended version of Chapter 11 of the OSPAR Background Document on Biological Effects Monitoring Techniques.

Chapter 11 DNA adducts

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 environmental carcinogens and DNA, termed DNA adducts, can be utilized as a biological marker of exposure to genotoxic compounds. DNA adducts can be removed by cellular repair processes or by cell death, but during chronic exposures they often reach steady state concentrations in carcinogen target tissues such as the liver. As a consequence, DNA adducts have several important features which make them suitable as biomarkers of carcinogen 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, metabolizm, 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 demonstrated 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 metabolizm 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 (Ictalurus 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 \( ^{32} \text{P} \)-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 \( 10^{\text{-}10^{10}} \) undamaged nucleotides from 5-10 \( \mu \text{g} \) 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 \( ^{32} \text{P} \)-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 carcinogens (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 freshwater 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 viviparus*) 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 demonstrated that at certain contaminated sites the prevalence of DNA adducts are associated with the prevalence of toxicopathic 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 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-postlabelling 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 revealed 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 demonstrated 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-postlabelling 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 glad (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 enhance the sensitivity of the assay to detect 1 adduct in $10^6$-$10^{10}$ bases (Reichert et al., 1999). Following the adduct enrichment step, the 3’-monophosphates are radio-labelled at the 5’-hydroxyl using $^{32}$P-ATP and T4-polymerase DNA polymerase to form 3’, (32P)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 radiolabelled 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 found to be 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) 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-postlabelling 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 (Figure 1B), 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 ²³⁷P-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 ²³⁷P-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 10⁶ 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 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:response 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 herring and halibut exposed to oil (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 turbot at this oil concentration (Jonsson et al., in prep.). This value may be revised as new data to determine chronic effect levels in cod emerge.

The following issues are important and require consideration:

- ²³⁷P-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 $10^6$ normal nucleotides, which was converted to nmol adducts/mol DNA by dividing by 10.

The derived values for dab and flounder were ER-L 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 in Atlantic cod was 6 (see pt.13 above for method of estimation). This value is also indicative for flatfish (to be verified).

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 $^{32}$P-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-postlabelling 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, 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.


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.


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.


Lyons, B.P., Stentiford, G.D., Green, M., Bignell, J., Bateman, K., Feist, S.W., Goodsr, 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.


Chapter 5  PAH metabolites in bile

Compiled by Lars-Petter Myhre (IRIS-Akvamiljø) and Ketil Hylland (NIVA).

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 demonstrated 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 unmetabolized 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 metabolized 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 demonstrated 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 demonstrated 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 revealed 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 gadids, may metabolize PAHs more slowly as more will partition into fat, but this has not been documented experimentally.

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.

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

It is possible to establish global criteria for individual PAH metabolites. Baseline data for individual species may be used to test against to determine whether fish have been exposed to PAHs. As mentioned above, 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.

As outlined in Ariese et al., 2005 there are three main methods used for the quantification of PAH metabolites, fixed fluorescence/scan (FF), HPLC and GC-MS. Assessment criteria can be suggested for FF and HPLC in some marine fish species, based on observed concentrations in unpolluted areas and/or concentrations in fish kept for extended periods under controlled laboratory conditions. The suggested criteria have been developed using the 90 percentile of all available data for the species indicated.

Fixed fluorescence (microgram/g, pyrene equivalents)
- Atlantic cod (Gadus morhua): 1.1 (North Sea)
- dab (Limanda limanda): 0.15 (UK)*
- flounder (Platichthys flesus): 1.3 (Scottish reference sites)
- haddock (Melanogrammus aeglefinus): 1.9 (Barents sea)

* synchronous scan fluorescence rather than fixed fluorescence

HPLC (fluorescence detection; ng/g)
- Atlantic cod (Gadus morhua): 13 (Barents Sea)
- dab (Limanda limanda): 13 (ICON, Iceland)
- flounder (Platichthys flesus): 232 (NL)
- haddock (Melanogrammus aeglefinus): 10 (ICON; Iceland)

**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 QUASIMEME2.

**Acknowledgement**

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

**References**


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


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

<table>
<thead>
<tr>
<th>SPECIES</th>
<th>SUBSTANCE</th>
<th>TEST CONCENTRATIONS/AREA</th>
<th>EXPOSURE TIME</th>
<th>METABOLITE</th>
<th>BASELINE</th>
<th>CONTROL OR REFERENCE</th>
<th>EXPOSED /CONTROL</th>
<th>REFERENCE/COMMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cod</td>
<td>Feral fish</td>
<td>Barents Sea Baseline</td>
<td></td>
<td>Naph type</td>
<td>5.3 ug/ml</td>
<td></td>
<td></td>
<td>Aas et al., 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>0.8 ug/ml</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>0.4 ug/ml</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod</td>
<td>Feral fish</td>
<td>Egersund Baseline non polluted area</td>
<td></td>
<td>Naph type</td>
<td>6.1 ug/ml</td>
<td></td>
<td></td>
<td>Klungsøyr et al., 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>1.0 ug/ml</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>0.5 ug/ml</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod</td>
<td>Feral fish</td>
<td>Sleipner Baseline polluted area?</td>
<td></td>
<td>Naph type</td>
<td>5.9 ug/ml</td>
<td></td>
<td></td>
<td>Klungsøyr et al., 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>0.9 ug/ml</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>0.3 ug/ml</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod</td>
<td>Feral fish</td>
<td>Frey, ceased installation 10 000 m (ref) 2000 m - 200 m Baseline polluted area?</td>
<td></td>
<td>Naph type</td>
<td>3.9 ug/ml</td>
<td>1.1 - 1.1</td>
<td></td>
<td>Beyer et al. 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>0.6 ug/ml</td>
<td>1.1 - 0.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>0.3 ug/ml</td>
<td>0.9 - 0.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod</td>
<td>Feral fish</td>
<td>Barents sea Baseline</td>
<td></td>
<td>Naph type</td>
<td>2.15 ug/g</td>
<td></td>
<td></td>
<td>Sundt, 2002</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>1.63 ug/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>0.69 ug/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod</td>
<td>Feral fish</td>
<td>Barents sea Baseline</td>
<td></td>
<td>Naph type</td>
<td>5.8 ug/g</td>
<td></td>
<td></td>
<td>Aas &amp; Børseth, 2002</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>1.7 ug/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>0.8 ug/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod</td>
<td>Laboratory</td>
<td>1 ppm crude oil Statfjord B</td>
<td>14 days</td>
<td>Naph type</td>
<td>3.9 ug/g</td>
<td>7.5 - 23.7 - 31.4</td>
<td></td>
<td>Aas et al., 2002</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>2.6 ug/g</td>
<td>3.6 - 10.6 - 13</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>1.0 ug/g</td>
<td>1.7 - 2.4 - 2.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod</td>
<td>Laboratory</td>
<td>0.06 - 0.25 - 1 ppm Oil average 3, 7, 14, 24 days</td>
<td></td>
<td>Naph type</td>
<td>53.1 ug/g</td>
<td>0.7 - 2.3 - 2.9</td>
<td></td>
<td>Skadsheim et al., 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>7.0 ug/g</td>
<td>1 - 2.9 - 3.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>1.0 ug/g</td>
<td>1.1 - 1.5 - 1.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Species</td>
<td>Substance (Lab/Field)</td>
<td>Test Concentrations/Area</td>
<td>Exposure Time</td>
<td>Metabolite</td>
<td>Baseline</td>
<td>Control or Reference</td>
<td>Exposed/Control</td>
<td>Reference/Comments</td>
</tr>
<tr>
<td>------------------</td>
<td>------------------------</td>
<td>---------------------------</td>
<td>---------------</td>
<td>------------</td>
<td>----------</td>
<td>----------------------</td>
<td>-----------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Laboratory</td>
<td>Oil 0.06 - 0.25 - 1 ppm</td>
<td>30 days</td>
<td>Naph type</td>
<td>7.1 fi</td>
<td>5.1 - 9.5 - 227.5</td>
<td></td>
<td>Aas et al., 2000</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>2 fi</td>
<td>6.4 - 12.7 - 43.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>0.8 fi</td>
<td>2.3 - 3.6 - 9.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Laboratory</td>
<td>PW Oseberg, 1:1000 - 1:200 - 0.2 ppm oil - 0.2 ppm oil + PAHmix</td>
<td>15 days</td>
<td>Naph type</td>
<td>12.6 ug/ml</td>
<td>1.3 - 2.5 - 3.6 - 5.4</td>
<td></td>
<td>Sundt, 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>4 ug/ml</td>
<td>1.7 - 3.7 - 4.1 - 17.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>1.8 ug/ml</td>
<td>1.3 - 1.8 - 1.5 - 2.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Field, Caged</td>
<td>North Sea - Statfjord, 10000 m - 2000m - 500 m</td>
<td>5.5 weeks</td>
<td>Naph type</td>
<td>7.5 ug/ml</td>
<td>0.7 - 0.9 - 1.6</td>
<td></td>
<td>Aas et al., in press</td>
</tr>
<tr>
<td></td>
<td></td>
<td>German bight G4 (Ref) G1 - G2 - G3</td>
<td></td>
<td>Pyren type</td>
<td>3.1 ug/ml</td>
<td>0.8 - 0.9 - 1.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>1.2 ug/ml</td>
<td>0.8 - 1 - 1.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Field, Caged</td>
<td>North Sea - Troll, 1000 m - 500m</td>
<td>6 weeks</td>
<td>Naph type</td>
<td>4.6 ug/ml</td>
<td>0.7 - 0.8 - 1.5 - 2.5</td>
<td></td>
<td>Børseth et al., 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>2.4 ug/ml</td>
<td>1.1 - 1.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>0.9 ug/ml</td>
<td>1.1 - 1.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Field, Caged</td>
<td>North Sea - Tampen, 10000 - 2500 - 1000 - 500</td>
<td>6 weeks</td>
<td>Naph type</td>
<td>8.8 ug/ml</td>
<td>1.0 - 1.5 - 1.2 - 1.2</td>
<td></td>
<td>Hylland et al., 2005</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>1.4 ug/ml</td>
<td>0.9 - 0.7 - 0.8 - 0.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haddock (Melanogrammus aeglefinus)</td>
<td>Feral fish</td>
<td>Egersund</td>
<td>Baseline non polluted area</td>
<td>Naph type</td>
<td>5.1 ug/ml</td>
<td></td>
<td></td>
<td>Klungsøy et al., 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>1.4 ug/ml</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>0.7 ug/ml</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haddock (Melanogrammus aeglefinus)</td>
<td>Feral fish</td>
<td>Sleipner</td>
<td>Baseline polluted area?</td>
<td>Naph type</td>
<td>6.8 ug/ml</td>
<td></td>
<td></td>
<td>Klungsøy et al., 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>1.9 ug/ml</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>0.8 ug/ml</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haddock (Melanogrammus aeglefinus)</td>
<td>Feral fish</td>
<td>Statfjord</td>
<td>Baseline polluted area?</td>
<td>Naph type</td>
<td>11.2 ug/ml</td>
<td></td>
<td></td>
<td>Klungsøy et al., 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>2.5 ug/ml</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>0.7 ug/ml</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haddock (Melanogrammus aeglefinus)</td>
<td>Feral fish</td>
<td>Barents sea</td>
<td>Baseline polluted area?</td>
<td>Naph type</td>
<td>2.52 ug/g</td>
<td></td>
<td></td>
<td>Sundt, 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>1.69 ug/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>0.77 ug/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Species</td>
<td>Substance (Lab/Field)</td>
<td>Test Concentrations/Area</td>
<td>Exposure Time</td>
<td>Metabolite</td>
<td>Baseline</td>
<td>Control or Reference</td>
<td>Exposed /Control</td>
<td>Reference/Comments</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------------------</td>
<td>--------------------------</td>
<td>---------------</td>
<td>------------</td>
<td>----------</td>
<td>----------------------</td>
<td>------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Haddock (Melanogrammus aeglefinus)</td>
<td>Feral fish</td>
<td>Barents sea</td>
<td></td>
<td>Naph type</td>
<td>2.0 ug/g</td>
<td></td>
<td></td>
<td>Aas &amp; Børseth, 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>1.3 ug/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>0.6 ug/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haddock (Melanogrammus aeglefinus)</td>
<td>Feral fish</td>
<td>Frøy, ceased installation 10 000 m (ref) 2000 m - 200 m</td>
<td>Baseline polluted area?</td>
<td>Naph type</td>
<td>5.6 ug/ml</td>
<td>1.3 - 2.2</td>
<td></td>
<td>Beyer et al., 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>1.4 ug/ml</td>
<td>1.4 - 0.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>0.75 ug/ml</td>
<td>1.8 - 0.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sheepshead minnow</td>
<td>Laboratory</td>
<td>North sea oil A 0.1 - 0.4 - 0.7 ppm</td>
<td>5 weeks</td>
<td>Naph type</td>
<td>6916</td>
<td>2.3 - 6.2 - 9.3</td>
<td></td>
<td>Bechmann et al., 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>569</td>
<td>2.5 - 5 - 6.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>107</td>
<td>4 - 13.1 - 19.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sheepshead minnow</td>
<td>Laboratory</td>
<td>North sea oil B 0.1 - 0.9 - 5.6 ppm</td>
<td>6 weeks</td>
<td>Naph type</td>
<td>18164</td>
<td>1.8 - 4.3 - 12.5</td>
<td></td>
<td>Bechmann et al., 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>438</td>
<td>5.6 - 12.6 - 30.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>110</td>
<td>12.6 - 42.7 - 123.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sheepshead minnow</td>
<td>Laboratory</td>
<td>2 - 14 - 214 ppb</td>
<td>5 weeks</td>
<td>Naph type</td>
<td>267280</td>
<td>0.9 - 2.2 - 18.6</td>
<td></td>
<td>Bechmann et al., 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>9926</td>
<td>0.9 - 1.5 - 9.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>5152.7</td>
<td>3 - 17.4 - 207</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polar cod (Boreogadus saida)</td>
<td>Laboratory, feral fish 1.5 ppm StatfjA oil, baseline, control</td>
<td>14 days</td>
<td>Naph type</td>
<td>16.0 ug/g</td>
<td>2</td>
<td>16.9</td>
<td></td>
<td>Sundt &amp; Bechmann, 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren type</td>
<td>0.9 ug/g</td>
<td>5,5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>BaP type</td>
<td>0 ug/g</td>
<td>0</td>
<td>1,8</td>
<td></td>
</tr>
</tbody>
</table>
Table 2. PAH-metabolites in marine fish – measured by GC-MS.

<table>
<thead>
<tr>
<th>Species</th>
<th>Substance (Lab/Field)</th>
<th>Test Concentrations</th>
<th>Exposure Time</th>
<th>Metabolite</th>
<th>Baseline</th>
<th>Control or Reference</th>
<th>Exposed/Control</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Feral fish</td>
<td>Barents sea</td>
<td>baseline</td>
<td>Naph sum</td>
<td>150.6 ng/g</td>
<td></td>
<td></td>
<td>Aas &amp; Børseth, 2002</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>61.2 ng/g</td>
<td></td>
<td></td>
<td>Sundt, 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren</td>
<td>4.6 ng/g</td>
<td></td>
<td></td>
<td>Klungsøy et al., 2003</td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Feral fish</td>
<td>Barents sea</td>
<td>baseline</td>
<td>Naph sum</td>
<td>1285 ng/g</td>
<td></td>
<td></td>
<td>Klungsøy et al., 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>220 ng/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren</td>
<td>3.5 ng/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Feral fish</td>
<td>Egersund</td>
<td>Baseline non polluted area</td>
<td>Naph sum</td>
<td>2005.1 ng/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>230.2 ng/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren</td>
<td>3.9 ng/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Feral fish</td>
<td>Sleipner</td>
<td>Baseline polluted area?</td>
<td>Naph sum</td>
<td>1296.1 ng/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>197.8 ng/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren</td>
<td>0 ng/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Feral fish</td>
<td>Statfjord</td>
<td>Baseline polluted area?</td>
<td>Naph sum</td>
<td>1361.7 ng/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>351.1 ng/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren</td>
<td>4.0 ng/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Laboratory</td>
<td>0.06 - 0.25 - 1 ppm Oil</td>
<td>average 3, 7, 14, 24 days</td>
<td>Naph sum</td>
<td>2549 ng/g</td>
<td>4.6 - 13.4 - 23.6</td>
<td>Skadsheim et al., 2004</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>691 ng/g</td>
<td>7.7 - 22.9 - 34.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren</td>
<td>27 ng/g</td>
<td>7.3 - 16.2 - 25.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Laboratory</td>
<td>0.06 - 0.25 - 1 ppm Oil</td>
<td>average 3, 17, 31 day</td>
<td>Naph sum</td>
<td>5702 ng/g</td>
<td>4 - 13.3 - 12.7</td>
<td>Skadsheim et al., 2004</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>377 ng/g</td>
<td>10.5 - 40.3 - 48.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren</td>
<td>5 ng/g</td>
<td>8.6 - 63 - 88.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Field, Caged</td>
<td>North Sea - Statfjord, 500 - 2000 - 10000 m</td>
<td>6 weeks</td>
<td>Naph sum</td>
<td>1150 ng/g</td>
<td>3.0 - 2.0 - 1.3</td>
<td>Aas et al., in press</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>340 ng/g</td>
<td>3.5 - 2.7 - 2.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Field, Caged</td>
<td>North Sea - Troll, 1000 m - 500m</td>
<td>6 weeks</td>
<td>Naph sum</td>
<td>1515.1 ng/g</td>
<td>1.1</td>
<td>1.1 - 1.2</td>
<td>Børseth et al., 2004</td>
</tr>
<tr>
<td>Species</td>
<td>Substance (lab/field)</td>
<td>Test concentrations</td>
<td>Exposure time</td>
<td>Metabolite</td>
<td>Baseline</td>
<td>Control or reference</td>
<td>Exposed/control</td>
<td>Reference</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>--------------------------------</td>
<td>---------------------</td>
<td>---------------</td>
<td>------------</td>
<td>----------</td>
<td>---------------------</td>
<td>-----------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Field, Caged</td>
<td>North Sea -</td>
<td>6 weeks</td>
<td>Phen sum</td>
<td>327.2 ng/g</td>
<td>1,6</td>
<td>2.1 - 2.0</td>
<td>Hylland et al., 2005</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tampen, 10000 -</td>
<td></td>
<td></td>
<td>173.2 ng/g</td>
<td>1.2</td>
<td>0.9 - 1.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2500 - 1000 -</td>
<td></td>
<td>Naph sum</td>
<td>965.3 ng/g</td>
<td>0.9 - 1.7 - 0.9 - 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>500</td>
<td></td>
<td>Phen sum</td>
<td>934.5 ng/g</td>
<td>1.4 - 3 - 1.8 - 1.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren</td>
<td>3.7 ng/g</td>
<td>0 - 0 - 0.5 - 0.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Field, Caged</td>
<td>North Sea - Statfjord, 10000 m - 2000 m - 500 m</td>
<td>5.5 weeks</td>
<td>Phen sum</td>
<td>228 ng/g</td>
<td>0.2</td>
<td>0.9 - 1.1 - 0.9</td>
<td>Aas et al., in press</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Naph sum</td>
<td>228 ng/g</td>
<td>0.8</td>
<td>1 - 1 - 1.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>482 ng/g</td>
<td>0.9</td>
<td>0.7 - 0.8 - 0.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren</td>
<td>28 ng/g</td>
<td>10.2</td>
<td>29.5 - 31.1 - 41.5</td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>Field, Caged</td>
<td>German bight G4 (Ref G1 - G2 - G3)</td>
<td>5.5 weeks</td>
<td>Phen sum</td>
<td>228 ng/g</td>
<td>0.8</td>
<td>1 - 1 - 1.9</td>
<td>Aas et al., in press</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Naph sum</td>
<td>228 ng/g</td>
<td>0.8</td>
<td>1 - 1 - 1.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>482 ng/g</td>
<td>1.0</td>
<td>0.7 - 0.8 - 0.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren</td>
<td>28 ng/g</td>
<td>0.0</td>
<td>0 - 0 - 0</td>
<td></td>
</tr>
<tr>
<td>Haddock (Melanogrammus aeglefinus)</td>
<td>Feral fish</td>
<td>Egersund</td>
<td>Baseline non polluted area</td>
<td>Naph sum</td>
<td>1346.9 ng/g</td>
<td></td>
<td></td>
<td>Klungsøy et al., 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>526.8 ng/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren</td>
<td>5.7 ng/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haddock (Melanogrammus aeglefinus)</td>
<td>Feral fish</td>
<td>Sleipner</td>
<td>Baseline polluted area?</td>
<td>Naph sum</td>
<td>1111.5 ng/g</td>
<td></td>
<td></td>
<td>Klungsøy et al., 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>331.5 ng/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren</td>
<td>10.4 ng/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haddock (Melanogrammus aeglefinus)</td>
<td>Feral fish</td>
<td>Statfjord</td>
<td>Baseline polluted area?</td>
<td>Naph sum</td>
<td>1279.7 ng/g</td>
<td></td>
<td></td>
<td>Klungsøy et al., 2003</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>331.9 ng/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren</td>
<td>3.1 ng/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haddock (Melanogrammus aeglefinus)</td>
<td>Feral fish</td>
<td>Barents sea</td>
<td></td>
<td>Naph sum</td>
<td>1474 ng/g</td>
<td></td>
<td></td>
<td>Sundt, 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>165 ng/g</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polar cod (Boreogadus saida)</td>
<td>Laboratory, feral fish</td>
<td>1.5 ppm StatfjA oil, baseline, control</td>
<td>14 days</td>
<td>Naph sum</td>
<td>1330 ng/g</td>
<td>1,3</td>
<td>114</td>
<td>Sundt &amp; Bechmann, 2004</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Phen sum</td>
<td>538 ng/g</td>
<td>0,9</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Pyren</td>
<td>52 ng/g</td>
<td>14,6</td>
<td>60</td>
<td></td>
</tr>
</tbody>
</table>
Annex 8: Revised Background document on Bioassays (Chapter 7)

Chapter 7  (Water) Bioassays including proposals for assessment tools and criteria

Executive summary

Applicability across the OSPAR maritime area. In vivo and in vitro 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 specific in vitro bioassays, which can be applied to salt water, brackish water and freshwater, 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.

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 are 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 is currently complied by the Netherlands and the UK and will be submitted for publication to ICES in 2007.

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 in the case of water bioassay applications in coastal and estuarine areas, but further work and agreement are 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 (whole sediment/pore water) in the OSPAR maritime area.

**Assessment of the applicability of water 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 freshwater 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; Matthesen 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/IJC Bremerhaven Workshop on biological effects of contaminants in the North Sea and the ICES BECELAG Workshop on biological effects in pelagic ecosystems have clearly demonstrated the potential applicability of a variety of *in vitro* and *in vivo* bioassays to coastal and offshore water column and micro surface layer monitoring (Stebbing et al., 1992; Hylland et al., 2002, 2006).

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 vitro tests and in vivo bioassays 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 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 demonstrated 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, sediment elutriate bioassays and pore 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, *Skeletronema* 72-hr 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 is currently complied by the Netherlands and the UK. Following review and approval by the ICES Working Group on the Biological Effects of Contaminants (WGBEC) in March 2007 it will be published in the ICES *Techniques in Marine Environmental Sciences* series on Biological Effects of Contaminants.

Synergism between CEMP 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 recently 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 and WFD in the case of bioassay applications in coastal and estuarine areas, but that further work and agreement are needed.

Thresholds and assessment tools

General

Thresholds are currently developed and will be available soon (ICESWGBEC 2007 meeting). 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, sediment bioassays, in vivo and in vitro 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. For method to derive the background (assessment) levels of whole sediment bioassays, see background document on sediment bioassays. Pore water and elutriate bioassay can be considered with water bioassays insofar that these techniques involve a sediment manipulation procedure followed by the use of small volume water in vivo bioassays. Assessment criteria for in vitro bioassays are currently being developed and are mentioned shortly below.

Background responses and assessment criteria for (water) bioassays currently in JAMP

Water bioassays

The species recommended for water bioassays are:

- Copepod (Tisbe battaglioni and Acartia sp); 48hr exposure using mortality as the endpoint.
- Bivalves (Crassostrea gigas, Mytilus spp) embryos: 24 hr exposure using Percent Net Response as the endpoint.
- Sea urchin (Paracentrotus lividus): 24 hr 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.

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 20% for the bivalve embryo bioassay. These figures however need to be defined and further established (see also Table 3 below). Above these values would be the warning level and at 100% this would be categorized as a level of serious concern. Responses at the warning level would prompt further sampling and assay 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.

Sediment pore water and elutriates

Any of the four water bioassays listed above can be used to assess the toxicity of pore water and elutriates. The procedures used to prepare the pore and elutriates are well documented and described elsewhere. For the assessment of elutriate data, confounding factors that must be considered are: volume of sediment-water ratio, ammonia, sulphide, sediment quality – (i.e. sand or mud). For the assessment pore water data confounding factors that must be considered are: salinity, pH, DO, ammonia and sulphide. Data is produced and assessed in the same manner as for the water bioassays above and the results may be expressed in term of EC50 values and/or toxic units, depending on the purpose and the objectives of the study. Toxic elutriates and pore water can be diluted for testing if appropriate.

In vitro assays

In vitro bioassays are being developed for use with water, and sediment bioassays. In general this requires sample manipulation and/or concentration techniques, and clean-up using extraction procedures in analog to chemical compounds. 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 appro-
appropriate to develop the background responses and assessment criteria for these techniques. This needs to be progressed within the current ICES OSPAR framework.

**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 batagliai*, 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 test 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).

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

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

<table>
<thead>
<tr>
<th>Test</th>
<th>Lab</th>
<th>Average</th>
<th>0.1 PERC</th>
<th>0.5 PERC</th>
<th>0.9 PERC</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mussel embryo</td>
<td>IEOV</td>
<td>Control</td>
<td>14.1</td>
<td>12.0</td>
<td>13.5</td>
<td>16.5</td>
</tr>
<tr>
<td>Copepods</td>
<td>CEFAS</td>
<td>Control</td>
<td>1.3</td>
<td>0.0</td>
<td>0.0</td>
<td>5.0</td>
</tr>
<tr>
<td><em>Tisbe</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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 freshwater 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 demonstrate 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, metabolizm 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 allow 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 covers 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 and sediments 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 micro-scale 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 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 normative 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 – freshwater, salt water or brackish water – can be tested using the same (freshwater 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 20–100 ml taken from a discrete water sample of say 2 litres. Water extraction procedures require a larger sample volume (egg 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.

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 bioassays. 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):

- 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 PAF5 (PAF = potentially affected fraction);
- 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 freshwater, 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 & 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).

The tables/shaded text below detail the metrics for surface water.

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 for 3 acute or chronic tests from different taxonomic groups).

<table>
<thead>
<tr>
<th>ACUTE TESTS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>NE (negligible effect):</td>
<td>in 3 acute tests effect = 0 (in practice &lt; EC50), Cf = 100</td>
</tr>
<tr>
<td>MPE (maximum permissible effect):</td>
<td>in 3 acute tests effect = 0 (in practice &lt; EC50), Cf = 10</td>
</tr>
<tr>
<td>SE (serious effect): in 1 acute test effect ≥ EC50, Cf = 10 or</td>
<td>in 2 acute tests EC20&lt; effect &lt; EC50, Cf = 10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHRONIC TESTS</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>NE (negligible effect):</td>
<td>in 3 chronic tests effect = 0, Cf = 10</td>
</tr>
<tr>
<td>MPE (maximum permissible effect):</td>
<td>in 3 chronic tests effect = 0, Cf = 1</td>
</tr>
<tr>
<td>SE (serious effect):</td>
<td>in 1 chronic test effect ≥ EC50, Cf = 1 or</td>
</tr>
<tr>
<td></td>
<td>in 2 chronic tests NOEC &lt; effect &lt; EC50, Cf = 1</td>
</tr>
</tbody>
</table>

EC50 = Mean effective concentration, produces a 50% effect in the bioassay

NOEC = no-observed-effect concentration

Method: 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 3 chronic tests must be available, including at least 2 ‘whole sediment’ tests.
- As for surface water, the MPE is: in 3 chronic tests effect = 0. If a negative effect is measured in at least 1 of 3 chronic tests, the MPE is exceeded.
- The SE level is reached when an effect ≥EC₅₀ is measured in 1 chronic test (on the ‘as is’ sample), or an effect between NOEC and EC₅₀ in 2 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<EC₅₀) 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 box). 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 3 bioassays.

Table 1: Indication of toxicity in surface water of the Western Scheldt estuary on basis of 3 different bioassay responses allowing a preliminary effect assessment as proposed in Maas et al., 2003.

<table>
<thead>
<tr>
<th>LOCATION</th>
<th>DATE</th>
<th>Cf (ECF50)*</th>
<th>Cf (MTE)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Daphnia</td>
<td>Algae</td>
</tr>
<tr>
<td>SvOD-1</td>
<td>12-2-00</td>
<td>42</td>
<td>20</td>
</tr>
<tr>
<td>SvOD-2</td>
<td>9-4-00</td>
<td>28</td>
<td>16</td>
</tr>
<tr>
<td>SvOD-3</td>
<td>11-6-00</td>
<td>54</td>
<td>2.4</td>
</tr>
<tr>
<td>SvOD-4</td>
<td>2-8-00</td>
<td>56</td>
<td>3.5</td>
</tr>
<tr>
<td>SvOD-5</td>
<td>17-10-00</td>
<td>96</td>
<td>4.5</td>
</tr>
<tr>
<td>SvOD-6</td>
<td>15-12-00</td>
<td>87</td>
<td>9</td>
</tr>
<tr>
<td>SvOD-1</td>
<td>13-01-05</td>
<td>95</td>
<td>20</td>
</tr>
<tr>
<td>SvOD-2</td>
<td>9-03-05</td>
<td>87</td>
<td>30</td>
</tr>
<tr>
<td>SvOD-3</td>
<td>2-05-05</td>
<td>127</td>
<td>17</td>
</tr>
<tr>
<td>SvOD-4</td>
<td>27-6-05</td>
<td>197</td>
<td>14</td>
</tr>
<tr>
<td>SvOD-5</td>
<td>23-8-05</td>
<td>251</td>
<td>10</td>
</tr>
<tr>
<td>SvOD-6</td>
<td>19-10-05</td>
<td>94</td>
<td>12</td>
</tr>
<tr>
<td>W.Scheldt Vlissingen</td>
<td>4-6-03</td>
<td>416</td>
<td>52</td>
</tr>
<tr>
<td>W Scheldt Honte</td>
<td>4-6-03</td>
<td>180</td>
<td>56</td>
</tr>
<tr>
<td>W Scheld Termeuzen</td>
<td>4-6-03</td>
<td>403</td>
<td>28</td>
</tr>
<tr>
<td>W Scheldt Hansweert</td>
<td>2-6-03</td>
<td>243</td>
<td>16</td>
</tr>
<tr>
<td>W Scheldt Boe i s.v WO3</td>
<td>2-6-03</td>
<td>271</td>
<td>15</td>
</tr>
<tr>
<td>Scheldt Bath</td>
<td>3-6-03</td>
<td>271</td>
<td>9</td>
</tr>
<tr>
<td>Schaar vo Doel (SvoD)</td>
<td>3-6-03</td>
<td>92</td>
<td>9</td>
</tr>
<tr>
<td>Scheldt Antwerpen</td>
<td>18-6-03</td>
<td>144</td>
<td>2</td>
</tr>
</tbody>
</table>

corrected for recovery

expected chronic effect in surface water:

green = neglicable 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 di-oxins applying in other countries and observed effects in reference sediments in Dutch coastal waters.

Alternatively, using field data on DR CALUX levels in sediment (using extraction procedure with silica) from the UK and Netherlands, a provisional baseline level is
set at 10 and a serious concern level at > 40 ng Si-TEQ/kg dw. These figures are based on the following data: a UK study (Thomas and Balaam, 2002) with data from 28 sites which include 6 major UK estuaries (Humber, Southampton Water, Thames, Mersey, Tees and Tyne). Of the 28 sites, 4 were above 40 ng/kg dw. with a maximum of 88 ng/kg recorded at a known contaminated site. Values of less than 10 were recorded at 10 sites and the lowest of these values ranged between 1 and 4 (6 sites). The lowest values are regarded as being a good estimate for a background level. A Dutch study on North Sea surface sediments (Klamer et al., 2005) included DR CALUX measurements in surface sediments from major Dutch estuaries and offshore sites up to 70 miles offshore and the Oyster Grounds. From the 10 sites sampled, a range of Si-TEQ values was observed between 6 and 27 ng/kg dw. with values at 3 offshore sites below 10 ng/kg (6, 9 and 8 respectively). The ranges of TEQs in dredged sediments from rivers in the Dutch coastal zone were 12–70 ng Si-TEQ/kg dw, and on average 24 ng TEQ/kg dw (Schipper et al., 2009). Refinement of these provisional ACs will be pursued when more data will become available. 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 indicated 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 to 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 revealed 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, 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 demonstrates 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**

- *In vivo* and *in vitro* 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 (sediment/pore 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 calculated and/or proposed but further work in this area is required.
References


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


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=original sample (Cf=1) ea. Marine Environmental Research, 53: 327–356.
Annex 9: Revised Background document on Cytochrome P4501A activity (EROD) (Chapter 3)

Chapter 3 Cytochrome P4501A activity (EROD)

Reviewed by Fetsl S., Maes, T., Cooreman, K. and Martínez-Gómez, C.

Introduction

Martínez-Gómez, C. (IEO, Spain), Maes, T. (Cefas, UK) and Cooreman, K (ILVO, Belgium) reviewed and updated the Background Document on Cytochrome P4501A activity (EROD) and the initial approach on Assessment Criteria (AC) initiated at the joint ICES/OSPAR WKIMON III meeting (2007).

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.

The cytochrome P450 1A family of enzymes are responsible for the primary metabolism of planar polycyclic 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; Goksöy et al., 1991; Whyte et al., 2000) and standardized by ring testing (BEQUALM).

Several endogenous and exogenous factors have been demonstrated 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 revealed 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 (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.

CYP1A expression is suppressed in spawning females due to interference of 17β-estradiol (E2) (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 recrudescence. Therefore interference of environmental estrogens on CYP1A induction can be assessed.
Dietary factors can be potentially important for the induction of CYP1A. First, of course, AhR ligands can be presented to the organism through the food. Second, 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.

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

**Background responses**

Baseline levels of EROD in six marine species have been estimated from results derived from the joint ICES/OSPAR WKIMON III meeting in 2007 and recent data submitted to ICES database (Table 1). The baseline value for Atlantic cod has been suggested to be 80 pmol/min/mg microsomal protein, for dab 40 pmol/min/mg protein 59, for flounder 15 pmol/min/mg protein 59, for dab 3.7 pmol/min/mg protein 59, for four-spotted 12 pmol/min/mg microsomal protein, for dragonet 144 pmol/min/mg microsomal protein and for red mullet 85 pmol/min/mg microsomal protein. The fish were for 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).

**Assessment criteria**

As many factors are known to influence EROD and it is not feasible to correct for all in the design, it is advisable always to include an appropriate reference group in studies that include EROD as an endpoint. Experience suggests that an EROD value in most marine species above twice the upper limit of baseline values (90 percentile) indicate an ecosystem influenced by planar organic contaminants.

**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 WKIMC REPORT 2009 and the workshop SGIMC 2009.

Relevant literature (marine and freshwater fish)


Table 1. EROD Background responses 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 for these species.

<table>
<thead>
<tr>
<th>EROD ASSESSMENT CRITERIA†</th>
<th>SAMPLING SEASON</th>
<th>BOTTOM WATER TEMPERATURE RANGE</th>
<th>SIZE LENGTH</th>
<th>SEX</th>
<th>EROD ACTIVITY (pmol/min/mg prot)</th>
<th>EROD ACTIVITY (pmol/min/mg prot)</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>* MICROSOMES SUB-FRACTION</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dab (Limanda limanda)</td>
<td>August-November</td>
<td>[10-18]</td>
<td>12-25</td>
<td>Females and/or males</td>
<td>&lt; 30†</td>
<td>&lt; 152†</td>
<td>1034</td>
</tr>
<tr>
<td>European flounder</td>
<td>August-November</td>
<td>[10-18]</td>
<td>20-25</td>
<td>Females and/or males</td>
<td>&lt; 14†</td>
<td>&lt; 24†</td>
<td>30</td>
</tr>
<tr>
<td>(Platichthys flesus)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cod (Gadus morhua)</td>
<td>August-November</td>
<td>[10-18]</td>
<td>30-45</td>
<td>Females and/or males</td>
<td>&lt; 78†</td>
<td>&lt; 151†</td>
<td>74</td>
</tr>
<tr>
<td>Four spotted megrim</td>
<td>September-October</td>
<td>[11.7-12.7]</td>
<td>18-22</td>
<td>Females and/or males</td>
<td>&lt; 12†</td>
<td>&lt; 13†</td>
<td>317</td>
</tr>
<tr>
<td>(Lepidorhombus boscii)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Plaice (Pleuronectes</td>
<td>January</td>
<td>[18.5-22.5]</td>
<td>7-10</td>
<td>Males</td>
<td>&lt;3.71†</td>
<td>&lt;9.49†</td>
<td>116</td>
</tr>
<tr>
<td>platessa)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dragonet (Callionymus</td>
<td>September-October</td>
<td>[12.0-12.8]</td>
<td>15-22</td>
<td>Females and/or males</td>
<td>&lt; 144*</td>
<td>&lt;202*</td>
<td>159</td>
</tr>
<tr>
<td>lyra)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Red mullet (Mullus</td>
<td>April</td>
<td>[13.3-15.3]</td>
<td>12-18</td>
<td>Males</td>
<td>&lt; 85*</td>
<td>&lt;208*</td>
<td>40</td>
</tr>
<tr>
<td>barbatus)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Dose-response, background response and sensitivity in experimental studies with gadoid fish.

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

<table>
<thead>
<tr>
<th>SPECIES</th>
<th>SUBSTANCE(S)</th>
<th>LOWEST-HIGHEST CONCS</th>
<th>EXPOSURE TIME</th>
<th>BASELINE/CONTROL (LEVEL/ACTIVITY)</th>
<th>INDUCTION (FOLD)</th>
<th>REFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rockling, Ciliata mustella</td>
<td>Crude oil (Gullfaks; M.V. Braer spill, Shetland)</td>
<td>85000 tons spill</td>
<td>3 months after spill</td>
<td>~160 pmol/min/mg ± 50</td>
<td>~9</td>
<td>(George et al., 1995)</td>
</tr>
<tr>
<td></td>
<td>129 ± 38 ng/g dry wt. of PAHs (selected 2- and 3-ring) detected in muscle.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roundnose grenadier, Coryphaenoides rupestris</td>
<td>i.a. PAHs and PCBs</td>
<td>260 ± 20 (Male) ~170 (Female) pmol/min/mg</td>
<td>~2 (530 ± 70 (male) and ~350 (female) pmol/min/mg)</td>
<td>(Lindesjoo et al., 1996)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hake, Urophycis spp.</td>
<td>Pollution (PAH) from oil platforms (Gulf of Mexico) &lt;100m from platforms</td>
<td>10.9 ± 6.4 and 11.7 ± 10.5 pmol/min/mg (&gt;3000 m from platforms)</td>
<td>&lt;1 (10.6 ± 3.8 and 10.5 ± 7.1 pmol/min/mg)</td>
<td>(McDonald et al., 1996)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Annex 10: Recommendations

WKIMC met in Aberdeen from 14–16 October 2009, and makes the following recommendations:

<table>
<thead>
<tr>
<th>Number</th>
<th>Recommendation</th>
<th>For Follow up By</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.</td>
<td>That the updated assessment criteria for biological effects of contaminants be forwarded WGBEC and SGIMC for a view to it being recommended for acceptance by OSPAR.</td>
<td>ICES Secretariat</td>
</tr>
<tr>
<td>3.</td>
<td>That the updated SGIMC Work Programme be forwarded to ICES/OSPAR as a framework for the further development of the work through SGIMC and WGBEC.</td>
<td>ICES Secretariat</td>
</tr>
<tr>
<td>4.</td>
<td>That the proposals in the WKIMC report on approaches to integrated assessments of biological effects measurements be forwarded to WGBEC and SGIMC for further discussion and elaboration.</td>
<td>ICES Secretariat</td>
</tr>
<tr>
<td>5.</td>
<td>It is recommended that SGIMC 2010 should consider whether background assessment should be considered robust only where the following criteria are met: background information on age, sex and seasonal variation are available, data for 2/3 reference areas is available and data for &gt;100 individuals is available. Where these conditions are not met, simultaneous sampling of study and reference sites should be conducted and current proposed reference values are to be used for guidance to aid interpretation only.</td>
<td>SGIMC</td>
</tr>
<tr>
<td>6.</td>
<td>That the standard mathematical procedure for derivation of (background) assessment criteria used routinely in the assessment of chemical monitoring data should also be applied, where appropriate, to biological effects monitoring data.</td>
<td>SGIMC</td>
</tr>
<tr>
<td>7.</td>
<td>That a secure audit trail for the datasets used for derivation of background assessment criteria, and other assessment criteria, and associated literature where relevant, should be established, possibly through the ICES Datacentre.</td>
<td>Ian Davies, ICES Datacentre</td>
</tr>
</tbody>
</table>