ICES/OSPAR SGIMC REPORT 2011

ICES ADVISORY COMMITTEE

ICES CM 2011/ACOM:30

REF. ACOM, OSPAR

Report of the Study Group on Integrated Monitoring of Contaminants and Biological Effects (SGIMC)

14-18 March 2011

Copenhagen, Denmark

International Council for the Exploration of the Sea

Conseil International pour l'Exploration de la Mer

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

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

Recommended format for purposes of citation:

ICES. 2011. Report of the Study Group on Integrated Monitoring of Contaminants and Biological Effects (SGIMC), 14–18 March 2011, Copenhagen, Denmark. ICES CM 2011/ACOM:30. 265 pp. https://doi.org/10.17895/ices.pub.19280786

Contents

Executive summary

The joint ICES/OSPAR SGIMC (Study Group on Integrated Monitoring of Contaminants and Biological Effects) met for five days in March 2011 at ICES HQ. The meeting was co-chaired by Ian M Davies (UK, OSPAR) and Dick Vethaak (NL, ICES). Eleven participants represented eight nations.

SGIMC develops advice and supporting documentation to meet the needs of both organizations for an integrated approach to the monitoring of chemical contaminants and biological effects for the purposes of assessment of environmental quality/status in relation to contaminants. The output also has strong application to the development of monitoring strategies and assessment criteria for Descriptor 8 of Good Environmental Status under Marine Strategy Framework Directive (MSFD).

The meeting reviewed a range of draft OSPAR Background Documents on biological effects monitoring methods and associated assessment criteria, developed an overarching chapeau document for the integrated approach, provided advice on an integrated assessment process, and developed advice for OSPAR on the application of the integrated approach to MSFD monitoring and assessment. Other tasks, including development of the ICES contaminants database (DOME) and review of draft TIMES series methodological documents were also addressed.

The main body of the report provides a framework for an extensive series of Annexes, many of which are recommended for transfer to OSPAR as advice in response to OSPAR request 2008/8. The main outcomes of the meeting were;

a) A series of 20 documents completed or updated documents on the strategy for the integrated approach, the relevant biological effects with assessment criteria, and a data assessment procedure, recommended for adoption by OSPAR;

b) Review of three draft TIMES methods documents, which were forwarded to the TIMES editorial system.

c) Advice to OSPAR MIME/HASEC on biological effects measurements relevant and available for use in GES assessments under MSFD descriptor 8.

d) Agreement to compile SGIMC advice into an ICES Cooperative Research Report.

e) A review of the SGIMC work programme for 2010–2011, concluding that it has been almost fully completed. Recommendation that outstanding items are referred to ICES WGBEC for completion and that SGIMC be discontinued.

While the supporting documentation and assessment criteria for both chemical and biological measurements are now complete, there is little experience of the application of an integrated assessment scheme covering both types of data. SGIMC recommended that the assessment process be applied to ICON project data, and also to appropriate national datasets as soon as possible.

1 Opening of the meeting

The meeting was opened at ICES HQ at 0900 hr on Monday 14 March 2011. The meeting was co-chaired by Ian Davies (UK, OSPAR) and Dick Vethaak (NL, ICES), and was welcomed to ICES HQ by Claus Hagebro. The list of participants is given in Annex 1.

2 Adoption of the agenda

The draft agenda was adopted without amendment. The Terms of Reference for the meeting require SGIMC to report by 15 April 2011 for the attention of ACOM and OSPAR. It was noted that the primary task to be completed prior to the intended discontinuation of the Study Group was the completion of the SGIMC contribution to OSPAR Request 2008/8, building on work carried out over the preceding two years.

An additional request had been received from OSPAR MIME for assistance in the development of advice to HASEC on a list of biological effect techniques, which could, from a scientific point of view, act a targets and indicators in the OSPAR area for good environmental status Descriptor 8 (hazardous substances) under the MSFD. This was added to the ToR provided in ICES Resolution 2010/2/ACOM31.

A further request had been received from ICES MSFDSG, directed at all EGs in 2011 as follows, and was added to the ToR at this meeting:

- Identify elements of the EGs work that may help determine status for the 11 Descriptors set out in the Commission Decision (available at http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:232:0014:0024:EN:PDF;
- Provide views on what good environmental status (GES) might be for those descriptors, including methods that could be used to determine status.

3 To receive and finalize Background Documents and draft assessment criteria from ICES WGBEC, as indicated in the SGIMC Work Programme

A series of Background Documents and draft TIMES series documents were received from WGBEC 2011, as indicated in Table 1 below. The main tasks were to review and edit the documents and assess their suitability for offering to OSPAR as ICES advice. Table 1 summarizes the actions taken by SGIMC on the documents, and the consequent status and location of the documents in Annexes 4–13.

Thanks are offered to Janina Barsiene, Brett Lyons and Aleksanders Rybakovas for their work on the micronucleus assay Background Document.

Thanks are also offered to Ionan Mariogomez, Miren Cajaraville and John Bignell for their work on the mussel histopathology Background Document, and to Grant Stentiford for his work on the Background Document on intersex in fish.

Table 1.

Annex number	Document	Action	Outcome
4	Background Document on Protocols for extraction, cleanup and solvent exchange methods for small-scale bioassays.	Reviewed and edited	Approved by SGIMC and recommended as advice for OSPAR. Included in report as Annex
5	Background Document on Intersex (ovotestis) measurement in marine and estuarine fish	Reviewed and edited. Assessment criteria (Background responses and EAC-equivalent) added to the Background document	Approved by SGIMC and recommended as advice for OSPAR. Included in report as Annex
6	Background Document on Supporting parameters for biological effects measurements in fish and mussels	Reviewed and edited	Approved by SGIMC and recommended as advice for OSPAR. Included in report as Annex
7	Background Document on Acetylcholinesterase assay as a method for assessing neurotoxic effects in aquatic organisms	Reviewed and edited	Approved by SGIMC and recommended as advice for OSPAR. Included in report as Annex
8	Background Document on Histopathology of mussels *Mytilus* sp. for health assessment in biological effects monitoring	Reviewed and edited	Approved by SGIMC and recommended as advice for OSPAR. Included in report as Annex
	Mussel histology TIMES document	Reviewed and edited	Approved by SGIMC subject to edits and passed to TIMES editor

Annex number	Document	Action	Outcome
9	Background Document on Micronucleus assay as a tool for assessing cytogenetic/DNA damage in marine organisms	Reviewed and edited	Approved by SGIMC and recommended as advice for OSPAR. Included in report as Annex
10	Background Document on Comet assay as a method for assessing DNA damage in aquatic organisms	Reviewed and edited	Approved by SGIMC and recommended as advice for OSPAR. Included in report as Annex
11	Background Document on Sediment seawater elutriate and pore-water bioassays with early developmental stages of marine invertebrates.	Reviewed and edited	Approved by SGIMC and recommended as advice for OSPAR. Included in report as Annex
12	Background Document on Sediment seawater elutriate and pore-water bioassays with copepods (*Tisbe*, *Acartia*), mysids (*Siriella*, *Praunus*), and decapod larvae (*Palaemon*).	Reviewed and edited	Approved by SGIMC and recommended as advice for OSPAR. Included in report as Annex
13	Background Document on Whole sediment bioassays with amphipods (*Corophium* sp) and *Arenicola marina*	Assessment Criteria updated and incorporated in BG	Approved by SGIMC and recommended as advice for OSPAR. Included in report as Annex
	Draft ICES-TIMES document: Biological effects of contaminants: Receptor H4IIE-Luciferase (DR-Luc) cell bioassay for screening of dioxins and/or dioxin-like compounds in environmental samples. C.A. Schipper, P.E.G. Leonards, H.J.C. Klamer, K.V. Thomas, A.D. Vethaak	Reviewed	Approved by SGIMC and recommended for publication in ICES TIMES series after resolution of points raised in review. Passed to WGBEC TIMES contact.
	Draft ICES-TIMES document: Protocols for extraction, cleanup and solvent exchange methods for small-scale bioassays. Hans Klamer, Knut-Erik Tollefsen, Steven Brooks John Thain	Reviewed	Approved by SGIMC and recommended for publication in ICES TIMES series after resolution of points raised in review. Passed to WGBEC TIMES contact.

4 To receive and finalize Background Documents and draft assessment criteria arising from the 2010 meeting, as indicated in the SGIMC Work Programme

A series of Background Documents and TIMES series papers were received or developed during the meeting, as indicated in Table 2 below. The main tasks were to review and finalize the documents and assess their suitability for offering to OSPAR as ICES advice. Table 2 summarizes the actions taken by SGIMC on the documents, and the consequent status and location of the documents.

Table 2.

Annex	Document	Action	Outcome
14	Background Document on DNA adducts, including BAC responses and EAC-equivalent Assessment Criteria	Document reviewed and edited, and assessment criteria added	Approved by SGIMC and recommended as advice for OSPAR. Included in report as Annex
15	Background Document on *in vitro* DR-Luc/DR-CALUX® bioassay for screening of dioxin-like compounds in marine and estuarine sediments	Document reviewed and edited, Background responses and EAC-equivalent Assessment Criteria added to the document	Approved by SGIMC and recommended as advice for OSPAR. Included in report as Annex
	DR-Luc/DR-CALUX® ICES-TIMES document (comments to be received from WGBEC and WGMC)	Document reviewed and edited	Approved by SGIMC subject to edits and sent to TIMES editor for WGBEC
16	Background document on Metallothionein (MT) in blue mussel (*Mytilus edulis*, *Mytilus galloprovincialis*)	Assessment criteria added to the BG document	Approved by SGIMC and recommended as advice for OSPAR. Included in report as Annex
17	Background Document on Water bioassays (to exclude *in vitro* bioassays)	Document reviewed and edited, and assessment criteria reviewed	Approved by SGIMC and recommended as advice for OSPAR. Included in report as Annex
18	Background Document on Externally visible fish diseases, macroscopic liver neoplasms and liver histopathology	Document reviewed and edited. Assessment criteria added	Approved by SGIMC and recommended as advice for OSPAR. Included in report as Annex
	Draft ICES-TIMES document: Histopathology of mussels *Mytilus* sp. for health assessment in biological effects monitoring. J Bignall, S Feist, M Caraville, I Margomez, A Villalba	Reviewed	Approved by SGIMC and recommended for publication in ICES TIMES series after resolution of points raised in review. Passed to WGBEC TIMES contact.

Establishing Environmental Assessment Criteria for Vitellogenin (Vtg)

The OSPAR Background Document on Vtg written in 2007 thoroughly reviewed the data available for establishing Environmental Assessment Criteria (EAC). The conclusion at the time was that an EAC could not 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, notably flounder and cod. SGIMC reviewed the literature on which this was based and new publications since 2007 and concluded that it was still not possible to establish an EAC for any species. Therefore, the Background Concentration (BC) of 0.13 μg/ml for Vtg in male flounder, and the tentative BC of 0.23 μg/ml proposed for cod remain as the only values proposed for assessment criteria.

Progress and opportunities for *in vitro* bioassays of endocrine disrupting compounds

Part of the work of SGIMC was to consider the derivation of assessment criteria for the vitellogenin (Vtg) assay, as described above. Other approaches to endocrine disruption are available for example the ER-Luc/CALUX method.

In order to determine the potential risk of endocrine disrupting compounds in the marine environment, *in vitro* bioassays, in addition to DR-Luc are required. These refer in particular to *in vitro* receptor assays for measuring estrogenicity and androgenicity, i.e. ER-Luc, YES and YAS. These *in vitro* bioassays are required as components of the suite of measurements in the JAMP Specific Guidelines for Biological Effects of endocrine estrogenic compounds and for the overall integrated monitoring approach as developed by SGIMC, and should therefore be adopted by OSPAR and included in the JAMP/CEMP when the necessary Background Documents, etc. are available. Background information on the applicability of these in vitro bioassays is provided in the background document for water *in vivo* bioassays.

In past years, ER-Luc, YES and YAS have been routinely used to determine (anti) estrogenic or (anti) androgenic activity in complex field samples including extracts of water, production water, suspended particulate matter, sediment and fish bile (Garcia-Reyero *et al*., 2001; Murk *et al*., 2002; Legler *et al*., 2002a, 2002b; Thange *et al*., 2000; Houtman *et al*., 2004; Thomas *et al*., 2006; Johnston *et al*., 2007). The methods for conducting ER-Luc, YES and YAS are fully described and validated in the scientific literature (see table below). The extraction methods for these assays are available through the ICES TIMES series document on extraction protocols. Work on assessment criteria and QA for these methods is, however, still limited and need to be progressed, ideally via ICES WGBEC and international programmes such as BEQUALM.

Method	Reference
ER-Luc/ER CALUX	Legler *et al*., 1996; 2002a; 2002b.
YES (Yeast Estrogen Screen)	Routledge and Sumpter, 1996; de Boever *et al*., 2001.
YAS (Yeast Androgen Screen)	Sohoni and Sumpter, 1998.

Recommendation: That ICES WGBEC should further progress the work on *in vitro* bioassays for endocrine disruption, with a view to OSPAR adopting these assays as components of the JAMP specific biological effect guidelines for estrogenic compounds and the overall integrated chemical-biological effects monitoring approach.

References

De Boever P, Demare W, Vanderperren E, Cooreman K, Bossier P, Verstraete W. Optimization of a yeast estrogen screen and its applicability to study the release of estrogenic isoflavones from a soygerm powder.. Environmental Health Perspectives. 2001, 109 (7): 691–697.

Garcia-Reyero N, Grau E, Castillo M, Lopez de Alda MJ, Barcelo D, Pina B. 2001. Monitoring of endocrine disruptors in surface waters by the yeast recombinant assay. Environ Toxicol Chem 20:1152–1158.

Houtman CJ, Cenijn PH, Hamers T, Lamoree MH, Legler J, Murk AJ, Brouwer A. 2004. Toxicological profiling of sediments with *in vitro* bioassays with emphasis on endocrine disruption. Environ Toxicol Chem 23:32–40.

Legler J, Van den Brink CE, Brouwer A, Murk AJ, Van der Saag PT, Vethaak AD, Van der Burg B. 1999. Development of a stably transfected estrogen receptor-mediated luciferase reporter gene assay in the human T47D breast cancer cell line. Toxicol Sci 48:55–66.

Legler J, Dennekamp M, Vethaak AD, Brouwer A, Koeman JH, Van der Burg B, Murk AJ. 2002a. Detection of estrogenic activity in sediment-associated compounds using *in vitro* reporter gene assays. Sci Total Environ 293:69–83.

Legler J, Jonas A, Lahr J, Vethaak AD, Brouwer A, Murk. AJ. 2002b. Detection of estrogenic activity in human urine and fish bile with the *in vitro* ER-CALUX reporter gene assay. Environ Toxicol Chem 21:473–479.

Murk AJ, Legler J, van Lipzig MMH, Meerman JHN, Belfroid AC, Spenkelink A, van der Burg B, Rijs GBJ, Vethaak D. 2002. Detection of estrogenic potency in wastewater and surface water with three *in vitro* bioassays. Environ Toxicol Chem 21: 16–23.

Johnson, I., Hetheridge, M. and Tyler, C. R. 2007. Assessment of (anti-) oestrogenic and (anti-) androgenic activities of final effluents from sewage treatment works. Science report-SC020118/SR, Environment Agency, Bristol, UK, pp. 61.

Routledge EJ, Sumpter, JP. 1996. Estrogenic activity of surfactants and some of their degradation products assessed using a recombinant yeast screen. Environ Toxicol Chem 15:241–248.

Sohoni P, and Sumpter JP. 1998. Several environmental oestrogens are also antiandrogens. J. Endocrinol 158(3):327–339.

Tanghe T, Devriese G, Verstraete W. 2000. Use of a recombinant yeast oestrogen assay for the determination of oestrogenic activity in aquatic environmental samples of Flanders. In: Endocrine Disrupting Compounds: Wildlife and Human Health Risks, Symposium proceedings, October 27, 1998, The Hague, the Netherlands, pp 44–48.

Thomas. K.V., Hurst M.R., Reynolds W., Thain J.E. 2006. *In vitro* bioassay testing of produced and surface water extracts. In Biological effects of contaminants in marine pelagic ecosystems (editors Hylland, Lang and Vethaak), SETAC, ISBN: 13: 978-1-880611-84-8. Belgium.

5 To review and comment on the outcome of the joint MED-POL/SGIMC Workshop held in Italy in 2010

SGIMC reviewed and commented on the final report on ICES/OSPAR WKLYS on the quality and interpretation of lysosomal stability data report and assessment criteria for LMS using NRR assay.

C. Martínez-Gómez (Spain) explained that during WKLYS the main aspects of the operational procedure to harmonize the use of the Neutral Red Retention assay were identified and it was concluded that a further discussion and consensus will be necessary, in terms of monitoring and intercomparison purposes, to use this technique through the ICES/OSPAR and MED POL area. The main reason is that, apart from the analytical procedure described in UNEP/RAMOGE, 1999 for MED POL programme, a new analytical procedure based on image analysis is currently being recommended in MED POL training courses to assess NRR time in mussels, being endpoints obtained in a different way.

During WKLYS, an illustrated draft document was produced to be used through the ICES/OSPAR area, with proposals for improved consistency in the interpretation of observations under microscope when NRR is used. This draft illustrated document was discussed in WKLYS with A. Viarengo, M. Moore and it was sent A. Kohler and D. Lowe for their comments.

Few participants from OSPAR area attended WKLYS and the progress with the data quality and interpretation was less than expected. However, a substantial progress was done on identification of the potential sources of variation affecting results between laboratories. During last WGBEC 2010, countries assessing LMS in monitoring programmes (Finland, Norway, Denmark, Ireland and Sweden) reported information about technical procedure used to establish LMS in mussels, including the NRR time endpoint.

The dataset used to calculate and establish ACs for lysosomal membrane stability is actually not available as conclusions were drawn from review of a large number of papers, and many years of practical experience (Professor M. Moore). During WGBEC 2011 it was proposed that the TIMES document, including assessment criteria, and corresponding Background Document are reviewed during 2011–2012.

A new proposal for NRR assay intercalibration exercise was proposed by using virtual slides in a similar way to the BEQUALM Fish Disease Measurement programme which has been successful during the last three years.

SGIMC 2011 highlighted the necessity of a review of NRR assessment criteria and ICES TIMES document.

ACTION: TIMES ICES protocol and corresponding Background Document should be amended/extended in relation with some practical details and illustration after feedback from D. Lowe and A. Kohler (ICES TIMES authors) have been received. It was agreed to conduct it through contact with ICES TIMES authors to be presented to WGBEC at their 2012 meeting.

6 To check the advice of SGIMC sent from ICES to OSPAR last year and resolve possible confusion over the integrated sampling schemes, if necessary

SGIMC 2011 reviewed the advice given to OSPAR by ICES after SGIMC 2010 and found that some errors had been made in the transcription of Figures representing the integrated schemes for fish and shellfish sampling/monitoring (Annexes 9 and 10 of the SGIMC 2010 report), at least partially due to file format giving rise to difficulties in editing. The Figures have been corrected, and the revised documents are included at Annex 20 of the current report, and SGIMC recommends that they are passed to OSPAR as updates.

7 To finalize work on OSPAR request 2008/8: Completion of draft merged guidelines for integrated monitoring and assessment of contaminants and their effects + technical annexes including one on survey design)

SGIMC continued with the preparation of a series of documents required to complete the merged guidelines for integrated monitoring. In particular, SGIMC developed an over-arching approach to ecosystem assessment for contaminants and their effects. This covered water sediment, fish and shellfish (mussels and gastropods) as the monitoring matrices and identified the key chemical and biological effects measurements that were relevant in each case.

The measurements covered chemistry, subcellular responses, tissue level and whole organism responses and were classified according to their degree of development into a core set of measurements, and additional measurements. The over-arching Guideline is attached as Annex 21.

Key steps in preparing methods for consideration for inclusion in OSPAR monitoring programmes are the availability of Background Documents, assessment criteria, and quality assurance procedures. SGIMC therefore continued to prepare Background Documents, as indicated in their Work programme in the SGIMC 2010 report, and to receive other documents from WGBEC for review. These Background Documents include assessment criteria in the form of Background responses (for all methods) and EAC-analogues (where appropriate). The documents also generally include sections on quality control, and on the availability of external QA. These documents therefore meet the requirements for methods to become available for adoption into the CEMP.

The availability of supporting documentation for biological effects measurements in the form of Background Documents is summarized in Annex 22, together with information on assessment criteria and quality assurance. The values of the assessment criteria for biological effects measurements are tabulated in Annex 23.

The Background Documents completed at this meeting are listed in Tables 1 and 2, and included in Annexes as indicated.

8 To consider MSFD matters raised by OSPAR

a) consider specific request by OSPAR MIME

- To clarify that point d) on the SGIMC 2011 terms of reference (Annex 3a of the SGIMC 2010 report) relating to "completing the draft merged Guidelines for the Integrated Monitoring and Assessment of Contaminants and their effects and drafting the necessary Technical Annexes […]" includes the finalizing of the actual guidelines for integrated chemical and biological effects monitoring, setting out the agreed principles as chapeau for the Technical Annexes;

The report of this meeting does include finalized guidelines for an strategy for integrated monitoring of contaminants and biological effects (Annex 21) to act as chapeau as suggested above.

- To request SGIMC to further develop the attached starting point document (MIME 2010 Annex 7) and prepare advice on a consolidated list of biological effects techniques which, from a scientific point of view, could act as targets and indicators in the OSPAR area for good environmental status descriptor 8 (hazardous substances) under the MSFD.

OSPAR has set up a coordination mechanizm by which all expert groups/Committees are invited to develop advice on characteristics, relevant to the regions or subregions of the OSPAR maritime area, for GES descriptors and on methodologies for determining targets and associated indicators. This should help Contracting Parties to regionally coordinate their national setting of targets and indicators. As part of this task, MIME has been invited to advise on developing a list of biological effects techniques relevant to the OSPAR maritime area under Descriptor 8. OSPAR MIME had requested assistance to help them develop advice to HASEC on a list of biological effect techniques would could, from a scientific point of view, act a targets and indicators in the OSPAR area for good environmental status Descriptor 8 (hazardous substances) under the MSFD.

SGIMC noted that formulation of Descriptor 8 was that "Concentrations of contaminants are at levels not giving rise to pollution effects". The Task Group set up by ISPRA interpreted this to mean that the concentrations of contaminants should not exceed established quality standards (e.g. EQSs, EACs) and that the intensity of biological effects attributable to contaminants should not indicate harm at organism or higher levels of organization.

The subsequent Commission Decision (2010/477/EU) on "criteria and methodological standards on good environmental status of marine waters expressed this as follows:

Progress towards good environmental status will depend on whether pollution is progressively being phased out, i.e. the presence of contaminants in the marine environment and their biological effects are kept within acceptable limits, so as to ensure that there are no significant impacts on or risk to the marine environment.

8.1 Concentrations of contaminants

- Concentration of the contaminants mentioned above, measured in the relevant matrix (such as biota, sediment and water) in a way that ensures comparability with the assessments under Directive 2000/60/EC

8.2 Effects of contaminants

- Levels of pollution effects on the ecosystem components concerned, having regard to the selected biological processes and taxonomic groups where a cause/effect relationship has been established and needs to be monitored.

In turn, these are being translated into indicators as, for example (UK):

a) Indicator 8.1.1 Concentrations of contaminants in water, sediment or biota are not increasing and do not exceed environmental target levels identified on the basis of ecotoxicological data as outlined within community legislation and other obligatory agreements (such as OSPAR).

b) Indicator 8.2.1. Biological effects of contaminants are below environmental target levels considered to result in harm at organism, population, community and ecosystem levels as outlined within community legislation and other obligatory agreements.

It is clear that assessment for Descriptor 8 will require both chemical and biological effects measurements. In the case of effects measurements, ICES/OSPAR have established a number of prerequisites for the inclusion of effects measurements in OSPAR programmes such as the CEMP. These are that:

- an OSPAR Background Document has been prepared, and
- assessment criteria have been agreed.

In addition, comparability of data will be greatly improved if quality control/quality assurance schemes are available. SGIMC consider that it is appropriate that similar prerequisites are applied to methods for MSFD purposes.

SGIMC, and its predecessor WKIMON, have been working for some years on the definition of integrated monitoring procedures for chemical and biological effects measurements. The current recommendations are in Annex 21 to this report. The status of the various effects measurements, and their interrelations in integrated assessments related to fish, mussels, sediment, water and gastropods are indicated in this document. The maturity of the various techniques for monitoring and for assessment and the prospect for their practical linking with chemical concentrations in an assessment context have been important factors in the development of the integrated schemes.

SGIMC therefore recommend that a similar approach is adopted in relation to Descriptor 8. The ecosystem components of the integrated approach to the assessment of contaminants and their biological effects described in Annex 21 can be readily adapted to the needs of Descriptor 8. The integrated monitoring scheme can be summarized in series of diagrams referring to assessment matrices, such as water, sediment, fish and shellfish. Appropriate core chemical and biological effects measurements are listed, for which Background Documents and assessment criteria have been prepared, together with a smaller number of potential additional measurements.

SGIMC therefore developed the MIME document (MIME 2010, Annex 7, Starting point for biological effects relevant to good environmental status in the OSPAR area), and the amended document is attached as Annex 24.

SGIMC noted that a coherent and integrated assessment scheme was an essential element of integrated monitoring, and this is covered in Agenda item 8 ii) below.

b) The study group proposes that the assessment of integrated monitoring data conducted according to the guidelines proposed here provides an appropriate tool for the assessment of Good Environmental Status for Descriptor 8 of MSFD 'concentrations of contaminants are at levels not giving rise to pollution effects'. However, an integrated assessment framework had not yet been agreed during the course of the SGIMC workplan. The group reviewed previous proposals for assessment frameworks and developed a scheme building on the experience of the assessment of contaminants data for sediment, fish and shellfish in OSPAR contexts. This proposal is described in more detail with a worked example (using artificial data) in Annex 25.

The proposed process is informed initially by the individual assessment of determinands (contaminants or effects) in specific matrices at individual sites against the defined assessment criteria (BAC and EAC). Such assessment criteria for biological effects have been developed over recent years and are included in OSPAR Background Documents, and for contaminants have been used by OSPAR groups, for example in the QSR 2010. Initial comparisons determine whether the determinand and site combinations are <BAC (blue), between the BAC and EAC (green) or >EAC (red). This summarized indicator of status for each determinand can then be integrated over a number of levels: matrix (sediment, water, fish, mussel, gastropod), site and region and expressed with varying levels of aggregation to graphically represent the proportion of different types of determinands (or for each determinand, sites within a region) exceeding either level of assessment criteria.

Such an approach has several advantages. The integration of data can be simply performed on multiple levels depending on the type of assessment required and the monitoring data available. The representation of the assessment maintains all the supporting information and it is easy to identify the causative determinands that may be responsible for exceeding EAC levels. In addition, any stage of the assessment can be readily unpacked to a previous stage to identify either contaminant or effects measurements of potential concern or sites contributing to poor regional assessments.

This approach builds on the OSPAR MON regional assessment tool developed for contaminants. The development of BAC and EAC equivalent assessment criteria for biological effects, which represent the same degree of environmental risk as indicated by BACs and EACs for contaminants, allows the representation of these monitoring data alongside contaminant data using the same graphical representation approach. The inclusion of biological effects data to the system adds considerable value to the interpretation of assessments. Where sufficient effects monitoring data are available, confidence can be gained that contaminants are not having significant effects even where contaminant monitoring data are lacking. In instances where contaminant concentrations in water/sediment are >EAC, a lack of EAC threshold breach in appropriate effects data can provide some confidence that contaminant concentrations are not giving rise to pollution effects (due for example to lack of availability to marine biota). Similarly, the inclusion of effects data in the assessment framework can indicate instances where contaminants are having significant effects on biota, but have not been detected or covered in contaminant-specific chemical monitoring work.

The assessment framework described in Annex 25 provides an appropriate tool for assessment of environmental monitoring data to determine whether Good Environmental Status is being achieved for Descriptor 8 of MSFD (concentrations of contaminants are at levels not giving rise to pollution effects). Determinands with EAC or

EAC equivalent assessment criteria provide appropriate indicators with quantitative targets. The assessment of contaminant and effects monitoring data against these EAC-level assessment criteria provides information both on concentrations of contaminants likely to give rise to effects and the presence/absence of significant effects in marine biota. Thresholds for the proportions of determinands falling below EAC thresholds could be used to help determine GES (95% compliance is proposed initially).

Further details of the proposed assessment framework are given in Annex 25. The proposed scheme requires testing with real monitoring data before it could be adopted.

9 To receive and review preliminary reports of the results of the ICON project, including the application of integrated assessment

There was insufficient time to address this item in any detail. However, SGIMC noted that the data were now almost complete, and that they should present a good opportunity for application of the integrated data assessment process described below. ICON aim to undertake this assessment by late June 2011.

10 To consider the collation of the combined outputs from SGIMC into a single report

SGIMC recommended that the output advice to OSPAR covering the work of SGIMC on integrated monitoring should be brought together in a single document, with a view to publication in an appropriate ICES series. Preliminary discussions with the editor of ICES publications (Bill Anthony) indicated that the Cooperative Research Report series would be the most appropriate vehicle.

The chairs agreed to progress this intersessionally, and SGIMC agreed a Category 1 Resolution to prepare the way for the publication (Annex 19).

11 To identify elements of the EGs work that may help determine status for the 11 Descriptors set out in the Commission Decision

The integrated monitoring and assessment contaminants and their biological effects in the theme of SGIMC, and also matches closely the requirements of MSFD Descriptor 8 (Concentrations of contaminants are at levels not giving rise to pollution effects). Some members of SGIMC have been involved in the Task Group for Descriptor 8 under the ICES/ISPRA MSFD implementation project which reported about one year ago.

The work of SGIMC is very directly relevant to status assessment under Descriptor 8. In particular,

- The philosophy of integrated monitoring, as outlined in the draft Guideline for OSPAR (Annex 21).
- The schemes for integrated monitoring of contaminants and effects in water, sediment, and biota
- The assessment criteria developed for biological effects, which have been incorporated in a large number of Background Documents, and which are summarized in Annexes 22 and 23.
- The proposed scheme for integrated assessment of contaminants and their effects described in Section 8 ii) and Annex 25 of this report.
- The Technical Annexes for biological effects that have been updated by SGIMC and WKIMON over the last few years and are included as Annexes to the annual reports.

12 To provide views on what good environmental status (GES) might be for those descriptors, including methods that could be used to determine status

SGIMC agreed that GES should be related to the assessment criteria for effects and contaminants that have been variously developed by OSPAR and ICES over recent years. These are Background Concentrations, Background Assessment Concentrations, and Environmental Assessment Criteria for contaminant concentrations, and their analogues for biological effects measurements. The assessment criteria for effects that are currently available are summarized in Annexes 22 and 23, and explained in more detail in the relevant Background Documents.

SGIMC consider that GES should be related to concentrations and the intensity of effects being less than EACs.

Section 8 ii) and Annex 25 of the SGIMC 2011 report describe an integrated assessment scheme that combines chemical and biological effects data through the coherent set of assessment criteria that have been developed by ICES/OSPAR, including those for biological effects developed by SGIMC. The scheme describes how measurements of various parameters in various environmental matrices at various stations can be progressively summarized into simple visual representations of status at different degrees of data aggregation. At the highest level, data for both contaminant concentrations and their effects can be represented at MSFD Regional level by a single three colour “traffic light”. SGIMC consider that the critical boundary for GES assessment is the green–red boundary, representing comparisons with EACs. SGIMC recommend that GES be expressed as some high percentage compliance with this boundary.

SGIMC consider that 100% compliance is impractical, as it amounts to a “one out all out” approach, and is therefore highly susceptible to perturbations by a small number of errors in sampling, analysis or data handling, or short-term variations in environmental quality. SGIMC therefore suggest that 95% compliance at the highest level of data aggregation would be an appropriate threshold for GES compliance.

13 To close the Study Group, and make recommendations for continuation of ICES work on integrated monitoring of contaminants and biological effects in the light of the developments in the definition of GES under MSFD, particularly for Descriptor 8

SGIMC noted that most of the tasks identified at their first meeting in 2009 had been completed, and that a large volume of advisory text had been made available to OSPAR. An analysis of progress against the SGIMC Work Plan for 2009–2011 (Annex 27) shows that the vast majority of the tasks identified by SGIMC have been completed, many in association with WGBEC. SGIMC recommend that the few remaining tasks are transferred to WGBEC for completion. The work of SGIMC had moved the foundations for integrated chemical and biological effects assessments forward significantly, building on the earlier work of WKIMON and working closely with colleagues in WGBEC.

The completion of a large body of Background Documents and assessment criteria was particularly appropriate at this time, as monitoring programmes and data assessment procedures for the Marine Strategy Framework Directive are a high priority for EU countries and Regional Conventions at this time. The integrated approach brought forward by SGIMC is directly applicable to status assessment under Descriptor 8 of GES.

It is inevitable that technique will develop in the coming years, and that new data will highlight where assessment criteria need to be reviewed. SGIMC recommend that WGBEC complete the minor tasks which remain outstanding from the SGIMC programme, and keep a watching brief in the subject area, particularly regarding the need for periodic review of assessment criteria to match the MSFD assessment cycle.

14 Any other business

14.1 Interactions with ICES DataCentre

a) Members of SGIMC spent time with the ICES DataCentre responding to specific questions regarding the parameters required to enable ICES to accept data on CALUX/Luc methods, and provided advice on the Comet assay.

In doing so, it was noted that the DataCentre was also not yet able to accept data for several other biological effects measurements that are included in the integrated monitoring schemes. SGIMC therefore provided advice to the DataCentre on data requirements for micronucleus assay, mussel histopathology and gametogenesis, stress on stress, mussel condition, intersex in fish, and reproductive success in fish. WGBEC has formed an intersessional subgroup to work with the DataCentre to complete the tasks.

A series of conversations were held between SGIMC and the ICES DataCentre to identify the additional capability that will be required of DOME in order to accept data on:

- DR-Luc
- COMET Assay
- Micronucleus assay
- Stress on stress
- Mussel condition
- Histology
- Parasites in organisms

The details of the requirements are tabulated in Annex 26.

b) In conversation with the DataCentre, it became clear that the seawater section of the contaminants database had been relatively little used (compared to the sediment and biota sections), and that it had not received the same attention and development as the other sections. The development of Descriptor 8 of the MSFD may well lead to more data on concentrations of contaminants in seawater coming available and being submitted to ICES, raising the potential for coordinated assessment exercises if the data are coherent and of good quality. In order to develop the seawater database:

SGIMC recommends that the seawater contaminants section of the ICES database should be reviewed, in collaboration with the ICES DataCentre, and updated, giving particular attention to parameter fields, station names/locations, gross errors (e.g. units), uncertainty in identification of contaminants, and the opportunity for improving the QA of data being submitted, for example through automated checking of data at the time of entry, with a view to the seawater data becoming more available for assessment in MSFD and other contexts.

c) SGIMC thank the DataCentre for the provision of extracts of fish disease data during the meeting to enable calculations of assessment criteria.

15 Adoption of the report and closure of the meeting

It was agreed that the Co-chairs would lead on completion of the Meeting Report. The meeting was closed at 1300 h on 18 March 2011.

Annex 1: List of participants

Name	Address	Phone/Fax	E-mail
Kris Cooreman	Institute for Agricultural and Fisheries Research (ILVO) Ankerstraat 1 8400 Oostende Belgium	Phone +32 59569820 Fax +32 59330629	kris.cooreman@ilvo.vlaanderen.be
Ian M. Davies Co-chair	Marine Scotland Marine Laboratory PO Box 101 AB11 9DB Aberdeen United Kingdom	Phone +44 1224 295468 Fax +44 1224 295511	Ian.Davies@scotland.gsi.gov.uk
Michelle Giltrap	Trinity College Dublin Zoology Department College Green Dublin 2 Ireland	Phone +353 1 8964187	giltrapm@tcd.ie
Matt J. Gubbins	Marine Scotland 375 Victoria Rd Aberdeen AB11 9DB United Kingdom	Phone +44 1224 876544 Fax +44 1224 295511	Matthew.Gubbins@scotland.gsi.gov.uk
Ketil Hylland	University of Oslo Dept of Biology PO Box 1066 N-0316 Oslo Norway	Phone +47 22857315	ketil.hylland@bio.uio.no
Thomas Maes	Centre for Environment, Fisheries and Aquaculture Science (Cefas) Pakefield Road NR33 0HT Lowestoft Suffolk United Kingdom	Phone +44 1502 524433 Fax +44 1502 513865	thomas.maes@cefas.co.uk

Name	Address	Phone/Fax	E-mail
Concepcion Martinez-Gomez	Instituto Español de Oceanografía Centro Oceanográfico de Murcia PO Box 22 E-30740 San Pedro del Pinatar Murcia Spain	Phone +34 968180500 Fax +34 968184441	concepcion.martinez@mu.ieo.es
John Thain	Centre for Environment, Fisheries and Aquaculture Science (Cefas) Weymouth Laboratory The Nothe, Barrack Road DT4 8UB Weymouth Dorset United Kingdom	Phone +44 (0)1305 20 6617 Fax +44 (0)1305 20 6601	john.thain@cefas.co.uk
Dick Vethaak Co-chair	Ecosystem Analysis and Assessment (EEA) Unit Marine and Coastal Systems Rotterdamseweg 185 2600 MH Delft The Netherlands	Phone +31 015-2858659/+31 065-1232412	dick.vethaak@deltares.nl
Katrin Vorkamp MCWG Chair	National Environmental Research Institute Frederiksborgvej 399 DK-4000 Roskilde Denmark	Ph +45 463 01859 Fax:+45 463 01114	kvo@dmu.dk
Werner Wosniok	University of Bremen Institut für Statistik, FB 3 PO Box 330 440 DE-28344 Bremen Germany	Phone +49 42 12 1863784 Fax +49 42 12 18 63799	wwosniok@math.uni-bremen.de

Annex 2: Terms of Reference

The **Study Group on Integrated Monitoring of Contaminants and Biological Effects** (SGIMC), chaired by Ian M Davies, UK and Dick Vethaak, The Netherlands, will meet at ICES Headquarters, 14–18 March 2011 to:

a) Receive Background Documents and draft assessment criteria from ICES WGBEC, as indicated in the SGIMC Work Programme, assess their usefulness in integrated assessments, and finalize the documents.

b) Receive Background Documents and draft assessment criteria arising from the 2010 meeting, as indicated in the SGIMC Work Programme, assess their usefulness in integrated assessments, and finalize the documents.

c) Review and comment on the outcome of the joint MEDPOL/SGIMC Workshop held in Italy in 2010.

d) Finalize work on OSPAR request 2008/8.

e) Further develop the integrated assessment framework proposed by SGIMC 2010 and undertake a trial assessment using example monitoring data.

f) Consider the collation of the combined outputs from SGIMC into a single report.

g) Close the Study Group, and make recommendations for continuation of ICES work on integrated monitoring of contaminants and biological effects in the light of the developments in the definition of GES under MSFD, particularly for Descriptor 8.

h) To receive and review preliminary reports of the results of the ICON project, including the application of integrated assessment.

i) MSFD matters:

Consider specific request by OSPAR MIME;

Update on MSFD activities including integration of different effects measurements into single expression of quality (input from WGBEC).

Additional ToR from MSFDSG:

j) Identify elements of the EGs work that may help determine status for the 11 Descriptors set out in the Commission Decision.

k) Provide views on what good environmental status (GES) might be for those descriptors, including methods that could be used to determine status.

Annex 2a: Agenda

Date	Approx time	Agenda item	Issue	Lead
Monday 14th March	09:30	1	Introduction by co-chairs, tour de table	IanD
	10:00	2	Adoption of agenda	IanD
		3	Receive Background Documents and draft assessment criteria from ICES WGBEC, as indicated in the SGIMC Work Programme on i) Extraction procedures for bioassay methods ii) Intersex in fish iii) BG document on supporting parameters iv) Acetylcholinestrease v) Mussel histology vi) Micronucleus and comet assay vii) *In vitro* YES/YAS, ER CALUX assays viii) Sediment and eluctriate bioassay for invert bioassays ix) Sediment and elutriate bioassays with copepods x) Update whole sediment bioassay AC to assess their usefulness in integrated assessments and finalize the documents	MattG + JohnT

Date	Approx time	Agenda item	Issue	Lead
		4	Receive Background Documents and draft assessment criteria arising from the 2010 meeting, as indicated in the SGIMC Work Programme on:	
			xi) DNA adducts BG responses and EAC-equivalent AC	BrettLand IanD
			xii) DR-Luc/CALUX BG document	DickV
			xiii) DR-Luc/CALUX BG responses and EAC-equivalent AC	DickV
			xiv) DR-Luc/CALUX ICES-TIMES document (comments to be received from WGBEC and WGMC)	DickV
			xv) VTG EAC-equivalent AC	IanD and DickV
			xvi) Intersex BG responses and EAC-equivalent AC	SteveF and IanD
			xvii) MT and ALA-Develop BC using recent data	KetilH
			xviii) Update chapter 7 BG document on Water bioassays (to exclude in vitro bioassays)	DickV
			and assess their usefulness in integrated assessments, and finalize the documents.	
		6	Review and comment on the outcome of the joint MEDPOL/SGIMC Workshop held in Italy in 2010.	ConcepcionM
			To check over the advice that was sent from ICES to OSPAR last year and resolve possible confusion over the integrated sampling schemes, if necessary.	IanD

Date	Approx time	Agenda item	Issue	Lead
		7	Finalize work on OSPAR request 2008/8: Completion of draft merged guidelines for integrated monitoring and assessment of contaminants and their effects + technical annexes including one on survey design)	DickV + Subgroup
		8	MSFD matters: i) consider specific request by OSPAR MIME ii) Update on MSFD activities including integration of different effects measurements into single expression of quality (input from WGBEC)	IanD +DickV All MattG
		9	To receive and review preliminary reports of the results of the ICON project, including the application of integrated assessment.	KetilH
		10	Consider the collation of the combined outputs from SGIMC into a single report.	IanD + DickV
		11	Close the Study Group, and make recommendations for continuation of ICES work on integrated monitoring of contaminants and biological effects in the light of the developments in the definition of GES under MSFD, particularly for Descriptor 8.	IanD + DickV + all
		12	Any other business	IanD
	18 March 13.00	13	Adoption of the report and closure of the meeting	IanD
			SGIMC will report by 15 April 2011 for the attention of ACOM.	

Annex 3: SGIMC Terms of Reference for the next meeting

The Study Group on Integrated Monitoring of Contaminants and Biological Effects (SGIMC) (Co-chairs: Dr Ian M Davies, UK and Dick Vethaak, NL) should be discontinued, having largely completed its work, and that any outstanding tasks be transferred to the Working Group on Biological Effects of Contaminants (WGBEC).

The members of the Study Group wish to thank the current, and past, chairs for their guidance through the series of meetings of SGIMC, and previously WKIMON.

Annex 3a: Recommendations

Recommendations from SGIMC 2011	Adressed to
1. That the documents listed in Table A3/1 below be forwarded to OSPAR in response to OSPAR Request 2008/8 with recommendation for adoption.	ACOM
2. That the documents listed in Table A3/2 below be progressed towards publication in the ICES TIMES series.	WGBEC link with ICES TIMES editor
3. That the compilation of advice provided or updated through SGIMC be published in the ICES Cooperative Research Report series. A Category 1 draft Resoultion is included at Annex 18	ICES Secretariat
4. That the outstanding tasks from the SGIMC work programme listed in Table A3/3 below be transferred to the WGBEC work programme.	ACOM
5. That the seawater contaminants section of the ICES database should be reviewed, in collaboration with the ICES DataCentre, and updated, giving particular attention to parameter fields, station names/locations, gross errors (e.g. units), uncertainty in identification of contaminants, and the opportunity for improving the QA of data being submitted, for example through automated checking of data at the time of entry, with a view to the seawater data becoming more available for assessment in MSFD and other contexts. That this task be included in the MCWG work programme.	MCWG
6. That the updated MIME document (MIME 2010, Annex 7, Starting point for biological effects relevant to good environmental status in the OSPAR area) included in this report as Annex 24 be sent to OSPAR Secretariat (Andrea Weiss).	ACOM
7. That the text in Sections 11 and 12 of this report be offered to ICES MSFDSG as a summary of the links between the work of SGIMC and MSFD GES. The co-chairs of SGIMC can provide further details on request.	ACOM
8. That the TIMES ICES protocol and corresponding Background Document on lysosomal stability (neutral red assay) should be amended/extended to include some practical details and illustration. Intersessional work should lead to a new draft being presented to WGBEC 2012 for review.	WGBEC
9. That ICES DataCentre should undertake work, with the support of members of WGBEC, to enhance the capability of DOME to accept data on an expanded range of biological effects data, as presented in Annex 26 to the SGIMC 2011 report.	ICES Secretariat, WGBEC
10. That SGIMC be discontinued, as it has completed the vast majority of its work progamme.	ACOM, OSPAR

Table A3/1. Documents recommended to be forwarded to OSPAR as ICES advice.

Annex number	Document
4	Background Document on Protocols for extraction, cleanup and solvent exchange methods for small-scale bioassays.
5	Background Document on Intersex (ovotestis) measurement in marine and estuarine fish
6	Background Document on Supporting parameters for biological effects measurements in fish and mussels
7	Background Document on Acetylcholinesterase assay as a method for assessing neurotoxic effects in aquatic organisms
8	Background Document on Histopathology of mussels *Mytilus* sp. for health assessment in biological effects monitoring
9	Background Document on Micronucleus assay as a tool for assessing cytogenetic/DNA damage in marine organisms
10	Background Document on Comet assay as a method for assessing DNA damage in aquatic organisms
11	Background Document on Sediment seawater elutriate and pore-water bioassays with early developmental stages of marine invertebrates.
12	Background Document on Sediment seawater elutriate and pore-water bioassays with copepods (*Tisbe*, *Acartia*), mysids (*Siriella*, *Praunus*), and decapod larvae (*Palaemon*).
13	Background Document on Whole sediment bioassays with amphipods (*Corophium* sp) and *Arenicola marina*
14	Background Document on DNA adducts, including BAC responses and EAC-equivalent Assessment Criteria
15	Background Document on in vitro DR-Luc/DR-CALUX® bioassay for screening of dioxin-like compounds in marine and estuarine sediments
16	Background document on Metallothionein (MT) in blue mussel (*Mytilus edulis*, *Mytilus galloprovincialis*)
17	Background Document on Water bioassays (to exclude *in vitro* bioassays)
18	Background Document on Externally visible fish diseases, macroscopic liver neoplasms and liver histopathology
20	Updating Annexes 9 and 10 to the SGIMC 2010 report, i.e. updating a) Technical Annex on sampling and analysis for integrated chemical and biological effects monitoring in fish and shellfish b) Technical Annex for Mussel (*Mytilus* sp.) OSPAR Integrated Monitoring
21	DRAFT Guidelines for the Integrated Monitoring and Assessment of Contaminants and their effects
22	Biological effects techniques relevant to the ecosystem components for integrated monitoring and assessment of chemical and biological effects data. Status regarding availability of Baclground Documents, assessment criteria, and quality assurance
23	Assessment criteria for biological effects measurements

Table A3/2. Draft TIMES documents forwarded to TIMES editor with recommendation for publication, after resolution of comments raised.

1) Draft ICES-TIMES document: Biological effects of contaminants: Receptor H4IIE-Luciferase (DR-LUC) cell bioassay for screening of dioxins and/or dioxin-like compounds in environmental samples. C.A. Schipper, P.E.G. Leonards, H.J.C. Klamer, K.V. Thomas, A.D. Vethaak.
2) Draft ICES-TIMES document: Histopathology of mussels *Mytilus sp.* for health assessment in biological effects monitoring. J Bignall, S Feist, M Caraville, I Margomez, A Villalba.
3) Draft ICES-TIMES document: Protocols for extraction, cleanup and solvent exchange methods for small-scale bioassays. Hans Klamer, Knut-Erik Tollefsen, Steven Brooks, John Thain.

Table A3/3. Outstanding tasks identified by SGIMC 2011, and recommended for transfer to the WGBEC work programme.

1) Further progress the work on *in vitro* bioassays for endocrine disruption, for example through the development of a Background Document, with a view to OSPAR adopting these assays as components of the JAMP specific biological effect guidelines for estrogenic compounds and the overall integrated chemical-biological effects monitoring approach.
2) The TIMES ICES protocol for the Neutral Red Retention assay should be amended/extended in response to the WKLYS 2010 with some practical details and illustrations, incorporating feedback requested from D. Lowe and A. Kohler (ICES TIMES authors). It was agreed to conduct it through contact with ICES TIMES authors, to be presented to WGBEC 2012.
3) The completion of TIMES documents on mussel histopathology, DR-Luc and extraction protocols.

Annex 4. Technical Annex: Protocols for extraction, cleanup and solvent exchange methods for small-scale bioassays

Hans Klamer[1], Knut-Erik Tollefsen[2], Steven Brooks[2] John Thain[3]

[1]. Centre for Water Management. PO Box 207, 9750 AE Haren, the Netherlands. e-mail: hans.klamer@rws.nl

[2]. Norwegian Institute for Water Research (NIVA). Gaustadalleen 21, 0349 Oslo, Norway.

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

Contents

1. Introduction

The aims of this document are as follows:

- To produce standardized protocols for bioassay extractions;
- To enhance consistency of applications between laboratories;
- To ensure applicability throughout OSPAR maritime area, including in estuarine waters;
- To ensure comparability of reported data for assessment purposes.

History:

- This document has been developed from a previous review, and relates particularly to background documents on water and sediment bioassays and *in vitro* bioassays prepared by ICES expert groups WGBEC and SGIMC. This paper describes a recommended methodology for extraction protocols for use of small-scale *in vitro* and *in vivo* bioassays

Scope:

- This procedure will be used to provide samples for measurements of toxicity in environmental samples and assessment of their potential environmental risk. Other applicable approaches include Toxicity Identification Evaluation (TIE)/ Effects Directed analysis (EDA), and toxicity tracking of effluent and produced water discharges.
- Extraction of aqueous, solid and fish bile samples.
- Preparation of extracts for *in vivo* bioassays including: Mussel and Oyster embryo, Tisbe, Daphnia, Nitocra, Acartia, Sea urchin embryo, fish embryo, algal growth, algal PAM, macrophyte germination);
- Preparation of extracts for *in vitro* bioassays (*e.g.* Microtox, Mutatox, YES, YAS, DR/ER/AR-CALUX, TTR, *umu*-C, Ames-II, fish cell lines).

2. Extraction protocols

In this chapter, extraction protocols will be presented covering a range of types of sample: solid, aqueous or fish bile. Depending on the bioassay that will be used, differences in extraction solvent and, in particular, sample cleanup (Section 2.5), may be applied.

Klamer *et al*., 2005, proposed the following operational definitions of solid and aqueous samples:

- *solid samples*: particulate material, sediments, sludges, aerosols, suspended solids, and soils;
- *aqueous samples*: surface or deep waters, wastewater, sediment pore water, potable water, rain, snow, ice.

Before detailed protocols are presented, the basic layout of each extraction and cleanup protocol is given below.

Solid samples

Protocol steps	Comment
1. Sample preparation	Sample sieved when necessary (e.g. sediment), dried and homogenized.
2. Extraction of crude sample	Accelerated Solvent Extraction (ASE) or Soxhlet extraction. Solvents: dichloromethane (DCM) or hexane with methanol or acetone as modifier.
3. Concentration of crude extract	Automatic (e.g. Turbovap® or manual) concentration to smaller volume, typically less than 5 mL. Remove co-extracted water if necessary
4. Cleanup of crude extract	Gel Permeation Chromatography (GPC) with DCM for broad-spectrum contaminant profiling. Reversed or normal phase HPLC for more selectivity. Sulphur removal may be necessary.
5. Concentration of cleaned extract	Automatic (e.g. Turbovap® or manual) concentration to smaller volume, typically less than 1 ml. Final test solvent (e.g. DMSO or methanol may be added as keeper.)

Aqueous samples

Protocol steps	Comment
1. Sample preparation	Sample filtered and/or pH-adjusted when necessary.
2. Extraction of crude sample	Solid Phase Extraction (SPE) with resin (e.g. XAD) or cartridge containing adsorbents (C8, C18, lichrolut™, POCIS)
3. Concentration of crude extract	Automatic (e.g. Turbovap® or manual) concentration to smaller volume, typically less than 5 ml
4. Cleanup of crude extract	Gel Permeation Chromatography (GPC) with DCM for broad-spectrum contaminant profiling. Reversed or normal phase HPLC for more selectivity
5. Concentration of cleaned extract	Automatic (e.g. Turbovap® or manual) concentration to smaller volume, typically less than 1 ml. Final test solvent (e.g. DMSO or methanol may be added as keeper.)

Fish bile samples

Protocol steps	Comment
1. Sample preparation	Thaw on ice.
2. Pretreatment of crude sample	Deconjugation with a mixture of water, sodium acetate buffer and beta-glucuronidase–arylsulfatase. Total volume typically 1.5 ml.
3. Extraction of pretreated sample	pH treatment with 100 µl 1N HCl, extraction with 2 mL ethyl acetate.
4. Cleanup of crude extract	Precipitate any formed protein using isopropanol. Centrifugate. Repeat extraction.
5. Concentration of extract	Manual concentration to dryness of combined ethyl acetate phases using N2, solvent exchange into 50 µl DMSO.

2. 1 Protocol for extraction of dried, solid samples with Accelerated Solvent Extraction (5 g sample). Steps are numbered S.1, S.2, etc.

S.1. Assemble the ASE cells. Add a small layer of dried silica until cellulose filter is no longer visible;

S.2. Weigh approximately 5 gramme dried sample in the ASE cells (weighing accuracy mass ± 0.1%);

S.3. Fill the ASE cells with dried silica and compact the content of the cells with the engraver pen. Close the cell and firmly twist the end-cap on the ASE cell;

S.4. Extract the sample using the following ASE settings:

Solvent	Pressure	Temp	Preheat time	Static time	Flush volume	Purge time	Static Cycles
	[psi]	[°C]	[min]	[min]	[ml]	[sec]	
Hexane/Acetone 9:1 v:v	2000	100	5	5	60	90	3
DCM or DCM/modifier**	2000	45–100*	5	5	60	90	1–3*

*** Set temperature to 45 to 50°C and # of cycles to 3 for use with ER-CALUX and similar tests.**

**** methanol or acetone.**

S.5. If water is co-extracted, dry the extract using anhydrous sodiumsulphate. Rinse with solvent. Evaporate the extract (until approximately 2–5 ml is left), in an automatic or manual set-up;

S.6. Proceed to solvent exchange (Section 3.3) or store the crude extract at -20°C until further use.

2.2 Protocol for extraction of aqueous samples with Solid Phase Extraction devices. Steps are numbered A.1, A.2, etc.

Extraction

A.1. Assemble the SPE cartridge. For samples up to 20L, a single column set-up is used. A Teflon tube is filled with glass wool to remove particulates and then the SPE columns are filled with methanol and attached in series with the C8 column first, followed by the ENV+. For 100L samples, a multi column system is used, where six Teflon tubes are set up as with the single column system, but then attached to a manifold, allowing one sample to pass through all six columns simultaneously.

A.2. Set up the pressure system. From the pressure source, the air line passes through an air filter and then into a manifold. This allows for more than one vessel to be run at any given time, and also the airline diameter to be reduced. This line is then connected to the pressure vessel via a needle valve, ensuring the correct inlet/outlet is used (the inlet for the air is just a hole in the top of the vessel; the outlet has a pipe which goes to the bottom). From the outlet, another tube is connected which goes into the top of the single column system or manifold for the multi column set-up.

A.3. Once the pressure lines are set up, the air line can be switched on, ensuring first that all needle valves are closed. The pressure should be no greater than 2 bar. The valve can then slowly be opened to allow a flow of approximately 40 ml min^{-1} through the columns.

A.4. Once the entire sample has passed through the column, allow the columns to dry by passing air through them. Label each column with sample site. Wrap in hexane rinsed foil and store in a freezer at -20 °C. Samples can be stored in the freezer for up to two months before elution.

Elution

A.5. Remove columns from the freezer and, while they are thawing, solvent rinse two glass sample collection tubes per column. Label the sample tubes.

A.6. In a fume cupboard, place the columns in the vacuum unit, with a Teflon tap. Fit a length of vacuum-proof hose to the unit, attaching the other end to a waste barrel. Another length of hose should run from the barrel to a vacuum pump.

A.7. Wash the columns with 10 ml RO or milliQ water. This will help to remove salt from saline samples.

A.8. Ensure columns are dry by sucking under vacuum for 10 min, or until there is no visible water dripping through the columns (whichever is longer).

A.9. Place a labelled collection tube under each column in a rack.

A.10. Elute each column with 10 ml DCM. Add 1 ml DCM to the column and allow to soak for 1 min with the tap closed. Open the tap and allow the solvent to drip through. Repeat this three times with 1 ml, 4 ml and 4 ml DCM respectively.

A.11. Remove the tube from under each column and replace it with a clean one. Repeat Section A. 7 with methanol.

A.12. Reduce the samples in volume to approximately 1 ml, and then combine the four fractions of each sample (C8 DCM, C8 methanol, ENV+ DCM, EMV+ methanol). For 100 l samples, there will be six of each type of column. Combine all fractions.

A.13. There may be some water in the samples. This will form a layer or droplets in the DCM. If this is the case, take a glass column and packed with hexane washed anhydrous sodium sulphate. Add the samples to the top of the column. Elute with 5 ml DCM and collect in a labelled tube.

A.14. Blow down each extract to approximately 5 ml using e.g. a Turbovap at 30°C, 5 psi oxygen free nitrogen. From this point, aliquots of samples can be solvent exchanged into the appropriate solvent depending on the assay in question (see paragraph 3.3). Transfer sample into a glass Store extracts in freezer at -20°C. Samples can be stored for a maximum of one year.

2.3 Protocol for extraction of fish bile samples. Steps are numbered B.1, B.2, etc.

The extraction procedure described below is taken from the work by Legler *et al.*, 2002.

Extraction

B.1. Thaw bile samples.

B.2. Transfer 100 µl of bile to glass test tubes.

B.3. Add 700 µl sodium acetate buffer (100 mM, pH 5.0 at 37°C), followed by 600 µL distilled water and 40U of β-glucuronidase–arylsulfatase (from *H. pomatia*).

B.4.4 Incubate tubes overnight (17–18 h) in a water bath (37°C, gentle shaking).

3. Cleanup

3.1 Broad-spectrum cleanup

Cleanup procedures are applicable to all crude extracts. However: the user has to choose between two fundamentally different cleanup principles: *broad spectrum* or *target* cleanup.

Gel Permeation Chromatography with DMC as eluting solvent provides a sample with contaminants having a *broad spectrum* of physico-chemical properties. GPC separates on molecular volume and may therefore be used to easily remove, *inter alia*, humic acids and lipids. GPC column material, however, also has a secondary retention mechanizm, based upon electronic interaction between the column material and the extracted compound. This secondary mechanizm is used for removal of molecular sulphur (as S8) from the crude extract, using DCM as eluting solvent. GPC cleanup requires careful calibration using a series of different compounds. This type of cleanup has successfully been applied to very different *in vitro* bioassays: Microtox, Mutatox, (anti)DR-CALUX, (anti)ER-CALUX, *umu*-C (e.g. by Klamer *et al.*, 2005 and Houtman *et al.*, 2004).

C.1. Set up of GPC equipment. For semi-preparative cleanup, large-diameter columns may be used in series, e.g. polystyrene-diphenylbenzene copolymer columns (PL-gel, 5 or 10µm, 50Å, 300x25 mm or 600x7.5 mm, preferably in a thermostatic housing at 18ºC, with a PL-gel pre-column 5 or 10 µm 50x7.5mm). Use an HPLC pump with 10 ml/min dichloromethane as eluens.

C.2. Calibration. When necessary, determine the elution profile of individual compounds by injection of 2 ml of standard solutions (concentration 0.5–10 mg/L) and assessment of retention times at peak maximum and peak shape.

C.3. Set-up of the fraction collector. As a rule of thumb, the elution of *parathion* may be used to trigger the start of the collection of the cleaned sample, while the collection is stopped just before sulphur (as S8) elutes (elution of the extract is monitored using a UV detector at 254 nm.) This range, however, should be carefully monitored using a several reference compounds (in DCM solution). Examples of compounds that may be included in this mixture are: sulphur, pyrene and ethyl-parathion. Depending on the particular application, other reference compounds may be needed (see e.g. Houtman *et al*., 2004).

C.4. Inject crude extract in 200–2000 μL batches, depending on the capacity of the GPC column (semi prep 25 mm column may be loaded with 2000 μL). Concentrate the collected sample fractions, proceed to solvent-exchange (see paragraph 3.3) or store at -20ºC until further use.

3.2 Selective or dedicated cleanup

Selective cleanup using adsorption chromatography (e.g. reversed or normal-phase liquid chromatography, with or without modifying additives like KOH, $AgNO_3$).

DR CALUX

The clean-up of crude extracts for DR CALUX measurements can be done with an acid silica column combined with TBA sulphur clean-up. The protocol for the DR-CALUX cleanup is as follows:

TBA sulphite solution

Wash a 250 ml separation-funnel with hexane, fill the funnel with 100 ml HPLC water and dissolve 3.39 grammes TBA

Rinse the solution three times with 20 ml hexane.

Dissolve 25 gramme sodium sulphite in the washed solution.

Store the solution in a dark bottle (Maximum storage time 1 to 2 weeks).

Sulphur clean-up

Add 2.0 ml TBA-sulphite solution and 2.0 ml isopropanol to the extract, mix for 1 minute on a vortex. Sulphur clean-up is complete if precipitation is visible. Add an extra 100 mg sodium sulphite if no precipitation is present and mix during 1 minute on a vortex. Repeat the addition if necessary.

Add 5 ml of HPLC-grade water, mix for 1 minute on a vortex.

Let the layers separate during approximately 5 minutes, transfer the hexane layer to a clean collection vial.

Add 1 ml hexane to the extract and mix during 1 minute on the vortex. Let the layers separate layers en transfer the hexane layer to the clean collection vial. Repeat this step. Evaporate the hexane until approximately 1 ml is left.

Acid silica clean-up

Prepare a solution of hexane/diethylether (97/3; v/v)

Place a small piece of glass wool in a separation. As the performance of the following steps is column-dependent (see Annex 1 for column layout).

Fill the column with 5 grammes of 33% silica and tremble the cells with the engraver pen. Add 5 grammes of 20% silica and tremble the column once more. Add a small amount of dried sodium sulphate to the top of the column.

Elute the column with 20 ml hexane/diethylether solution.

Bring the extract on the column as soon as the meniscus reaches the sodium sulphate. Wash the collection vial of the extract twice with approximately 1 ml hexane/diethylether solution.

Place a clean collection vial under the column and elute the column with 38 ml hexane/diethylether.

Evaporate the hexane until less than 1 ml is left.

Proceed to solvent exchange (see below, 3.3)

ER CALUX

This section describes the cleanup of deconjugated fish bile extract for use in the ER-CALUX assay. Steps are numbered B5, B6, etc, referring to the fish bile extraction procedure above.

B.5. Add 100 µl 1N HCl to each glass test tube containing the deconjugated bile sample (see B.1., above). Stir well (Vortex).

B.6. Add 2 ml ethyl acetate to each test tube. Vortex for 1 min, followed by centrifugation for 5 min at 3800 rpm.

B.7. Remove the ethyl acetate fraction using a Pasteur pipette and transfer this to a new test tube. If protein formation is observed between the water and solvent phases, precipitate this protein by adding 500 µl of isopropanol after centrifugation.

B.8. Repeat steps B6 and B7 three times, with exception of the isopropanol-step.

B.9. Concentrate the collected ethyl acetate fractions and evaporate to a small drop under a gentle N_2 gas flow at 37°C.

B.10. Transfer the concentrated extract to a conical glass vial.

B.11. Rinse the glass test tube three times with ethyl acetate, and transfer the rinses to the conical vial.

B.12. Evaporated the ethyl acetate to dryness at 37°C under a gentle stream of nitrogen.

B.13. Proceed to solvent exchange (see below, 3.3).

3.3 Solvent exchange

Klamer and van Loon (1998) and Bakker *et al*. (2007) developed criteria and evaluated co-solvents for bioassays. The ideal co-solvent or carrier solvent used for ecotoxicity testing should meet the following criteria: (1) effective: sufficiently high solubility of target compounds, (2) water-miscible: the carrier solvent must be water-miscible, and (3) non-toxic: the carrier solvent should have little or no adverse effects on test organisms or cells at typical test concentrations in aqueous media (usually 0.1% v/v). The authors tested ten different solvents, with the following final ranking for the first five solvents:

Solvent	Final rank
Dimethylsulfoxide (DMSO)	1
2-Propanol	2
Acetone	2
Methanol	4
Ethanol	5

The following general solvent-exchange protocol is applicable to all five solvents:

1) Transfer the remaining cleaned extract to a conical vial and evaporate until a small meniscus of it is left (approximately 20 µl).

2) Wash the collection vial twice with at least 0.5 ml DCM or other appropriate solvent, and transfer this to the conical vial (evaporate between washes; do not let the vial fall dry).

3) Evaporate the extract until the meniscus reaches the bottom of the conical vial and then add 50 µl of co-solvent.

4 Preparation of extract test dilutions for *in vivo* bioassay

The following procedure should be employed when using the prepared extract(s) for standard *in vivo* bioassay testing. This approach is focused on microscale tests with a typical test volume of no more than 5 ml.

Once prepared using the above extraction procedure, the extract must be stored at minus 20°C degrees C until bioassayed, and should not be stored for longer than twelve weeks.

A stock solution is made with the concentrated extract using the appropriate dilution water (i.e. aerated seawater or freshwater), from which an appropriate series of concentrations will be prepared. The preparation of the stock solution is important: typically 5 ml of extract in solvent is concentrated by evaporation to 20 ul. The concentration series must be made up on the day of testing and the ratio between the concentrations should not exceed 2.2 (usually log).

The stock solution must be shaken vigorously, stirred on a magnetic stirrer for at least 30 minutes or placed in the ultrasonic bath for ten minutes to ensure that all of the chemical/compound(s) within the extract are in solution. The solvent concentration in the final test solution must not exceed 0.1 ml/L with all test concentrations containing the same amount of solvent. A solvent control of the appropriate solvent at the same concentration must be used. All controls and test concentrations must have at least three replicates. The salinity, pH, temperature and dissolved oxygen concentration of the test concentrations must be checked prior to testing and corrected to within the specific parameters of the bioassay as appropriate.

Where possible, the concentrations selected should cover a range from low concentrations with no effect on the test organism relative to the control, intermediate effects, and complete 100% effect. Clearly, this may require an initial sighting test prior to conducing a definitive test. This will enable the calculation of the NOEC, LOEC and EC_{50} values with greater precision.

Preparation of extracts for cell lines

DMSO is the recommended solvent for use with cell line exposures. The concentration of solvent in the final test volume should not exceed 1% (v/v).

Confounding factors

For small test volumes, evaporation of the test solution can be a problem as the volume to air-surface ratio is high, and particularly if the test temperature is high e.g. >15 degrees C. Precautions should be taken to avoid evaporation and also the contaminant crossover that can occur in multiwell plates. In this respect, a short exposure time is desirable: Test duration is typically not greater than 48 h, although there are some exceptions, such as bioassays with algae which may need a 72 hr exposure.

The surface area to volume ratio of the test container is high and some contaminants may preferentially adhere to surfaces such as polystyrene. For this reason, glass test containers should be used in preference to plastic.

5. Conclusions

Whatever the matrix, extraction procedures generally produce small volumes and therefore small-scale bioassay procedures are required for testing. In most cases, the recommended procedures are adapted from well-established protocols. The choice of test species will depend on the purpose of the study and the availability of test organism.

Bioassays frequently used for testing extracts are shown below:

	Test organism	Test volume (ml)	Number of organisms/cells per test vessel	Reference
In vivo	Mussel embryo	1–5	50 per ml	ASTM724
	Oyster embryo	1–5	50 per ml	ASTM724
	Sea urchin	1–5	40 per ml	ASTM1563
	Microalgae (freshwater and seawater)	1–5	5x106 cells /L	ISO8692, ISO10253
	Macrophyte germination	1–5	500-1000 zygotes per ml	Brooks *et al.*, 2008
	Daphnia	1–5	1 per test vessel	ISO6341
	Acartia / Nitocra	5	5 per test vessel	ISO 14669
	Tisbe	5	5 per test vessel	ISO14669
	Fish embryo	2–5 ml	1 per 2ml test vessel	OECD draft guideline
In vitro	YES, YAS, anti-YES, anti-YAS	200 µl	0.8 x 106 cells/ ml	Tollefsen *et al.*, 2007
	ER calux	200 µl	5-10 x 105cells/ml	Legler *et al.*, 2003
	Primary cell cultures	200 µl	5x105 cells/ml	Tollefsen *et al.*, 2003
	Cell lines	200 µl	5-10 x 105cells/ml	

Matrix	Procedure	Bio-assay	Reference
Sediment	ASE, DCM, acetone	ER-Calux	Houtman *et al.*, 2007

In all of the above test methods, appropriate reference materials should be tested as stated in the specific test protocols.

6. References

Bakker, J.F., M.J. Bulzunce-Segarra, R. Castro, M van de Heuvel-Greve, H.J.C. Klamer, W. Brack, R. Altenburger, V. Poulsen, K.V. Thomas, P.E.G. Leonards. 2007. Effect directed analysis and toxicity identification evaluation. In: Sediment quality and impact assessment of pollutants. Sustainable Management of Sediment Resources. Chapter 5. D. Barceló, and M. Petrovic (Eds.), Elsevier, Amsterdam, 333 pp., ISBN: 9780-444-51962-7.

Besselink, H.T., C.A. Schipper, H.J.C. Klamer, P.E.G. Leonards, H.J.M Verhaar, E.C. Felzel, A.J. Murk, J.E. Thain, K. Hosoe, G. Schoeters, J. Legler, and A. Brouwer. 2004. Intra- and inter-laboratory calibration of the DR CALUX (T) bioassay for the analysis of dioxins and dioxin-like chemicals in sediments. *Environ. Toxicol. Chem.* 23(12), 2781–2789.

Houtman, C.J., P.H. Cenijn, T. Hamers, M.H. Lamoree, J. Legler, A.J. Murk, and A. Brouwer. 2004. Toxicological profiling of sediments using in vitro bioassays, with emphasis on endocrine disruption. *Environ. Toxicol. Chem.* 23(1), 32–-40.

Houtman CJ, Leonards PEG, Kapiteijn W, Bakker JF, Brouwer A, Lamoree MH, Legler J, Klamer HJC. 2007. Sample preparation method for the ER-CALUX bioassay screening of (xeno-)estrogenic activity in sediment extracts. Marine Pollution Bulletin, 54.

ISO. 1999. ISO 14669 Water Quality - Determination of acute lethal toxicity to marine copepods (*Copepoda*, *Crustacea*).

Klamer, H.J.C., K. Thomas, K-E. Tollefsen. 2005. *Extraction procedures applicable to small-scale bioassays*. Discussion paper, Working Group on Biological Effects of Contaminants.

Klamer, H.J.C., P.E.G. Leonards, M.H. Lamoree, L.A. Villerius, J.E.Åkerman, J.F. Bakker. 2005. A chemical and toxicological profile of Dutch North sea surface sediments. *Chemosphere* 58, 1579–1587.

Legler J, Leonards P, Spenkelink A, Murk AJ. 2003. *In vitro* biomonitoring in polar extracts of solid phase matrices reveals the presence of unknown compounds with estrogenic activity. Ecotoxicology, 12: 239–249.

Legler, J., A. Jonas, J. Lahr, A.D. Vethaak, A. Brouwer and A.J. Murk. 2002. Biological measurement of estrogenic activity in urine and bile conjugates with the *in vitro* ER-CALUX reporter gene assay. *Environ. Toxicol. Chem.* 21(3), 473–479.

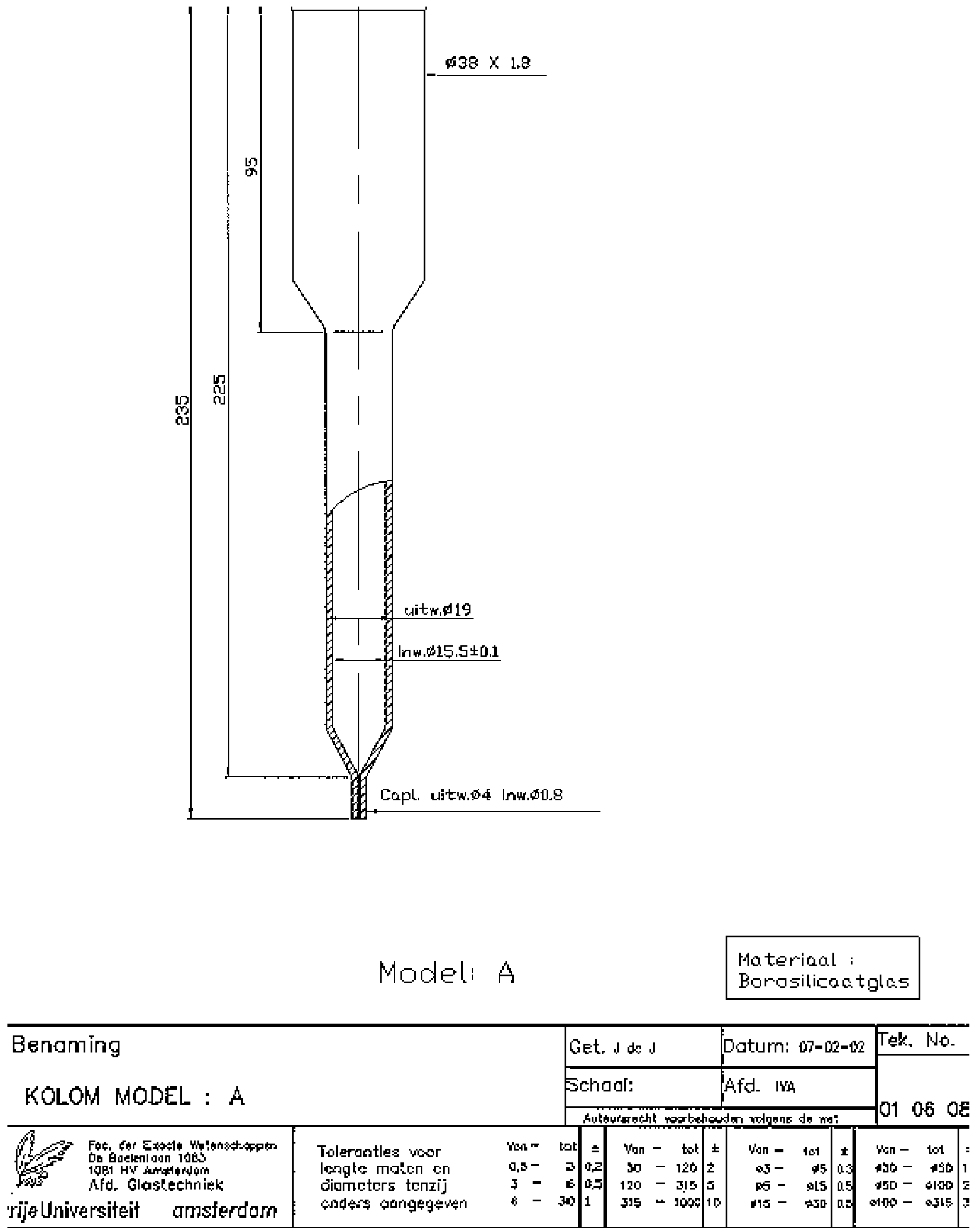
Leonards, P.E.G., I. van de Veen, M. Lohman, E. Felzel en S. Man. 2002. *Comparison of extraction and cleanup methods for use with the DR-CALUX assay*. RIVO, report C048/02, (*in Dutch*).

Tollefsen, K.-E., Mathisen, R. and Stenersen, J. 2003. Induction of vitellogenin synthesis in an Atlantic salmon (*Salmo salar*) hepatocyte culture: a sensitive *in vitro* bioassay for the (anti)estrogenic activity of chemicals. *Biomarkers*. 8, 394–407.

Tollefsen K-E, Harman C, Smith A, Thomas KV. 2007. Estrogen receptor (ER) agonists and androgen receptor (AR) antagonists in effluents from Norwegian North Sea oil production platforms. Original Research Article. Science of the total environment, 386: 134–144.

Thomas, K.V., J. Balaam, M. Hurst, Z. Nedyalkova, and O. Mekenyan. 2004. Potency and characterization of estrogen-receptor agonists in United Kingdom estuarine sediments. *Environ. Toxicol. Chem.* 23(2), 471–479.

Appendix 1: Lay-out of borosilica column for use with acid-silica cleanup



Annex 5: Background Document: Intersex (*ovotestis*) measurement in marine and estuarine fish

Compiled by G D Stentiford[1]

Summary

1) *Applicability across OSPAR maritime area*

The presence of susceptible host species utilized in monitoring programmes in marine and estuarine habitats of the OSPAR region make this an applicable measurement in field programmes. The requirement for the sampling of testis from male fish captured in such programmes and the assessment of these tissues by histology can be aligned with the sampling of other tissues currently assessed for fish diseases work (e.g. for liver cancer assessment). The epidemiological basis for the sampling of fish for intersex measurement is therefore aligned with other field sampling programmes for fish health.

2) *Status of quality assurance*

Formal QA for the measurement of intersex in marine and estuarine fish has not been carried out under existing programmes (such as BEQUALM) but published methods are available for the grading of intersex severity in flatfish collected from monitoring programmes. These methods would be directly applicable to QA programmes. The sampling of materials from epidemiological relevant numbers of animals is also well characterized in the literature and is outlined in this document.

3) *Influence of environmental variables*

Although sex determination can be influenced by environmental factors and age, there has been an historic linkage between sites with the highest prevalence of intersex fish, biomarkers for exposure to endocrine disrupting chemicals (e.g. vitellogenin), and anthropogenic contaminants known to elicit development of ovotestis in a range of test species.

4) *Assessment of thresholds*

Threshold assessment to indicate an impacted site has not previously been discussed for measurement of intersex (ovotestis) in male fish. However, based upon the reported prevalence of the condition in marine and estuarine fish from the OSPAR region, and the constraints inherent with the sampling of large populations for health effects, it would appear that a threshold of 5% prevalence (in external males) may be used to indicate impact. The epidemiological basis for this is discussed in this document.

5) *Proposals for assessment tools*

Given background data on quality assurance techniques for intersex measurement, it seems appropriate to propose a two-tier assessment tool. Tier 1 consists of an individual sample grading system for intersex severity based on the methodology presented by Bateman *et al*. (2004). Tier 2 consists of apparent prevalence estimates based upon a sampling regime designed to detect a 5% prevalence of intersex at 95% confi-

[1] Centre for Environment, Fisheries and Aquaculture Science (Cefas), Weymouth Laboratory, Dorset DT4 8UB, UK.

dence. Both of these tools can be combined to provide a population-level and individual-level assessment tool for the condition. Because intersex prevalence is likely to be negligible in non-impacted populations, survey designs are likely to be similar to that for fish disease measurement, whereby detection is based upon diseases present in a population at 5% prevalence (95% confidence). In this way, >5% prevalence would be considered the cut-off point for definition of an impacted population. The use of cohort-matching, similar to that for assessment of liver pathology in flatfish, is recommended to remove any confounding effects of age on intersex prevalence (e.g. use of fish of 4 years old) (Stentiford *et al*., 2010).

Assessment of the applicability of intersex measurement across the OSPAR maritime area

In recent years, a significant proportion of research into the biological effects of contaminants in the aquatic environment has been devoted to the study of endocrine disrupting chemicals (EDCs) of anthropogenic origin. EDCs have been widely reported to impair fertility, development, growth and metabolism in a range of animal groups (see Colborn *et al*., 1996). The effects of exposure of fish to such compounds include disturbed maturation and degeneration of the gonads, elevated concentrations of vitellogenin (egg yolk protein) in the plasma of male fish and the presence of intermediate or 'intersex' gonads (Gimeno *et al*., 1996). Using histological analysis, fish with the intersex condition are seen to possess oocytes within their normal testicular matrix (Sharpe, 1997; Bateman *et al*., 2004). Until the early 1990s intersex had only rarely been described from fish in the wild (Jafri and Ensor, 1979; Slooff and Kloowijk-Vandijk, 1982; Blachuta *et al*., 1991). However, the condition has now been detected in several wild freshwater and migratory species, including roach *Rutilus rutilus* (Jafri and Ensor, 1979; Purdom *et al*., 1994; Jobling *et al*., 1998), gudgeon *Gobio gobio* (van Aerle *et al*., 2001), barbel *Barbus plebejus* (Vigano *et al*., 2001), chub *Leuciscus cephalus* (Minier *et al*., 2000), bream *Abramis brama* (Slooff and Kloowijk-Vandijk 1982), white perch *Morone americana* (Kavanagh *et al*., 2002), stickleback *Gasterosteus aculeatus* (Gercken and Sordyl, 2002), coregonids (Mikaelian *et al*., 2002), grayling *Thymallus thymallus* (Blachuta *et al*., 1991) and Atlantic salmon *Salmo salar* (authors' pers. obs.). Furthermore, detection of elevated prevalence of intersex in some estuarine and marine species such as the European flounder *Platichthys flesus* (Allen *et al*., 1999a), Japanese flounder *Pleuronectes yokohamae* (Hashimoto *et al*., 2000), bothid flounder *Bothus pantherinus* (Amaoka *et al*., 1974), common eel *Anguilla anguilla* (Peters *et al*., 2001) and viviparous blenny *Zoarces viviparus* (Matthiessen *et al*., 2000; Stentiford *et al*., 2003) suggest that the effects of anthropogenic EDCs may extend beyond inland river systems to coastal and even offshore waters. This is supported by reports of elevated plasma vitellogenin and ovotestis in male Mediterranean swordfish *Xiphias gladius* (Fossi *et al*., 2001 and De Metrio *et al*., 2003, respectively), and the dab *Limanda limanda* (Scott *et al*., 2007; Stentiford and Feist, 2005, respectively). In terms of species of relevance to the OSPAR region, those in which intersexuality (ovotestis) have been described from marine and estuarine habitats include flounder (Allen *et al*., 1999b; Stentiford *et al*., 2003, Bateman *et al*., 2004), dab (Stentiford and Feist, 2005), viviparous blenny (Stentiford *et al*., 2003; Lyons *et al*., 2004), red mullet (Martin-Skilton *et al*., 2006) and the 3-spined stickleback (Gercken and Sordyl, 2002).

Status of quality assurance techniques for intersex measurement in marine and estuarine fish

Male fish with the intersex condition are seen to possess, at varying degrees of severity, oocytes within the testis; this being regarded as a phenotypic endpoint of endocrine disruption (both natural and anthropogenic) in male fish. Due to the fact that the testis may appear normal from external observations, histological examination of the testis is necessary to identify and grade individual cases of intersex and to estimate prevalence in a population. Intersex has been recorded histologically in all of those species listed above as relevant to marine and estuarine waters from the OSPAR region. It is important to consider quality assurance techniques for intersex measurement at two levels: 1. Individual (grading of intersex severity) and 2. Population (intersex prevalence).

Individual-level grading of intersex (ovotestis)

The most comprehensive assessment of ovotestis severity at the individual level has been presented by Bateman *et al*. (2004) for the European flounder. In this case, the study provided information on the different pathological manifestations of the intersex condition in flounder sampled from various estuarine and coastal waters of the United Kingdom and furthermore, described the development and application of an ovotestis severity index (OSI), calculated for individual histological sections of gonad. The development of this index provides pathologists with a robust tool for the grading of the intersex condition in flounder and potentially other fish species sampled in the OSPAR region.

The study by Bateman *et al*. (2004) utilized samples collected from monitoring programmes around the United Kingdom over a four-year period (1998–2002) and assessed externally classified male flounder of above 15 cm in length. For histology, whole gonads were removed and fixed in a 10% solution of neutral buffered formalin prior to processing to wax using standard protocols. In order to assess the distribution of oocytes throughout the testis, all specimens examined were step-sectioned longitudinally at 0.2-mm intervals throughout tissue at a thickness of 3 to 5 µm, mounted onto glass slides, and stained using haematoxylin and eosin (H and E). Sections were analysed by light microscopy. A total of 56 intersex cases were examined. All gonadal sections were viewed at low magnification using a x10 eyepiece and x10 objective lens, giving a total magnification of x100. Each gonadal tissue section was divided into a variable number of fields of view depending on the size of the sample. The number of fields of view comprising the whole tissue section was then used to construct a virtual grid, with each square on the grid corresponding to a field of view. Only fields of view that contained 100% tissue coverage were included in calculations of the OSI. Each field of the grid was then scored for the presence of oocytes, the distribution of these oocytes, and their stage of development (according to previously published criteria in other fish species). The overall OSI takes into consideration both the oocyte development stages present and their distribution throughout the testis (see Figure 2 from Bateman *et al*., 2004 below).

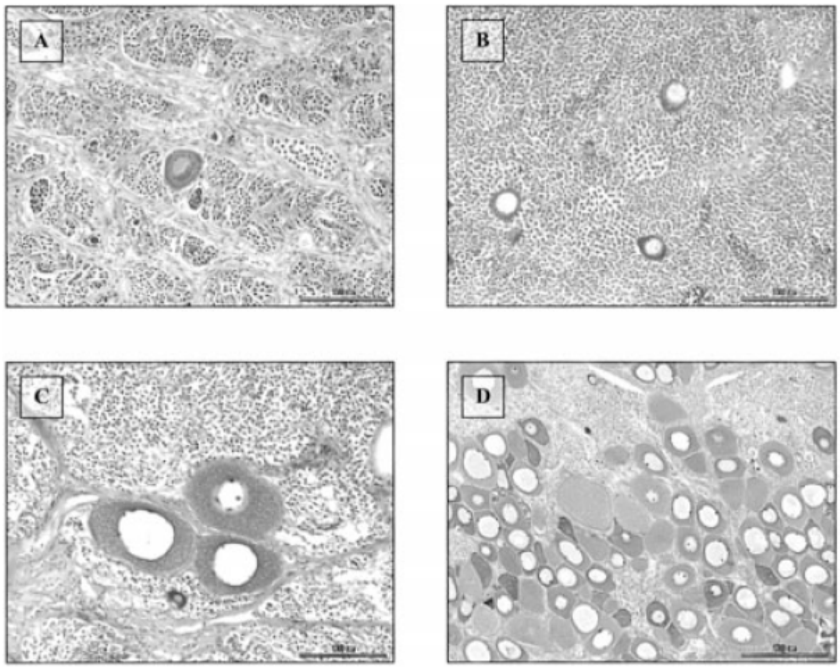


Fig. 2. Oocyte distribution patterns in ovotestis cases. (**A**) Focal, only single oocytes are present within field of view. (**B**) Diffuse distribution, more than one oocyte is present in a field of view and are not closely associated with neighboring oocytes. (**C**) Cluster distribution with more than one, but less than five, closely associated oocytes present within a field of view. (**D**) Zonal distribution, indicated by the presence of more than five closely associated oocytes within a field of view. In this case, oocytes in various stages of maturity can be seen. Haematoxylin- and eosin-stained sections. Scale bars = 100 μm.

In order to calculate the severity of the intersex condition within an individual section of gonadal material, an algorithm was formulated, incorporating the scores for development and distribution of oocytes within individual fields of view. This algorithm allowed for the calculation of the OSI for an individual section of gonad. The OSI was calculated as follows where D1 is the most advanced development stage of oocytes within a field of view (score 1–5), D2 is the distribution of oocytes within a field of view (score 1–4), and X is the total number of fields of view examined.

$$\mathrm{OSI} = \left(\frac{\sum \; [D_1 \cdot D_2]}{X} \right)$$

The OSI is a sum of the severity staging for each field of view in a section of gonad. By dividing this sum by the total number of fields of view in the whole section, the mean ovotestis severity per field of view can be obtained. For intersex flounder, this gives an overall OSI of >0 up to 20 (the maximum score, whereby each field of view contains over five vitellogenic oocytes in a zonal distribution). Testing this scoring system on field collected samples of flounder, Bateman *et al*. (2004) used the OSI scores from each gonad to create a broad grading system of: Absent (OSI = 0), stage 1 (OSI >0–5), stage 2 (OSI >5–10), and stage 3 (OSI >10–20). This was summarized as:

Table 2. Ovotestis ranking by stage based on histological appearance, proportion of fields containing oocytes, and the distribution and developmental stage of oocytes

Severity category	Histology	Proportion of fields of view with oocytes	Distribution and developmental stage of oocyte
Absent (score 0)	Testis structure is normal, with no oocytes present in section.	—	—
Stage 1 (score > 0–5)	Structure of the majority of the testis appears normal.	Generally below 50%	Single or multiple previtellogenic oocytes. Cortical alveolar or fully vitellogenic oocytes rarely are present.
Stage 2 (scores > 5–10)	Regions of the testis are altered, replacement of testicular material with oocytes.	Up to 75%	Majority of oocytes are previtellogenic, present in clusters or zones in high proportion of fields of view. Single or multiple vitellogenic oocytes.
Stage 3 (score > 10–20)	Majority of testis is disrupted, replacement of testicular material with oocytes in various stages of development.	Above 75%	Associated previtellogenic or vitellogenic oocytes through majority of section.

1) Population prevalence of intersex (*ovotestis*).

The second level of assessment of intersex (*ovostestis*) in marine and estuarine fish from the OSPAR region requires an indication of prevalence (or the total number of cases in the population, divided by the number of individuals in the population). Because it is problematic to define the number of individuals in a wild population of marine or estuarine fish, the estimation of prevalence (or so-called apparent prevalence) is therefore carried out by sampling a statistically significant number of animals from a population exceeding a presumed size (e.g. >10,000 individuals). The size of the sample required will also depend on necessity of detecting a given prevalence (e.g. 1%, 2%, 5%, etc.) and the confidence level of detecting this prevalence (e.g. 90%, 95%, 99%). Whereas the majority of studies examining the presence of intersex in wild populations do not appear to have followed statistical guidelines relating to the sampling of wild populations (e.g. see Simon and Schill, 1984), it is perhaps relevant that the approach to monitoring for intersex should follow that outlined in the chapter for fish diseases and as reported in studies such as those of Stentiford *et al*. (2009, 2010). In this context, sampling is designed to detect a disease prevalence of 5% at a confidence level of 95%. Using these figures, 59 individuals should be sampled if the population size is assumed to be 10 000 individuals. By using the same confidence of detecting lower prevalence of intersex, sample sizes would need to increase to 148 individuals (for 2% prevalence) and 294 individuals (for 1% prevalence). It should be noted however that where populations exceed 100 000, 500 000 or 1 000 000 individuals, sample sizes required to detect a 5, 2 and 1% prevalence at 95% confidence are considerably larger (597, 1494 and 2985 individuals, respectively). Clearly cost and conservation limitations will relate to most monitoring schemes so that these latter numbers become somewhat unfeasible. It is for this reason that presuming a population size of 10,000 and sampling to detect 5% prevalence at 95% confidence has been chosen for much of the fish disease work (Feist *et al*., 2004).

When considering apparent prevalence of intersex in a population of marine or estuarine fish sampled from the OSPAR region, it is useful to consider the reported prevalence range for the condition in relevant species. Available data for the key monitoring species are as follows:

Flounder (*Platichthys flesus*)	Up to 20% (Allen *et al*., 1999a) Up to 9% (Allen *et al*., 1999b) Up to 8% (Minier *et al*., 2000) Up to 8.3% (Stentiford *et al*., 2003)
Viviparous blenny (*Zoarces viviparus*)	Up to 27.8% (Gercken and Sordyl, 2002) Up to 25% (Stentiford *et al*., 2003)
Dab (*Limanda limanda*)	Up to 14.3% (Stentiford and Feist, 2005)
Red mullet (*Mullus barbatus*)	Up to 14.3% (Martin-Skilton *et al*., 2006)
Stickleback (*Gasterosteus aculeatus*)	Up to 12.5% (Gercken and Sordyl, 2002)

2) Given the fact that intersex appears to exist at a range of between 0 and 27.8% in different monitoring species, a sampling regime based upon detection of 5% prevalence at 95% confidence appears appropriate. Furthermore, multi-site work in several species (e.g. flounder and dab by Stentiford *et al*., 2003 and 2005, respectively) has demonstrated that intersex is detected at some sites and not at others when this regimen is utilized. This indicates that intersex, if present, occurs at below 5% at these latter sites. As such, for monitoring purposes, it could be proposed that 5% prevalence of intersex is considered to be 'above baseline', with all sites with a prevalence above this being further assessed for intersex severity using the OSI approach of Bateman *et al*. (2004). This gives a two-tiered assessment of intersex utilizing apparent prevalence in the population, and an indicator for severity in affected individuals.

Review of the environmental variables that influence the presence of intersex in marine and estuarine fish

3) While the link between the formation of intersex (ovotestis) and exposure to anthropogenic contaminants considered to be 'endocrine disrupters' has been demonstrated for several fish species (e.g. Gimeno *et al*., 1996, 1997), it is also known that intersex and sex reversal are not specific markers for estrogens but rather they have many causes (including androgens, aromatase inhibitors and even water temperature shifts). Recent work has also demonstrated a potential for age to affect the occurrence and prevalence of the condition in freshwater fish species (Jobling *et al*., 2009). For certain species utilized in monitoring programmes in the OSPAR region, there is a clear historical link between those sites where anthropogenic endocrine disrupters, direct biomarkers of endocrine disruption (e.g. VTG) and the presence of intersex in populations residing in those habitats are most pronounced (for example, see links between papers by Allen *et al*., 1999a,b and Stentiford *et al*., 2003 for estuarine flounder). Extending this relationship between cause and effect to offshore populations is not so clear although data presented by Scott *et al*. (2007) showing elevated VTG in dab sampled from certain North Sea sites do correspond to data presented by Stentiford and Feist (2005) for intersex in the same species from these sites. Complications in specifically linking the presence of a chronic marker (such as intersex) with more acute phase markers (such as VTG), or the burden of anthropogenic chemicals are not unique in this instance, with similar parallels being reported in liver cancers present in a consistent, but as yet unexplainable manner in multiyear samples of dab collected from offshore sites (Stentiford *et al*. 2009, 2010). Interestingly, those estuarine and offshore sites with the highest prevalence of liver pathologies (including cancer) are also those where intersex have been reported. However, because hatchlings and juveniles are likely to inhabit different grounds to those where adults are sampled

(Dipper, 1987) and it is at these early life stages at which sex is determined (and at which disruption may occur) (Gimeno *et al*., 1997; Devlin and Nagahama, 2002), the presence of fish with the intersex condition at the particular offshore sites may not necessarily reflect the presence of EDCs at the site but rather their presence at sites where hatching and early growth occurs. Future studies should be directed towards the measurement of intersex in fish of known age, or in earlier life stages residing at monitoring sites and at those sites identified at nursery grounds for the key monitoring species. Comparisons of the prevalence of the intersex condition in juvenile and adult fish of the same species may furthermore provide clarification on the population level effects of EDCs in the marine environment and on their long-term ecological effects on sensitive ecosystems. Coupled with studies on the population genetics of these species and the identification of specific spawning grounds for different adult stocks, the potential selective pressure imposed by endocrine disturbances may also be identified.

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

4) As stated above, given the fact that intersex appears to exist at a range of between 0 and 27.8% in different monitoring species, a sampling regime based upon detection of 5% prevalence at 95% confidence appears appropriate. Furthermore, multi-site surveys in several species (e.g. flounder and dab by Stentiford *et al*., 2003 and Stentiford and Feist, 2005, respectively) have demonstrated that intersex is detected at some sites and not at others when this regimen is utilized. This indicates that intersex, if present, occurs at below 5% at these latter sites. As such, for monitoring purposes, it could be proposed that 5% prevalence of intersex is considered to be 'above baseline', with all sites with a prevalence above this being further assessed for intersex severity using the OSI approach of Bateman *et al*. (2004). This gives a two-tiered assessment of intersex utilizing apparent prevalence in the population, and an indicator for severity in affected individuals. It also allows for the discounting of potential isolated cases of intersex that may occur due to genetic abnormalities or other causes.

Proposals for assessment tools

5) Given background data on quality assurance techniques for intersex measurement, it seems appropriate to propose a two-tier assessment tool. Tier 1 consists of an individual sample grading system for intersex severity based on the methodology presented by Bateman *et al*. (2004). Tier 2 consists of apparent prevalence estimates based upon a sampling regime designed to detect a 5% prevalence of intersex at 95% confidence. Both of these tools can be combined to provide a population-level and individual-level assessment tool for the condition. Because intersex prevalence is likely to be negligible in non-impacted populations, survey designs are likely to be similar to that for fish disease measurement, whereby detection is based upon diseases present in a population at 5% prevalence (95% confidence). In this way, >5% prevalence would be considered the cut-off point for definition of an impacted population. It is recommended that cohort-matching is applied when comparing fish captured from different geographic sites, similar to the manner carried out for assessment of liver pathologies (Stentiford *et al*., 2010).

Literature cited

Allen, Y., Matthiessen, P., Scott, A.P., Haworth, S., Feist, S.W., Thain, J.E. 1999b. The extent of oestrogenic contamination in the UK estuarine and marine environments—further surveys of flounder. Sci. Total Environ. 233, 5–20.

Allen, Y., Scott, A.P., Matthiessen, P., Haworth, S., Thain, J.E., Feist, S.W. 1999a. Survey of estrogenic activity in United Kingdom estuarine and coastal waters and its effect on gonadal development of the flounder *Platichthys flesus*. Environ. Toxicol. Chem. 18, 1791–1800.

Amaoka, K., Nishikawa, S., Tanaka, N. 1974. Sexual dimorphism and an abnormal intersexual specimen in the botid flounder *Bothus pantherinus*. Jpn. J. Ichthyol. 21, 16–20.

Bateman, K.S., Stentiford, G.D., Feist, S.W. 2004. A ranking system for the evaluation of intersex condition in European flounder (*Platichthys flesus*). Environ. Toxicol. Chem. 23, 2831–2836.

Blachuta, J., Witkowski, A., Kokurewicz, B. 1991. An hermaphrodite grayling, *Thymallus thymallus* (L.) from the Nysa Klodzka River (Lower Silesia, Poland). J. Fish. Biol. 38, 955–957.

Colborn, T., Dumanoski, D., Peterson Myers, J. 1996. Our stolen future. Abacus, London.

De Metrio, G., Corriero, A., Desantis, S., Zubani, D. *et al*. 2003. Evidence of a high percentage of intersex in the Mediterranean swordfish (*Xiphias gladius* L.). Mar. Pollut. Bull. 46, 358–361.

Devlin, R.H., Nagahama, Y. 2002. Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences. Aquaculture 208, 191–364.

Dipper, F. 1987. British sea fishes. Underwater World Publications, London.

Feist, S.W., Lang, T., Stentiford, G.D., Koehler, A. 2004. Use of liver pathology of the European flatfish dab (*Limanda limanda* L.) and flounder (*Platichthys flesus* L.) for monitoring. ICES Tech. Mar. Environ. Sci. 38, 1–48.

Fossi, M.C., Casini, C., Ancora, S., Moscatelli, A., Ausilli, A., Notarbartolo di Sciara, G. 2001. Do endocrine disrupting chemicals threaten Mediterranean swordfish? Preliminary results of vitellogenin and Zona radiata proteins in *Xiphias gladius*. Mar. Environ. Res. 52, 477–483.

Fry, D.M. 1995. Reproductive effects in birds exposed to pesticides and industrial chemicals. Environ. Health Perspect. 103, 165–171.

Gercken, J., Sordyl, H. 2002. Intersex in feral marine and freshwater fish from northeastern Germany. Mar. Env. Res. 54, 651–655.

Gimeno, S., Gerritson, A., Bowmer, T., Komen, H. 1996. Feminization of male carp. Nature 384, 221–222.

Gimeno, S., Komen, H., Venderbosch, P.W.M., Bowmer, T. 1997. Disruption of sexual differentiation in genetic male common carp (*Cyprinus carpio*) exposed to an alkylphenol during different life stages. Environ. Sci. Technol. 31, 2884–2890.

Hashimoto, S., Bessho, H., Nakamura, M., Iguchi, T., Fujita, K. 2000. Elevated serum vitellogenin levels and gonadal abnormalities in wild male flounder (*Pleuronectes yokohamae*) from Tokyo Bay, Japan. Mar. Environ. Res. 49, 37–53.

Jafri, S.I.H., Ensor, D.M. 1979. Occurrence of an intersex condition in the roach *Rutilus rutilus* L. J. Fish Biol. 14, 547–549.

Jobling, S., Nolan, M., Tyler, C.R., Brightly, G., Sumpter, J.P. 1998. Widespread sexual disruption in wild fish. Environ. Sci. Tech. 32, 2498–2506.

Jobling, S., Burn, R.W., Thorpe, K., Williams, R., Tyler, C. 2009. Statistical Modelling Suggests That Anti-Androgens in Wastewater Treatment Works Effluents are Contributing Causes of Widespread Sexual Disruption in Fish Living in English Rivers. Environmental Health Perspectives. doi: 10.1289/ehp.0800197.

Kavanagh, R.J., Balch, G.C., Kiparissis, Y., Niimi, A.J., Metcalfe, C.D. 2002. Gonadal intersex in white perch (*Morone Americana*) from the Lower Great Lakes region. In: Conference on Great Lakes Research, Winnipeg, Canada, June 2002. International Association for Great Lakes Research, Ann Arbor, MI, p 64.

Matthiessen, P., Allen, Y., Bignell, J., Craft, J. *et al*. 2000. Studies of endocrine disruption in marine fish—progress with the EDMAR programme. ICES, Copenhagen.

Martin-Skilton, R., Lavado, R., Thibaut, R., Minier, C., Porte, C. 2006. Evidence of endocrine alteration in the red mullet, *Mullus barbatus* from the NW Mediterranean. Env. Poll. 141, 60–68.

Mikaelian, I., De Lafontaine, Y., Harshbarger, J.C., Lee, L.L.J., Martineau, D. 2002. Health of lake whitefish (*Coregonus clupeaformis*) with elevated tissue levels of environmental contaminants. Environ. Toxicol. Chem. 21, 532–541.

Minier, C., Caltot, G., Leboulanger, F., Hill, E.M. 2000. An investigation of the incidence of intersex fish in Seine-Maritime and Sussex regions. Analusis 28, 801–806.

Peters, L.D., Doyotte, A., Mitchelmore, C.L., McEvoy, J., Livingstone, D.R. 2001. Seasonal variation and estradiol-dependent elevation of Thames estuary eel *Anguilla anguilla* plasma vitellogenin and comparisons with other United Kingdom estuaries. Sci. Total Environ. 279, 137–150.

Purdom, C.E., Hardiman, P.A., Bye, V.J., Eno, N.C., Tyler, C., Sumpter, J.P. 1994. Estrogenic effects of effluent from sewage treatment works. Chem. Ecol. 8, 275–285.

Scott, A.P., Sanders, M., Stentiford, G.D. Reese, R.A., Katsiadaki, I. 2007. Evidence for oestrogenic endocrine disruption in an offshore flatfish, the dab (*Limanda limanda* L.). Mar. Env. Res. 64, 128–148.

Sharpe, R.M. 1997. Do males rely on female hormones? Nature 390, 447–448.

Simon, R.C., Schill, W.B. 1984. Tables of sample size requirements for detection of fish infected by pathogens: three confidence levels for different infection prevalence and various population sizes. J. Fish Dis. 7, 515–520.

Slooff, W., Kloowijk-Vandijk, E. 1982. Hermaphroditism in the bream, *Abramis brama* (L.). J. Fish Dis. 5, 79–81.

Stentiford, G.D., Feist, S.W. 2005. First case of intersex (ovotestis) in the flatfish species, dab (*Limanda limanda*): Dogger Bank, North Sea. Mar. Ecol. Prog. Ser. 301, 307–310.

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

Stentiford, G.D., Bignell, J.P., Lyons, B.P., Feist, S.W. 2009. Site-specific disease profiles in fish and their use in environmental monitoring. Mar. Ecol. Prog. Ser. 381, 1–15.

Stentiford, G.D., Bignell, J.P., Lyons, B.P., Thain, J.E., Feist, S.W. 2010. Age at onset of fish diseases: application to assessment of marine ecological health status. Mar. Ecol. Prog. Ser. 411, 215-230.

van Earle, R., Nolan, M., Jobling, S., Christiansen, L.B., Sumpter, J.P., Tyler, C.R. 2001. Sexual disruption in a second species of wild cyprinid fish (the gudgeon, *Gobio gobio*) in United Kingdom freshwaters. Environ. Toxicol. Chem. 20, 2841–2847.

Vigano, L., Arillo, A., Bottero, S., Massari, A., Mandich, A. 2001. First observation of intersex cyprinids in the Po River (Italy). Sci. Total Environ. 269, 189–194.

Annex 6: Technical Annex: Supporting parameters for biological effects measurements in fish and mussels

A. Measurement of supporting metrics for fish: condition indices, GSI, HSI and age

Background

1. For all biological effect techniques within the OSPAR JAMP and OSPAR integrated strategy there is a requirement to report supporting parameters, and these include species, sex, fish length, whole fish weight, liver weight and gonad size. The measurement of gonad size and liver weight is used to provide an indication of reproductive state, and liver weight may also give an indication of general health and well-being. These measurements are used in indices relating gonad weight to whole body weight (Gonad Somatic Index - GSI) and liver weight to whole body weight (Liver Somatic Index - LSI or Hepato Somatic Index - HSI), explanations of these are described below. Both gonad and liver weight will change markedly throughout the year and for comparative purposes these seasonal variations must be taken into account for the interpretation of biomarker responses such as EROD and VTG for example. Additionally, the condition factor (CF) is a general indicator for fish condition, similarly the condition index (CI) for mussels.

2. ICES WGBEC recently reviewed the measurement of these metrics and their role and importance in fish monitoring programmes, and this is described below.

Summary of supporting parameters required for fish;

Parameter	Measurement	Comment
Live fish whole body weight	To 0.1g	Blotted dry
Length of fish	To nearest mm	
Liver weight	To 0.1 g	
Gonad weight	To 0.01g	In addition record sex
Gonad length	To nearest mm	In addition record sex
Age	Conducted on otoliths	All individuals sampled

General Overview: Organ size and related measurements

3. Organ sizes constitute a very elementary measurement. The measurements can be performed with a minimum of equipment, and the procedures are easy to undertake. At least for some species it is possible to analyse these variables on frozen material. With minimal instruction these measurements can be determined by personnel not regularly involved in biomarker analysis, although it is preferable to use personnel familiar with handling fish and able to perform simple dissection of fish.

4. Data of this type may be of relevance either in their own right, indicating adverse effects of various kinds where the toxic mechanizms are not fully understood as a result of xenobiotic exposure and/or, partly as a supporting variable to biomarkers conducted at the whole individual, tissue, cellular and subcellular levels. As for all biomarkers in use today, there is a strong need for quality assurance when these measurements are carried out.

5. One of the most important measurements in this field may be the development of gonads among female fish. This variable is best expressed as gonad size relative to

the somatic body weight (Gonad Somatic Index - GSI) and expressed as a percentage value. The best species to use are those where the gonads of juvenile and immature fish are different from adult fish and where there are distinct differences in the genders. For example, it is much easier when the morphology of the female ovary is a single structure while the male testes are paired bilaterally.

6. This offers the opportunity to investigate when the fish in relation to size and/or age are sexually immature or adult, or indeed have retarded gonad development (often termed sexually immature - SIM) as compared to normal sexual development. This can be expressed as a percentage of sexually immature females among the adult females, and represents the portion of fish with the extreme low value of the GSI value (usually below ~1%) and they have therefore a gonad with no or neglected development.

7. Analogous to the analysis of the gonad size is the liver size relative to the somatic body weight (Liver Somatic Index – LSI, or sometimes referred to as Hepato Somatic Index - HSI). It may be regarded as a parameter in its own right and also as a supporting variable for other biomarkers such as EROD.

8. Furthermore, growth (e.g. gramme/year) as shown in Kiceniuk and Khan, 1986; McMaster *et al*., 1991 and in Ericson *et al*., 1998, as well as the Condition Factor (CF - see reference to Foulton below) are relatively straightforward to determine and may be used as markers for adverse effects due to xenobiotic exposure. The measurement of the condition factor has not often been used in short exposure laboratory experiments, however, field observations over longer time periods indicate that it may be a valuable measure for adverse effects. (See review by van der Ost *et al*., 2003). Recent investigations related to the Fish Disease Index (WGPDMO, 2011) support this assumption.

During periods of high food intake and also in conjunction with the reproductive cycle an individual may have a higher gross weight at a particular length. This can be assessed by calculating the coefficient of condition (K) or by Fulton's condition factor (Bagenal and Tesch, 1978). This is calculated as follows:

$$K = \text{weight} / (\text{length})^3$$

The condition factor reflects the nutritional state or "well-being" of an individual fish and is sometimes interpreted as an index of growth rate.

9. Feeding status in fish may be reflected in the condition factor, and may be important for a number of different responses, and as such can be included in biomonitoring investigations.

Gonad size in fish – GSI

10. The reproductive process constitutes (one of) the most essential health signals for the individual animal, and when missing or impaired indicates an obvious risk for adverse effect both genetically and for population survival. Therefore, decreased sizes of the gonad, of one or both of the genders, indicate an apparent risk for a reduced reproductive potential.

11. Gonad size is measured as a percentage of somatic body weight, gonadosomatic index (GSI*), It has been demonstrated to be a variable that can be influenced by contaminants in a number of different polluted field studies. It should be underlined that the toxicological response observed for this variable could have originated from a

number of different toxicological reasons such as, tissue or cell death to more sophisticated regulatory endocrine mechanizms.

Measurement of GSI: record whole body weight of fish and gonad weight to 2 decimal places.

*GSI = (gonad weight x 100) /(total body weight[#]–gonad weight)
[#]subtract stomach content

12. Deviation in GSI levels could represent a permanent effect or impairment for the reproductive cycle for one or more years (Janssen *et al*., 1997; Vallin *et al*., 1999). Both scenarios will seriously affect reproductive potential. Examples of different pollution gradients were reduced gonads have been observed are in bleached kraft pulp mill effluents (Andersson I., 1988; Sandström *et al*., 1988; McMaster *et al*., 1991; Balk *et al*., 1993; Förlin *et al*., 1995), including using chlorine-free processes (Karels *et al*., 2001) and general pollution (Johnson *et al*., 1988; Noaksson *et al*., 2001). Laboratory exposure experiment where effect on the GSI value have been documented include petroleum mixtures (Truscott *et al*., 1983; Kiceniuk and Khan, 1986), specific PAHs (Thomas, 1988; Singh, 1989; Thomas and Budiantara, 1995), PCB mixture (Thomas, 1988), pesticides (Ram *et al*., 1986; Singh, 1989), and cadmium (Singh, 1989; Pereira *et al*., 1993).

13. There is no doubt that xenobiotics can affect gonad size through a number of different toxicological mechanizms. However, as for most biomarkers, a variable that shows a (annual) natural biological cycle it is essential that the normal background values are well known, and that the appropriate control material is used for comparison. For the GSI value it should be pointed out that during certain time times of the year the gonad development is very fast and that different GSI values are obtained only within a period of a few days/weeks. Analysis of the GSI in these time periods should be avoided. Baseline studies are important in order to evaluate suitable time periods for this variable (Förlin and Haux, 1990; Larsen *et al*., 1992).

14. A state of complete disruption of sexual maturation reflects an extreme situation of low GSI values, e.g. a state of condition when the adult (based on age and/or size) fish are unable to develop from the prepubertal condition to the sexually mature stage. Field observations demonstrating a delay or lack of gonad development has been observed include the following species; burbot (*Lota lota*) in the north coast of the Bothnian bay (Pulliainen *et al*., 1992), English sole (*Parophrys vetulus*) in generally polluted areas in Puget sound, USA (Johnson *et al*., 1988), perch (*Perca fluviatilis*) in the effluent water from pulp and paper mills in Baltic waters (Sandström *et al*., 1988; Sandström *et al*., 1994) as well as white sucker (*Catostomus commersoni*) in corresponding effluents in Ontario, Canada (McMaster *et al*., 1991). Studies have also shown that perch, roach (*Rutilus rutilus*), and brook trout (*Salvelinus fontinalis*) exposed to leachate from a public refuse dump in a Swedish freshwater system show corresponding adverse effects (Noaksson *et al*., 2001; Noaksson *et al*., 2002). Although the above cited field investigations are not all related to suspected PAH contamination, these kinds of disorders has been created in laboratory experiment using petroleum products and a pure naphthalene (Thomas and Budiantara, 1995).

GSI confounding factors

15. Although the measurement is robust and easy to perform there is a need to characterize and avoid confounding factors. For example female perch populations do not naturally spawn every year and the spawning frequency is affected by water temperature as indicated in Luksiene *et al*., 2000 and Sandström *et al*., 1995. Moreover, in

the closely related yellow perch (*Perca flavescens*) both photoperiod and temperature have been suggested to be of importance (Dabrowski *et al*., 1996). Therefore, GSI data should be interpreted with regard to the reproductive cycle for each species under investigation.

Liver size of female and/or male fish – LSI (HSI)

16. Liver size is measured in relation to somatic body weight, and is known as Liver Somatic Index (LSI* or HSI – see above).

<u>Measurement of LSI:</u> record whole body weight of fish and liver weight to two decimal places.

*LSI = (liver weight x 100) /(total body weight#–liver weight)
#subtract stomach content

17. LSI may be regarded as a relevant measurement because it has been documented to be affected by contaminants in a number of different polluted field studies. For example, in pollution gradients of paper and pulp mill effluents where increased LSI vales were observed (Andersson *et al*., 1988; Lehtinen *et al*., 1990; Hodson *et al*., 1992; Kloepper-Sams and Owens, 1993; Huuskonen and Lindström-Seppa, 1995; Förlin *et al*., 1995), as well as decreased LSI levels as reported by Balk *et al*. (1993), and Förlin *et al*. (1995). Other complex effluents shown to affect liver size in various fish species are: leakage water from public refuse dumps (Noaksson *et al*., 2001; 2002) and effluent from wastewater treatment plant (Kosmala *et al*., 1998).

17. Field situations where PAHs and/or organochlorines are suspected contaminants for increased liver size in various fish species are documented by: Sloff *et al*. (1983); Goksoyr *et al*. (1991); Kirby *et al*. (1999); Kirby *et al*. (1999); Beyer *et al*. (1996); Leadly *et al*. (1998); Stephensen *et al*. 2000). Laboratory experiments shown to affect liver size among different fish species from exposure to organochlorines have been documented by: Adams *et al*. (1990); Newsted and Giesy (1993); Otto and Moon (1995); Arnold *et al*. (1995); Gadagbui and Goksoyr (1996); Åkerblom *et al*. (2000), and for two-stroke outboard engine exhaust extract (Tjärnlund *et al*., 1996) and PAHs (Celander *et al*., 1994) as well as pesticides (Singh, 1989; Åkerman *et al*., 2003) and cadmium (Singh, 1989).

LSI Confounding factors

18. Although there is no doubt that xenobiotics could affect liver size as a result of different toxicological mechanizms it should be emphasized that, as for most biomarkers, control/reference fish should be analysed in close/direct parallel with the exposed site(s). In addition, seasonal variation is observed in different fish species (Koivusaari *et al*., 1981; Förlin and Haux, 1990; Larsen, 1992), and must be taken into account at all times. Besides the time of the year, factors (i.e. parameters) such as feeding behaviour, gender, maturity, age, size, temperature (George *et al*., 1990), photoperiod, parasites, among others, needs to be taken into considerations. Baseline studies are an important strategy to finally evaluate confounding factors (Balk *et al*., 1996).

19. Determination of age

It is essential to the interpretation and assessment of biological effect responses that the age of fish is known. This is particularly important for effect measurements such as fish diseases which may be more prevalent in older fish (Stentiford *et al*., 2010). Age is assessed by removing the otoliths of each fish sampled, and using standard procedures. These vary with species, and sometimes location, and specific guidance should be sought from relevant experts, or ICES. In some species, age may be more easily determined in scales or bone. Ideally age-size relationship (length and weight) should be known for several populations of fish species for longer time periods, because the growth of a fish species may vary in different populations and at different locations, and from year to year .

20. Interpretation of data

The GSI, LSI (HSI) and condition factor are described here as supporting parameters to assist the interpretation of contaminant related biological effect measurements. However, it should be noted that these supporting parameters in their own right may be influenced by a number of factors which should be described if known and these include: feeding behaviour, gender, maturity, development stage, age, water temperature, presence of parasitic infections and other disease, location and seasonality.

B: Measurement of supporting metrics for mussel: condition indices

Background

1. In Northern Europe mussels have their main spawning season in late winter to early spring e.g. February in the UK. During the onset of reproduction energy normally used in shell and somatic growth is fully utilized for gametogenesis. This is manifested by a marked increase in flesh weight relative to whole body weight which increases and reaches a maximum at spawning. Post-spawning, flesh weight relative to whole body weight is at a minimum. As a consequence flesh weight relative to whole body weight or internal shell volume may be regarded as an index of condition.

2. For all biological effect techniques within the OSPAR mussel integrated strategy there is a requirement to report supporting parameters, and these include mussel length, whole body weight and condition index.

Summary of supporting parameters required for mussels: at least ten animals per site, usually within a specific size e.g. 40–45 mm or similar depending on availability at the site.

Parameter	Measurement	Comment
Live whole animal weight	To 0.1g	Must be on animals taken from full immersion i.e including water in body cavity (not gaping). Also blotted dry
Length of animal / width	To nearest mm	
Wet flesh weight	To nearest 0.1g	Flesh excised from open shell and drained / blotted dry
Dry flesh weight	To nearest 0.01g	80 degrees C for 24 hr and constant dry weight
Wet shell weight	To nearest 0.01g	Blotted dry
Dry shell weight	To nearest 0.01g	80 degrees C for 24 hr and constant dry weight
Internal shell volume	To 0.1ml	Not generally conducted but provides a very accurate measure of condition.

Condition index

3. Condition Indices(CI) based on flesh weight relative to whole weight or shell have been used for several years, both in scientific research and in commercial fisheries and several methods are available (see Lutz, 1980, Aldrich and Crowley, 1986, Davenport and Chen, 1987). The methods may use wet flesh weight, whole weight and shell size and/or volume but these are less sensitive due to the difficulty in standardizing the degree of wetness. Indices using dry flesh weight are more accurate particularly when used in relation to internal shell volume.. Example of condition indices are given below;

CI "A" = 100 x Dry weight / Whole animal weight

CI "B" = 100 x Dry weight / Wet flesh weight

CI "C" = 100 x Dry weight / Internal shell volume

CI "D" = (Ratio of shell length:shell width) / dry weight

In general CI "A" is commonly used for convenience and ease of measurement but the most accurate assessment of condition is CI "C". Whatever condition index is used, it is high post-spawning and lower post-spawning when the animal is in poor condition and the flesh weight is greatly reduced relative to the whole animal weight and the volume of the internal shell cavity (Dix and Ferguson, 1984; Rodhouse *et al*., 1984).

4. It should be noted that condition indices will vary according to body size (Lutz *et al*., 1980). In addition, other factors such as the level of parasitic infection (Kent, 1979 and Thiessen, 1987) and aerial exposure can adversely affect the condition of mussels.

References

Adams, S.M., Shugart, L.R., Southworth, G.R., and Hinton, D.E. 1990. Application of bioindicators in assessing the health of fish populations experiencing contaminant stress. In: McCarthy, J.F., Shugart, R.L. (Eds.). Biomarkers of Environmental Contamination. Lewis Pub., pp 333–353.

Åkerblom, N., Olsson, K., Berg, A.H., Andersson, P.L., Tysklind, M., Förlin, L., and Norrgren, L. 2000. Archives of Environmental Contamination and Toxicology 38:225–233.

Åkerman, G., Amcoff, P., Tjärnlund., Fogelberg, K., Torrissen, O. and Balk, L. 2003. Paraquat and menadione exposure of rainbow trout (*Oncorhynchus mykiss*) –Studies of effects on the pentose-phosphate shunt and thiamine levels in liver and kidney. Chemico-Biological Interactions 142, 269–283.

Aldich, J.C. and Crowly, M. 1986. Conditions and variability in *Mytilus edulis* L. from different habitats in Ireland. Aquaculture 52: 273–286.

Andersson, T., Förlin, L., Härdig, J., and Larsson, Å. 1988. Physiological disturbances in fish living in coastal water polluted with bleached kraft pulp mill effluents. Canadian Journal of Fisheries and Aquatic Science 45: 1525–1536.

Arnold, H., Pluta, H.-J., and Braunbeck, T. 1995. Simultaneous exposure of fish to endosulfan and disulfoton *in vivo*: ultrastructural, stereological and biochemical reactions in hepatocytes of male rainbow trout (*Oncorhynchus mykiss*). Aquatic Toxicology 33: 17–43.

Bagenal, T.B. and Tesch, F.W. 1978. Age and Growth. Pages 101–136 in T.B. Bagenal, edit. Methods for assessment of fish production in freshwaters, 3rd edition. Blackwell Scientific Publications, Oxford, England.

Balk, L., Andersson, T., Förlin, L., Söderström, M., and Larsson, Å. 1993. Indications of regional and large-scale biological effects caused by bleached pulp mill effluents. Chemosphere 27: 631–650.

Balk, L., Larsson, Å. and Förlin, L. 1996. Baseline studies of biomarkers in the feral female perch (*Perca fluviatilis*) as tools in biological monitoring of anthropogenic substances. Marine Environmental Research, 42 (1–4), 203–208.

Beyer, J. 1996. Fish biomarkers in marine pollution monitoring; evaluation and validation in laboratory and field studies. Academic thesis, University of Bergen, Norway.

Celander, M., Näf, C., Broman, D., and Förlin, L. 1994. Temporal aspects of induction of hepatic cytochrome P450 1A and conjugating enzymes in the viviparous blenny (*Zoarces viviparus*) treated with petroleum hydrocarbons. Aquatic Toxicology 29: 183–196.

Dabrowski, K., Ciereszko, R.E., Ciereszko, A., Toth, G.P., Christ, S.A., El-Saidy, D., and Ottobre, J.S. 1996. Reproductive physiology of yellow perch (*Perca flavescens*): Environmental and endocrinological cues. Journal of Applied Ichtyology :12: 139–148.

Davenport, J. and Chen, X. 1987. A comparison of methods for the assessment of condition in the mussel (*Mytilus edulis* L.). J. Molluscan Stud., 53: 293–297.

Dix, T.G. and Fergussen, A. 1984. Cycles of reproduction and condition in Tasmanian blue mussel *Mytilus edulis planatus*. Aust. J. Mar Freshw. Res. 35; 307–313.Ericson, G., Lindesjöö, E. and Balk, L. (1998) "DNA adducts and histopathological lesions in perch (*Perca fluviatilis*) and northern pike (*Esox lucius*) along a polycyclic aromatic hydrocarbon gradient on the Swedish coastline of the Baltic Sea." Canadian Journal of Fisheries and Aquatic Sciences 55, 815–824.

Förlin, L., Andersson, T., Balk, L., and Larsson, Å. 1995. Biochemical and physiological effects in fish exposed to bleached kraft mill effluents. Ecotoxicology and Environmental Safety 30: 164–170.

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

George S, Young P, Leaver M, and Clarke D. Activities of pollutant metabolizing and detoxification systems in the liver of the plaice, *Pleuronectes platessa*: sex and seasonal variations in non-induced fish. Comp Biochem Physiol 1990; 96C: 185-192.

Gadagbui, B.K.-M., and Goksoyr, A. 1996. CYP1A and other biomarker responses to effluents from a textile mill in the Volta river (Ghana) using caged tilapia (*Oreochromis niloticus*) and sediment-exposed mudfish (*Clarias anguillaris*). Biomarkers 1: 252–261.

Goksoyr and 10 co-authors. 1991. Environmental contaminants and biochemical responses in flatfish from the Hvaler archipelago in Norway. Archives of Environmental Contamination and Toxicology 21: 486–496.

Hodson and 8 co-authors. 1992. Effects of bleached kraft mill effluent on fish in the St Maurice River, Quebec. Environmental Toxicology and Chemistry 11: 1635–1651.

Huuskonen, S., and Lindström-Seppä, P. 1995. Hepatic cytochrome P4501A and other biotransformation activities in perch (*Perca fluvitialis*): the effects of unbleached pulp mill effluents. Aquatic Toxicology 31: 27–41.

Janssen, P.A.H., Lambert, J.G.D., Vethaak, A.D., and Goos, H.J.Th. 1997. Environmental pollution caused elevated concentration of oestradiol and vitellogenin in the female flounder, *Platichthys flesus* (L.). Aquatic Toxicology 39: 195–214.

Johnson, L.L., Casillas, E., Collier, T.K., McCain, B.B., and Varanasi, U. 1988. Contaminant effects on ovarian development in English sole (*Parophrys vetulus*) from Puget Sound, Washington. Canadian Journal of Fisheries and Aquatic Science 45: 2133–2146.

Karels, A., Markkula, E., and Oikari, A. 2001. Reproductive, biochemical, physiological, and population responses in perch (*Perca fluviatilis* L.) and roach (*Rutilus rutilus* L.) downstream of two elemental chlorine-free pulp and paper mills. Environmental Toxicology and Chemistry 20: 1517–1527.

Kent, R.M.L. 1979. The influence of heavy infestations of *Polydora ciliate* on the flesh content of *Mytilus edulis*. J. Mar. Biol. Ass. UK. 59: 289–297.Kiceniuk, J.W., and Khan, R.A. 1987. Effect of petroleum hydrocarbons on Atlantic cod, *Gadus morhua*, following chronic exposure. Canadian Journal of Zoology 65:490–494.

Kirby, M.F., Matthiessen, P., Neall, P., Tylor, T., Allchin, C.R., Kelly, C.A., Maxwell, D.L., and Thain, J.E. 1999. Hepatic EROD activity in flounder (*Platichys flesus*) as an indicator of contaminant exposure in English estuaries. Marine Pollution Bulletin 38:676–686.

Kirby, M.F., Neall, P., and Tylor, P. 1999. EROD activity measured in flatfish from the area of the Sea Empress oil spill. Chemosphere 38: 2929–2949.

Kloepper-Sams P.J. and Owens J.W. 1993. Environmental biomarkers as indicators of chemical exposure. Journal of Hazardous Material. 35: 283–294.

Kosmala, A., Migeon, B., Flammarion, P., and Garric, J. 1998. Impact assessment of a wastewater treatment plant effluent using the fish biomarker ethoxyresorufin-O-deethylase: field and on-site experiments. Ecotoxicology and Environmental Safe 41: 19–28.

Koivusaari U, Harri M, Hanninen O. Seasonal variation of hepatic biotransformation in female and male rainbow trout, *Salmo gairdneri*. Comp Biochem Physiol, 1981; 70C: 149–157.

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

Leadly, T.A., Balch, G., Metcalfe, C.D., Lazar, R., Mazak, E., Habowsky, J., Haffner, D. 1998. Chemical accumulation and toxicological stress in three brown bullhead (*Ameiurus nebulosus*) populations of the Detroit river, Michigan, USA. Environmental Toxicology Chemistry 17: 1756–1766.

Lehtinen, K.-J., Kierkegaard, A., Jakobsson, E., Wändell, A. 1990. Physiological effects in fish exposed to effluents from mills with six different bleaching processes. Ecotoxicology and Environmental Safe 19: 33–46.

Luksiene, D., Sandström, O., Lounasheimo, L. and Andersson, J. 2000. The effects of thermal effluent exposure on the gametogenesis of female fish. Journal of Fish Biology 56: 37–50.

Lutz, R.A. 1980. Mussel Culture and Harvest: A North American Perspective, Elsevier Science Publishers, B.V. Amsterdam, 305pp.

Lutz, R. A., Incze, L. S., Porter, B., and J. K. Stotz. 1980. Seasonal variation in the condition of raft-cultivated mussels (*Mytilus edulis* L.), Proc World Mariculture Society, 11, 262–268.

Newsted, J.L., and Giesy, J.P. 1993. Effect of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) on the epidermal growth factor receptor in hepatic plasma membranes of rainbow trout (*Oncorhynchus mykiss*) Toxicology and Applied Pharmacology 119: 41–51.

Noaksson E, Tjärnlund U, Bosveld ATC, Balk L. Evidence for endocrine disruption in perch *Perca fluviatilis* and roach *Rutilus rutilus* in a remote Swedish lake in the vicinity of a public refuse dump. Toxicol Appl Pharmacol 2001; 174: 160–176.

Noaksson, E., Linderoth, M., Bosveld, A.T.C., Norrgren, L., Zebuhr, Y. and Balk, L. 2002. "Endocrine disruption in brook trout (*Salvelinus fontinalis*) exposed to leachate from a public refuse dump." The Science of the Total Environment, In Press.

Noaksson, E., Tjärnlund, U., Bosveld, A.T.C., and Balk, L. 2001. Evidence for endocrine disruption in perch *Perca fluviatilis* and roach *Rutilus rutilus* in a remote Swedish lake in the vicinity of a public refuse dump. Toxicology and Applied Pharmacology 174: 160–176.

Noaksson, E., Linderoth, M., Bosveld, A.T.C., Norrgren, L., Zebuhr, Y., and Balk, L. 2002. Endocrine disruption in brook trout (*Salvelinus fontinalis*) exposed to leachate from a public refuse dump." The Science of the Total Environment, In Press.

Otto, D.M.E., and Moon, T.W. 1995. 3,3′,4,4′-tetrachlorbiphenyl effects on antioxidant enzymes and glutathione status in different tissues in rainbow trout. Pharmacology and Toxicology 77: 281–287.

Pereira, J.J., Mercaldo-Allen, R., Kuropat, C., Luedke, D., and Sennefelder, G. 1993. Effect of cadmium accumulation on serum vitellogenin levels and hepatosomatic and gonadosomatic indices of winter flounder (*Pleuronectes americanus*). Archives of Environmental Contamination and Toxicology 24: 427–431.

Ram, R.N., and Sathyanesan, A.G. 1986. Effect of a mercurial fungicide on the gonadal development of the teleostean fish *Channa punctatus* (Bloch). Ecotoxicology and Environmental Safety 11: 352–360.

Rodhouse, P.G., Rhodes, C.M., Burrell, G.M., Hensey, M.P., McMahon, T., Ottway, B. and Ryan, T.H. 1984. Food resource, gametogenesis and growth of *Mytilus edulis* on the shore and in suspended culture; Killory Harbour, Ireland. J. Mar. Biol. Ass. UK: 59:289–297. Sandström O. Incomplete recovery in a coastal fish community exposed to effluent from a modernised Swedish bleached kraft mill. Can J Fish Aquat Sci 1994; 51: 2195–2205.

Sandström, O., Neuman, E., and Karås, P. 1988. Effects of bleached pulp mill effluent on growth and gonad function in Baltic coastal fish. Water Science Technology 20: 107–118.

Sandström O, Neuman E, Thoresson G. Effects of temperature on life-history variables in perch, *Perca fluviatilis*. J Fish Biol 1995; 47: 652–670.

Singh, H. 1989. Interactions of xenobiotics with reproductive endocrine functions in a Protogynous teleost, *Monopterus albus*. Marine Environmental Research 28: 285–289.

Sloff W., van Kreijl, C.F., Baars, A.J. 1983. Relative liver weights and xenobiotic-metabolizing enzymes of fish from polluted surface waters in the Netherlands. Aquatic Toxicology: 4: 1–14.

Stentiford, G.D., Bignell, J.P., Lyons, B.P., Thain, J.E. and Feist, S.W. 2010. Effect of age on liver pathology and other diseases in flatfish: implications for assessment of marine ecological health status. Marine Ecology Progress Series: 411: 215–230.

Stephensen, E., Svavarsson, J., Sturve, J., Ericson, G., Adolfsson-Erici, M., and Förlin. 2000. Biochemical indicators of pollution exposure in shorthorn sculpin (*Myoxocephalus scorpius*), caught in four harbours on the southwest coast of Iceland. Aquatic Toxicology: 48: 431–442.

Thiessen, B.FR. 1987. Infestation of *Mytilus edulis* by *Mytilicola intestinalis*. Ophelia 27: 77–86.

Thomas, P. 1988. Reproductive endocrine function in female Atlantic croaker exposed to pollutants. Marine Environmental Research 24: 179–183.

Thomas, P., and Budiantara, L. 1995. Reproductive life-history stages sensitive to oil and naphthalene in Atlantic Croaker. Marine Environmental Research 39: 147–150.

Truscott, B., Walsh, J.M., Burton, M.P., Payne, J.F., and Idler, D.R. 1983. Effect of acute exposure to crude petroleum on some reproductive hormones in salmon and flounder. Comparative Biochemistry and Physiology 75C: 121–130.

Tjärnlund, U., Ericson, G., Lindesjöö, E., Petterson, I., Åkerman, G. and Balk, L. 1996. Further studies of exhaust from two-stroke outboard engines on fish. Marine Environmental Research, 42: 267–271.

Vallin, L., Nissling, A., and Westin, L. 1999. Potential factors influencing reproductive success of Baltic cod, *Gadus morhua*: a review. Ambio 28: 92–99.

van der Oost, R., Beyer, J., and Vermeulen, N.P.E. 2003. Fish bioaccumulation and biomarkers in environmental risk assessment: a review. Environmental Toxicology and Pharmacology 13: 57–149.

WGPDMO. 2011. Report of the ICES Working Group on Pathology and Diseases of Marine Organisms. International Council for the Exploration of the Sea, Copenhagen.

Annex 7: Background Document (revised): Acetylcholinesterase assay as a method for assessing neurotoxic effects in aquatic organisms

Compiled by: Thierry Burgeot[1] , Joelle Forget-Leray[2] , Gilles Bocquené[1] and Lùcia Guilhermino[3], Concha Martinez-Gomez[4] and Kari Lehtonen[5]

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

[2]LEMA, Le Havre University, Rue P. Lebon, 76400 Le Havre, France

[3]CIIMAR-LA/CIIMAR Centro Interdisciplinar de Investigação Marinha e Ambiantal, Universidad do Porto, laboratorio de ecotoxicologia, 4050-123, Porto, Portugal

[4]Centro Oceanografico de Murcia, Instituto Espagñol de Oceanografia, 30740 San Pedro Del Pinatar, Spain

[5] Finnish Institute of Marine Research, FI 00561, Helsinki, Finland.

Background

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

AChE is present in most animals and is responsible for the rapid hydrolytic degradation of the neurotransmitter acetylcholine (ACh) into the inactive products choline and acetic acid. AChE has highest specificity for ACh of any other choline ester, while butyrylcholinesterase has the highest specificity for butyrylcholine or propylthiocholine. The inhibition of AChE leads to an accumulation of ACh which, in turn, overstimulates sensitive neurons at the neuromuscular junction which results in tonic spasm and tremors. The presence of AChE has been demonstrated in a variety of tissues of marine organisms including muscle and brain tissue of fish, adductor muscle, foot tissue, haemocytes and gills of shellfish, and abdominal muscle of crustaceans (Bocquené and Galgani, 1998). The highest activities have been found in the brain and muscle of fish, in the eye and muscle of prawn (Frasco *et al*., 2010). Molluscs in general show low activity (Bocquené *et al*., 1998). In vertebrates, neurotoxic poisoning with hyperactivity, tremors, convulsions and paralysis may finally lead to death.

Being an indicator of neurotoxic effects, AChE has traditionally been used as a specific biomarker of exposure to organophosphate and carbamate pesticides (e.g. Coppage and Braidech, 1976; Day and Scott, 1990; Bocquené and Galgani, 1998; Printes and Callaghan, 2004; Hoguet and Key, 2007). The existence of extremely low thresholds for induction of inhibitory effects on AChE suggests that detection is possible after exposure to low concentrations of neurotoxic insecticides (0.1 to 1 $\mu g l^{-1}$; Habig *et al*., 1986).

During the 1990s, there was a resurgence of interest concerning the use of ChEs as a biomarker. Its responsiveness has been demonstrated to various other groups of chemicals present in the marine environment including heavy metals, detergents and hydrocarbons (Zinkl *et al*., 1991; Payne *et al*., 1996; Guilhermino *et al*., 1998; Forget *et al*., 1999; Burgeot *et al*., 2001, Brown *et al*., 2004). Its usefulness as a general indicator of pollution stress in mussels from the Baltic Sea has recently been suggested and it has been used for this purpose (Schiedek *et al*., 2006; Kopecka *et al*., 2006, Barsiene *et al*., 2006).

Confounding factors

It is important to know the natural limits of variability in AChE activity in the species of interest to assess the significance of the observed depression in activity. Knowledge of possible variations related to sex, size, state of gonadal maturation and the influence of seawater temperature should be systematically determined. Also, the presence of different ChEs in the same tissue having different sensitivities to anti-cholinesterase agents may act as a confounding factor; therefore, prior characterization of the enzymes present is recommended (Garcia *et al*., 2000). AChE activity of juveniles of *Callionymus lyra* in the Atlantic sea and in *Serranus cabrilla* and *Mullus barbatus* in the Mediterranean Sea is higher than that of adults, but no differences were found between male and female in *Limanda limanda* in the Atlantic Ocean (Galgani and Bocquené, 1992).

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

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

The process and mechanizms of biological response in each organism require further investigation in specific habitats with specific chemical contamination. The mussel *Mytilus galloprovincialis* shows a great heterogeneity of esterases and a particular sensitivity to specific compounds such as paraoxon (Ozretic and Krajnovic-Ozretic, 1992; Brown, 2004). The alleged versatility of AChE inhibition as an effect criterion after exposure to detergents may be misleading and may underestimate the contamination potential of complex mixtures (Rodrigues *et al*., 2011). As for many other biomarkers, the hormesis effects cannot be ignored and represents a substantial scientific challenge. (Kefford *et al.*, 2008).

Enzymatic polymorphism has also been demonstrated in the oyster *Crassostrea gigas*, and two forms of AChE with different sensitivity to paraoxon have been described (Bocquené *et al.*, 1997). Thus, extraction of the sensitive form now identified in some organisms would provide greater precision for determination of AChE enzymatic activity than would an overall measurement of acetylcholinesterases. In addition to polymorphisms, ChEs of some invertebrates have been shown to have some differences in their properties compared to typical forms of vertebrates. For example, ChEs with properties of both AChE and pseudocholinesterases have been found in the gastropods *Monodonta lineata* and *Nucella lapillus* (Cunha *et al.*, 2007), in the sea urchin *Paracenthrotus lividus* (Cunha *et al.*, 2005), in *Artemia* sp (Varó *et al.*, 2002) and in some strains of *Daphnia magna* (Diamantino *et al.*, 2003).

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

Ecological relevance

AChE inhibition results in continuous and excessive stimulation of nerve and muscle fibres, producing tetany, paralysis and death. Sublethal exposure affecting AChE can alter the animal's behaviour and locomotive abilities (e.g. Vieira *et al.*, 2009), potentially affecting reproduction, fitness and survival. Therefore, AChE should be considered an ecologically relevant parameter, potentially affecting reproduction, fitness and survival. Evidence of modulation of AChE activity by organic chemicals, including fuel oil, has been described in marine organisms, including crustaceans (Signa *et al.*, 2008). The evaluation of the variations of AChE activity in different species allows characterization of neurotoxic effects of a wide spectrum of organic and inorganic contaminants in the marine environment.

Quality assurance

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

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

Standardization of the sampling strategy and regular intercalibration exercises on specific organisms sampled in the Atlantic Ocean, Mediterranean and the Baltic Sea are necessary for the widespread use of AChE in routine pollution monitoring.

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

Background Assessment Criteria (BAC) and Environmental Assessment Criteria (EAC)

Baseline levels of AchE activity in different marine species have been estimated from results derived from field studies in the Atlantic Ocean and the Mediterranean Sea (Table 1). Assessment criteria should be defined on regional basis, using available long-term data.Therefore, in order to understand and apply the AChE enzymatic activity as a biomarker of neurotoxic exposure, it is of fundamental importance to gain information on the natural background levels in non-contaminated organisms during at least two seasonal cycles. The baseline level (35 nmol.min^{-1} mg prot^{-1}) of the seasonal cycle of the mussel *Mytilus edulis* studied during three years along the Atlantic coast demonstrated a maximum of amplitude of 30% (Bocquené *et al.*, 2004).

Generally, it has been accepted that 20% reduction in AChE activity in fish and invertebrates indicates exposure to neurotoxic compounds (Zink *et al.*, 1987; Busby *et al.*, 1989). Depression of AChE activity by 20% to 50% indicates sublethal impact (Dizer *et al.*, 2001). In the field, several species have been found to have baseline AChE activities of the same order of magnitude in different studies/measurements (Table 1). However, differences between sea areas and seasons are obvious, e.g. with activity values in *Mytilus* spp. varying from 25 to 54 nmol min^{-1} mg protein^{-1}.

According to these observations, background assessment criteria (BAC) and environmental assessment criteria (EAC) were proposed using the 10th percentile of data. BACs are estimated from data from reference sites and describe the threshold value for the background level. Environmental Assessment Criteria (EACs) are usually derived from toxicological data and indicate a significant risk to the organism. EACs were calculated by subtracting 30% from the BAC values (Table 1) and represent a significant inhibition of AChE activity. EAC values characterize a sublethal impact. BACs and EACs should be estimated for different geographical regions, and include the effect of differences in water temperature.

Table 1. Assessment of acetylcholinesterase activity after *in vitro* and *in vivo* exposure of biomonitoring organisms in control laboratory conditions and field studies that have utilized common monitoring species collected from reference locations.

Organisms	Tissue	Reference location or control conditions	Sampling Season or month	Bottom Temperature or temperature range °C	BAC AChE 10th Percentile (activity nmol.min−1 mg prot−1)	EAC (activity nmol.min−1 mg prot−1)	Ref.
Invertebrates							
Mytilus gallorpovincialis	Gills	Wild mussels Mediterranean Sea in Spain	May–June	15–25	15	10	Campillo-Gonzalez (unpublished results)
Mytilus gallorpovincialis	Gills	Caging in field Mediterranean Sea –Carteau , France	Seasonal cycle	14–25	29	20	Bodin *et al.*, 2004
Mytilus edulis	Gills	Wild mussels Atlantic ocean (N.W. Portuguese)	Seasonal cycle		26	19	L.Guilhermino (unpublished results)
Mytilus edulis	Gills	Wild mussels Atlantic ocean (Loire estuary)	Seasonal cycle		30	21	Bocquené *et al.*, 2004
Vertebrates							
Plathichthys flesus	Muscle	French Atlantic ocean (Seine Bay)		15°C	235	165	Burgeot *et al.*, 2009
Plathichthys flesus	Muscle	French Atlantic ocean (Ster estuary-Brittany)		15°C	335	235	Evrard *et al.*, 2010
Limanda limanda	Muscle	French Atlantic ocean (Seine Bay)		15°C	150	105	Burgeot *et al.*, 2009
Mullus barbatus	Brain	Mediterranean Sea SE Spain (Màlaga-Almeria)	October	14 ºC	75	52	Martínez-Gómez, unpublished results
Mullus barbatus	Muscle	Mediterranean Sea (France, Spain, Italy)	*In situ*	18°C	155	109	Burgeot *et al.*, 1996, Bocquené, 2004

Future work

Standardized AChE measurement protocols and intercalibrations are required for the main species currently used in international marine biomonitoring programmes (OSPAR, HELCOM, MEDPOL and MSFD). An ICES TIMES series method document has been published (Bocquené and Galgani, 1998) and can be used as a basis of standardized procedure. Further information should be gathered to confirm baseline activity levels in specific habitats and different sentinel species in Europe. The BAC and EAC values must be considered as provisional and should be updated and revised when additional relevant data become available. BAC and EAC could also be derived for new species of interest and specific local studies.

References

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

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

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

Bocquené G., P. 2004. Biological effects of environmental pollution in aquatic ecosystem. BEEP WEB. NOE Global change and ecosystem 6.3.(Personal communication).

Bocquené G., Chantereau S., Raffin B., Minier C., Clérendeau C., Leskowicz A., Beausir E., Burgeot T. and Ménard D. 2004. The monitoring of the biological effects of the ERIKA oil spill on the blue mussel (*Mytilus edulis*). Aquatic Living Resources. 17.309–316.

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

Brown R.J., Galloway T.S., Lowe D., Browne M.A., Dissanayake A., Jones M.B. and Depledge M.H. 2004. Differential sensitivity of three marine invertebrates to copper assessed using multiple biomarkers. Aquat. Toxicol, 66. 267–278.

Burgeot T., G. Bocquené, C. Porte, A. Pfhol Leszkowicz, R.M. Santella, C. Raoux, J. Dimeet, and F. Galgani. 1996. Bio-indicators of pollutant exposure in the northwestern part of the Mediterranean Sea. Mar. Ecol. Prog. Ser.131: 125–141.

Burgeot T., G. Bocquené, E. His, F. Vincent, O. Geffard, R. Beiras, H. Goraguer and F. Galgani 2001. Procedures for cholinesterase determination in fish and mussel. Technical annex. In: Biomarkers in marine organisms; a practical approach. Edited by P. Garrigues H. Barth, C.H. Walker and J.F. Narbonne. Eds Elsevier.487p.

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

Busby D.G., White L.M., Pearce P.A. and Mineau P. 1989. Fenitrothion effects on forest songbirds: a critical new look. In: Environmental effects of fenitrothion use in forestry, Ed.: Ernst, W.R. et al Environment Canada Publication, no EN40-370/1989E, p.43–108.

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

Cunha I., Garcia L.M., Guilhermino L. 2005. Sea-urchin (*Paracentrotus lividus*) glutathione S-transferases and cholinesterase activities as biomarkers of environmental contamination. Journal of Environmental Monitoring 7: 288–294.

Cunha I., Mangas-Ramirez E., Guilhermino L. 2007. Effects of copper and cadmium on cholinesterase and glutathione S-transferase activities of two marine gastropods (*Monodonta lineata* and *Nucella lapillus*). Comparative Biochemistry and Physiology Part C: Toxicology and Pharmacology 145: 648–657.

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

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

Diamantino T.C., Almeida E., Soares A.M.V.M., Guilhermino, L. 2003. Characterization of cholinesterases from *Daphnia magna* Straus and their inhibition by zinc. Bulletin of Environmental Contamination and Toxicology 71: 219–225.

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

Evrard E., Devaux A., Bony S., Burgeot T., Riso R., Budzinski H., Ledu M., Quiniou L., Laroche J. 2010. Response of the European flounder *Platychthys flesus* to the chemical stress in estuaries: load of contaminants, gene expression, cellular impact and growth rate. Biomarkers. 1–17.

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

Frasco M.F., Fournier D., Carvalho F., Guilhermino, L. 2006. Cholinesterase from the common prawn (*Palaemon serratus*) eyes: catalitic properties and sensitivity to organophosphate and carbamate compounds. Aquatic Toxicology 77: 412–421.

Frasco M.F., Erzen I., Stojan J., Guilhermino L. 2010. Localization and properties of cholinesterases in the common prawn (*Palaemon serratus*): a kinetic-histochemical study. Biological Bulletin 218: 1–5.

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

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

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

Garcia L.M., Castro B., Ribeiro R., Guilhermino L. 2000. Characterization of cholinesterase from guppy (*Poecilia reticulata*) muscle and its *in vitro* inhibition by environmental contaminants. Biomarkers 5: 274–284.

Guilhermino, L.; Barros, P.; Silva, M.C.; Soares, A.M.V.M. 1998. Should the use of inhibition of cholinesterases as a specific biomarker for organophosphate and carbamate pesticides be questioned? Biomarkers 3:157–163.

Habig C., Di Guilio R.T., Nomeir A.A. and Abou Donia A.M.B. 1986. Comparative toxicity, cholinergic effects and tissue levels of S,S,S-tri-nButyl Phosphoritrithioate ‘DEF) to Channel catfish (*Ictarus punctatus*) and blue crab (*Callinectes sapidus*). Aquatic Toxicology, 9:193–206.

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

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

Kefford B.J., Zalizniak L, Warne M.J., Nugegoda D. 2008. Is the integration of hormesis and essentiality into ecotoxicology now opening Pandora’s Box? Environmental pollution 151:516–523.

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

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

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

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

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

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

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

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

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

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

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

Rodrigues S.R., Caldeira C., Castro B.B., Gonçalvez F., Nunes B. and Antunes S.C. 2011. Cholinesterase (ChE) inhibition un pumpkinseed (*Lepomis gibbosus*) as environmental biomarker: ChE characterization and potential neurotoxic effects of xenobiotics. Pesticide Biochemistry and Physiology. 99 : 181–188.

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

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

Tim-Tim A.L., Morgado F., Moreira S., Rangel R., Nogueira A.J., Soares A.M., Guilhermino L. 2009. Cholinesterase and glutathione S-transferase activities of three mollusc species from the NW Portuguese coast in relation to the 'Prestige' oil spill. Chemosphere 77: 1465–1475.

Varó I., Navarro J.C., Amat F., Guilhermino, L. 2002. Characterization of cholinesterases and evaluation of the inhibitory potential of chlorpyrifos and dichlorvos to *Artemia salina* and *Artemia parthenogenetica*. Chemosphere 48: 563–569.

Vieira L.R., Gravato C., Soares A.M.V.M., Morgado F., Guilhermino L. 2009. Acute effects of copper and mercury on the estuarine fish *Pomatoschistus microps*: linking biomarkers to behaviour. Chemosphere 76: 1416–1427.

Zinkl, J.G., Shea P.J., Nakampoto R.J. and Callman J. 1987. Technical and biological considerations for the analysis of brain cholinesterase of rainbow trout. Trans.Am. Soc. 116: 570–573.

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

Annex 8: Background Document: Histopathology of mussels (*Mytilus* spp.) for health assessment in biological effects monitoring

1. Background

Mussels have long been used for the measurement of pollutants and the biological effects of contaminants in the aquatic environment (Bayne, 1976; Goldberg, 1978; Widdows and Donkin, 1992; Granmo, 1995; Salazar and Salazar, 1995). They are widespread, sessile, possess the ability to accumulate chemicals and exhibit a wide range of biological responses. They are able to tolerate wide ranging salinity conditions and are also seen attached to piers and gravely substrata. This makes them well placed as a sentinel species in programmes designed to monitor the marine environment. Over the years numerous studies utilizing mussels have demonstrated the impact of anthropogenic inputs into the aquatic environment. Early studies such as the "Mussel Watch" programme (Goldberg, 1978) were primarily designed to evaluate pollution within coastal waters by measuring levels of pollutants within tissues of mussels (and other bivalves). In comparison, relatively few studies focused on the effect of these chemicals on their test organisms.

Over the years there has been increased emphasis placed on integrated assessments in national and international monitoring programmes within the Oslo-Paris Commission (OSPAR) region that incorporate both chemical analysis and their biological effects. A range of contaminants exist within the aquatic environment, which may elicit an assortment of biological responses. As such it is well established that integrated techniques provide a more robust approach for the overall health assessment of aquatic organisms and their environment, than the application of a single technique in isolation.

Histopathology (of aquatic organisms) is a valuable tool for providing health assessment of individuals and of populations because it incorporates measures of reproductive and metabolic condition, and allows for the detection of a range of pathogens that may affect morbidity and mortality. In addition to its role as a 'baseline' measure of health, histopathology has been employed to investigate changes related to PAH, PCB and heavy metal exposure in mussels (Sunila, 1984; Lowe and Pipe, 1987; Auffret, 1988; Kluytmans *et al.*, 1988; Marigómez *et al.*, 2006). Mussel histopathology has been designated a promising technique (tissue response) for inclusion within the "mussel integrated approach". It provides an effective set of tools for the detection and characterization of toxicopathic pathologies, which are increasingly being used as indicators of environmental stress, in addition to disease.

Histopathology is also complementary to other techniques used to monitor the biological effects of contaminants as it can help to dissociate markers of underlying health or disease condition from those associated with exposure to contaminants. The advent of genomic and post-genomic technologies increases the potential utility of histopathology in quality assurance and quality control of sample groups for analysis (e.g. by selecting homogenous groups attributes and to control for potential variation among individuals). This approach should help reduce uncertainties associated with the potential confounding effects of pathogens when trying to identify the specific effects of toxicant exposure on host gene, protein and metabolite profiles (Stentiford *et al.*, 2005; Ward, *et al.*, 2006; Hines *et al.*, 2007a). In this respect, it can be considered as a means to provide supporting information for measures (biomarkers) that specifically aim to assess historic exposure to, or effect of, a contaminant. Histopathology therefore provides a 'phenotypic anchor' against which this specific data can be assessed (Stentiford *et al*., 2005).

This ICES TIMES document provide a description of numerous health parameters that can be employed in monitoring programmes designed to assess the biological effects of contaminants. It also describes pathology that has been previously associated with contaminant exposure but may also result from exposure to pathogens.. Whereas the latter may initially seem misplaced in this document describing contaminant induced pathology, it is important to note that disease conditions of pathogen aetiology can result in pathology that may appear contaminant-related to the untrained eye. Therefore it is essential to an individual to posses the ability to be able to distinguish between contaminant- and pathogen-related pathology.

2. Sampling and dissection for formalin-fixed paraffin-embedded (FFPE) histology

When sampling mussels from the field, mussels should be carefully removed from their substratum by cutting the byssus threads with a pair of scissors. This will help to reduce stress that may act as a confounding factor when integrating with other sensitive biological effects techniques such as the Neutral Red Retention (NRR) assay. Mussels should be placed into a suitable insulating container and kept cool and moist during prompt transport back to the laboratory. This can be achieved by using a combination of ice-packs, wet paper towels and/or seaweed.

With integrated studies becoming more widespread, adopting a quality assurance approach is considered an important practice. So that potential post-sampling artefacts are minimized, mussels should be processed as soon as possible following removal from water. When dealing with samples distributed over a large geographical area (e.g. from national/international monitoring programmes), it may not always be possible to process samples immediately or relatively soon after. This is primarily because samples require lengthy transit to the laboratory thus delaying subsequent processing. Under these circumstances efforts should be made to keep the time from sampling until the time of processing, equivalent in duration between all samples. Currently, the number of individual mussels required for histology should be 50 although this may be refined in the final publication of the TIMES document for mussel histopathology (to be published imminently).

The dissection process is an extremely important stage in the histological process and it is crucial that it is conducted in a standardized manner. Standardized dissections ensure greater comparability between samples and simplify downstream histological analysis. It is essential to achieve good quality cross sections that are not too thick to ensure adequate penetration of tissues by the fixative. Mussels should be treated with care as not to cause any damage to any of the tissues. Any damage caused to tissues during dissection may prevent good quality cross sections being obtained.

In order to gain access to the visceral mass within the shell, hold the mussel with the posterior shell edge on a suitable work surface such as a dissection board. Insert scalpel blade into the midventral byssal cavity (do not insert too far as this will damage tissues situated along the dorsal shell edge) followed by a downward movement resulting in the cutting of the posterior adductor muscle. Carefully open the two shell halves to reveal the visceral mass. Using a scalpel or scissors, remove any byssus threads that may hinder any microtomy carried out at a later stage. Do not remove byssus threads by pulling (threading) as this may cause undue stress to the mussel. Starting with one shell half at first, carefully separate the mantle tissue from the inner shell surface using the flat edge of a scalpel blade. Care should be taken as not to "slice" the mantle with the scalpel blade itself. To an untrained individual, this can be challenging at first, however it is soon overcome. The most successful approach is to combine the use of "teasing" and "scraping". Brush aside the partially remove visceral mass into the remaining shell half and sever the posterior retractor muscles. Once complete the empty shell half can be removed from the remaining half by disassociation of the shell ligament (a simple twist of the empty shell will suffice). In a similar manner to previous, the mantle tissue should be teased away from the inner shell surface of the remaining shell half. This process can be made easier by resting the previously dissected tissue onto a work surface while working with the remaining tissue. Once complete the entire visceral mass should be removed from the remaining shell and placed onto a dissection board. Using a razor blade or scalpel, a slightly angled 3 mm slice across the ventral and posterior axis should be obtained towards the anterior end of the visceral mass. This will ensure that the main organs of interest (gonad, gills, mantle, digestive gland, kidney, foot) are incorporated into a single standardized section. Using forceps carefully transfer the cross section into a histo-cassette before placing into Davidson's Seawater Fixative or suitable alternative. The use of histo-cassettes is highly recommended due to their ability to ensure that the cross section remains intact during the fixation process. Allow fixation to proceed for a minimum of 24 hours with periodic agitation throughout. The use of a "rocker" facilitates this greatly.

3. Sampling and dissection for histochemistry

Histochemical techniques on frozen tissue sections (obtained by cryotomy) are needed in order to evaluate lysosomal alterations described below. As such further dissection is required when incorporating these techniques.

For cryotomy, a small cube of digestive gland should be dissected from a minimum of ten individual mussels and snap-frozen onto a cryotome chuck in two rows of five, using a suitable cryo-embedding compound such as OCT. Snap freezing can be achieved using liquid nitrogen or a commercially available cryobath. For better integration of data, it is possible to obtain frozen samples from the same mussels identified for Formalin fixed paraffin embedded (FFPE) histology. Chucks should be transported to the laboratory in dry-ice (if required) and subsequently stored at -80°C.

4. Histology

Formalin fixed paraffin embedded - histology is the most widely used histological process; however resin based embedding techniques can also be employed. For FFPE histology, tissues are dehydrated through a series of graded alcohols followed by clearing and embedding within paraffin wax. Finally, tissues are placed into moulds containing molten wax that are subsequently cooled to produce a rigid support medium (block) for microtomy. See Bignell *et al*. (2011) for detailed protocols.

Using a microtome, the face of the tissue blocks are "trimmed" or "faced" in order to expose the maximum surface area of the mussel embedded within the block. Occasionally, sand or residual byssus may be encountered during sectioning, which may prevent suitable sections being obtained. Under these circumstances, it may be possible to remove these artefacts from the block face using a small sharp implement such as a pin or needle. Care should be taken not to cause any unnecessary damage to the surrounding tissues. This ensures that all areas of interest are included during sectioning. Tissue sections are obtained at 3□m to 5 □m and floated onto a pre-heated water bath (35□C–40□C) containing a suitable tissue adhesive (e.g. StaOn, Surgipath, UK). Alternatively, commercially available slides that have been pretreated with saline or electrostatically charged can be used. Sections are adhered to a glass microscope slide by inserting the slide vertically into the water bath adjacent to floating section and lifting straight up. Following sectioning, slides should be dried overnight on a suitable hotplate. Alternatively, a section-dryer can be used which can decrease the time taken for slides to dry. Whatever drying method is employed, it is important to ensure that all moisture has been removed from slides prior to staining. Subsequently, sections are stained with haematoxylin and eosin (protocol provided in Annex 3) or a suitable alternative. Following staining, the end result should represent Figure X. This approach produces a uniform histological section that (a) incorporates all of the target organs of interest and (b) makes for a more simple microscopic examination due to the standardized orientation of the tissues and organs. Using a low magnification objective, the histopathologist should scan the slide for any abnormalities before further examination at higher magnifications. It is recommended to observe slides "blind" i.e. without prior knowledge to geographical location or exposure groups, in order to reduce bias that may otherwise be introduced to the interpretation.

Detailed sampling procedures are outlined in the ICES TIMES document.

5. Quality assurance

At present there is no quality assurance scheme in place for mussel histopathology. It is envisaged that this will be run in a similar manner to the BEQUALM Fish Disease Programme currently organized by Cefas.

6. Health parameter measurements

The following parameters can be measured quantitatively or semi quantitatively with histological techniques, cell type composition in digestive gland epithelium, digestive tubule epithelial atrophy and thinning, lysosomal alterations and inflammation and are described in detail below.

7. Cell type composition in digestive gland epithelium

Under normal physiological conditions the digestive cells outnumber basophilic cells, but under different stress situations, including exposure to pollutants, the relative occurrence of basophilic cells is apparently augmented (Rasmussen *et al.*, 1985; Lowe and Clarke, 1989; Cajaraville *et al.*, 1990, Marigómez *et al.*, 1990, 1998; 2006; Zorita *et al.*, 2006; 2007; Garmendia *et al.*, 2011b). Changes in cell type composition in the digestive gland epithelium constitute a common response in molluscs that may lead to disturbances in food digestion and xenobiotic metabolism and accumulation Marigómez *et al.*, 1998). These changes have been attributed to basophilic cell proliferation (Lowe and Clarke, 1989; Cajaraville *et al.*, 1989; Marigómez *et al.*, 1990), but it has been recently concluded that it mainly results from digestive cell loss and basophilic cell hypertrophy (Zaldibar *et al.*, 2007), which is a fast inducible and reversible response that can be measured in terms of volume density of basophilic cells (*VvBAS*). In clean localities and in experimental control conditions, *VvBAS* is usually below 0.1 $\mu m^3/\mu m^3$ but after exposure to pollutants *VvBAS* may surpass 0.12 $\mu m^3/\mu m^3$ (Marigómez *et al.*, 2006).

A stereological procedure is applied in order to quantify the volume density of basophilic cells (*VvBAS*) as a measure of digestive cells loss by counting on H/E stained digestive gland paraffin sections (Soto *et al*., 2002). Cell counts (digestive and basophilic cells) are made in one field randomly selected per mussel (n=10) to complete a total of ten counts per experimental group, with the aid of a drawing tube attached to a light microscope using a 20x objective lens. A Weibel graticule (multipurpose test system M-168) is used, and hits on basophilic cells and on remaining digestive epithelium are recorded to calculate *VvBAS* according to the Delesse's principle:

VvBAS ($\mu m^3/\mu m^3$)= x/(m+x);

where "x" is the number of hits on basophilic cells and "m" is the number of hits on digestive cells. The statistical signification of changes in *VvBAS* volume is determined according to parametric tests (e.g. ANOVA, Duncan's test for comparison between pairs of means; $p<0.05$). Assessment criteria should be considered as:

Background:	<0.12 $\mu m^3/\mu m^3$
Elevated:	0.12-0.18 $\mu m^3/\mu m^3$
High:	>0.18 $\mu m^3/\mu m^3$

8. Digestive tubule epithelial atrophy and thinning

The best documented cellular alteration in bivalves is apparent atrophy or "thinning" of the digestive gland epithelium. The digestive gland of mussels is greatly dynamic and plastic. The morphology of digestive alveoli undergoes severe changes even during normal physiological processes (i.e. trough every digestion cycle; Langton, 1975). Changes in the normal phasic activity may be attributed to environmental factors, such as food availability or saline and thermal stress (Winstead, 1995) as well as exposure to pollutants. Particularly, it has been widely demonstrated that molluscs exposed to pollutants exhibit a net mass loss in the digestive gland epithelium that gives rise to abnormal epithelial thinning and finally atrophy (Lowe *et al*., 1981; Couch, 1984; Lowe and Clarke, 1989; Vega *et al*., 1989; Cajaraville *et al*., 1992; Marigómez *et al*., 1993; Garmendia *et al*., 2011b). Atrophy and epithelial thinning constitute a non-specific fast inducible and slowly or not recoverable response to stressful environmental conditions that can be measured after semi-quantitative scoring (Kim *et al*., 2006) or after quantitative morphological analysis in terms of *MPTW* (mean proportion of tubule width; Robinson, 1983); or in terms of mean epithelial thickness (*MET*) and the relative parameters *MLR*/*MET* and *MET*/*MDR* (Lowe *et al*., 1981; Vega *et al*., 1989; Cajaraville *et al*., 1992; Marigómez *et al*., 1993; 2006; Garmendia *et al*., 2011b), where *MLR* is the mean luminal radius and MDR the mean diverticular radius. *MLR*/*MET* ratio is more sensitive than *MET* alone. The alterations in these parameters are used as tissue-level biomarkers in ecosystem health assessment (Garmendia *et al*., 2011b).

The following table describes a semi-quantitative scoring index for digestive tubule epithelial atrophy and thinning*.

Stage	Response	Description
0	None	Normal tubule thickness (0% atrophy). Lumen nearly occluded, few tubules exhibiting slight atrophy.
1	Low	Epithelium averaging less than one-half (50%) normal thickness (stage 0), most tubules show some atrophy although some tubules appear normal.
2	Elevated	Epithelium averaging about 50% of normal thickness (stage 0).
3	High	Epithelium thickness greater than one-half (50%) atrophied, most tubules affected. Some tubules extremely thin (fully atrophied).
4	Severe	Epithelium extremely thin (100% atrophied), nearly all tubules affected.

***adapted from Ellis (1996).**

Most commonly, a planimetric procedure has been applied to quantify changes in size and shape of the digestive alveoli (Vega *et al*., 1989) resulting in apparent epithelial thinning. A total of 50–100 tubular profiles per sample (two profiles per field in five fields per mussel in 5–10 mussels per sample) are recorded in an image analysis system attached to light microscope using a 20x objective lens. The five measurement fields are selected at given intervals throughout the tissue section, the direction of movement always following a zigzag pattern. Alternatively tubular profiles can be drawn with the aid of a drawing tube attachment to the light microscope and then digitized for data input into a computer. Other methods are also available because the final goal is just calculating the section areas of the lumen and the whole tubule profile, which can be done by image analysis systems (after data input into the computer), by hand (e.g. using millimetre paper), or by point counting onto a Weibel stereological graticle (Weibel ER, 1979). *MET*, *MLR* and *MDR* are quantified (in µm) and the ratios *MLR/MET* and *MET/MDR* (in µm/µm) are calculated as integrative measures of changes in the alveolar morphology, epithelial thinning included, as follows:

$$MET = 2(Ao - Ai)/(Po + Pi);$$
$$MLR = \sqrt{(Ai/\pi)}; \text{ and}$$
$$MDR = \sqrt{(Ao/\pi)};$$

where *Ao* is the section area of the whole tubule profile, *Po* is the perimeter of a circle with area *Ao*, *Ai* is the section area of the lumen profile and *Pi* is the perimeter of the corresponding circle with area *Ai*. The statistical signification of changes in these parameters is determined according to parametric tests (e.g. ANOVA, Duncan's test for comparison between pairs of means; $p<0.05$). *MLR/MET* values between 0.7 µm/µm (spring-summer) and 1.2 µm/µm (winter) have been recorded in *M. galloprovincialis* of reference localities in Southern Bay of Biscay, whereas after exposure to pollutants or stress in long-term laboratory manipulation *MLR/MET* surpasses 1.6 µm/µm (Marigómez *et al*., 2006).

9. Lysosomal alterations

Lysosomal responses are widely used as effect biomarkers indicative of the general stress provoked by pollution in the marine environment. Lysosomes are cell organelles containing acid hydrolases. The digestive cells of mussels possess a complex endo-lysosomal system that is primarily involved in the uptake and digestion of food materials as well as in processes of pollutant accumulation and detoxification. Endolysosomes and heterolysosomes occupy the majority of digestive cell cytoplasm and are reactive for marker hydrolases such as N-acetyl hexosaminidase, β-glucuronidase and acid phosphatase (Izagirre and Marigomez 2009a; Izagirre and Marigomez 2009b). Lysosomal responses to environmental stress fall into essentially three categories: increased lysosomal size, reduced membrane stability, and changes in lysosomal contents (Marigσmez and Baybay-Villacorta, 2003).

Lysosomal enlargement

Diverse sources of environmental stress (chemical pollution, salinity changes, elevated temperature, malnutrition, reproductive stress) are known to provoke an increase in the size of digestive cell lysosomes in mussels, often accompanied by increased enzyme activity and lysosome numbers, which may compromise intracellular digestion and detoxification capacity (Moore, 1985; 1988; Lowe, 1988; Cajaraville *et al*., 1989; 1995; Marigómez *et al*., 1995; 2005; 2006; Domoutsidou and Dimitriadis, 2001; Garmendia *et al*., 2011a). These lysosomal structural changes (LSC) have been commonly determined by image analysis of digestive gland cryotome sections where β-glucuronidase is employed as lysosomal marker enzyme. The final calculations of the structural parameters are in most cases based on the equations published by Lowe *et al*. (1981). The structural parameters are lysosomal volume density (*Vv*), surface density (*Sv*), surface-to-volume ratio (*S/V*) and numerical density (*Nv*). Although the four stereological parameters altogether provide complete information about the size, size class distribution and number of lysosomes in mussel digestive cells, Vv can be sufficient to detect changes in the size of the endo-lysosomal system and is therefore the most used parameter.

Stereological determination of lysosomal enlargement

The histochemical reaction for β-Gus is demonstrated as in Moore (1976) with the modifications described by Cajaraville *et al*. (1989). Slides are kept at 4°C for 30 minutes and then at RT for 5 minutes prior to staining. Sections (8 μm) are incubated in freshly prepared β-Gus substratum incubation medium consisting of 28 mg naphthol AS-BI-β-glucuronide (Sigma, N1875) dissolved in 1.2 ml 50 mM sodium bicarbonate, made up to 100 ml with 0.1 M acetate buffer (pH 4.5) containing 2.5% NaCl and 15% polyvinyl alcohol, for 40 minutes at 37°C in a shaking water bath. After incubation, slides are rinsed in a 2.5% NaCl solution for 2 minutes at 37°C in a shaking water bath and then transferred to a postcoupling medium containing 0.1 g Fast garnet GBC (Sigma, F8716) dissolved in 100 ml 0.1 M phosphate buffer (pH 7.4 containing 2.5% NaCl) for 10 minutes in the dark and at RT. Afterwards, the sections are fixed for 10 minutes at 4°C in Baker's formol calcium containing 2.5% NaCl and rinsed briefly in distilled water. Finally, sections are counterstained with 0.1% Fast green FCF (Sigma, F7252) for 2 minutes, rinsed several times in distilled water, mounted in Kaiser's glycerine gelatine and sealed with nail varnish. Then, *de visu* grading and scoring can be applied to grossly determine the extent of lysosomal enlargement (Lowe, 1988), which can be straightforward and very useful in cases of extreme symptoms. However, quantifying lysosomal enlargement by hand stereology (Cajaraville *et al*., 1989; 1992) or by image analysis (Marigómez *et al*., 2005; Izagirre and Marigómez, 2009) can provide evidence of more subtle lysosomal responses. Slides are viewed under a light microscope fitted with a ×100 objective lens. A Weibel graticule (multipurpose test system M-168) is used, and hits on digestive cell lysosomes and on digestive cell cytoplasm are recorded to calculate *VvLYS*, *SvLYS*, S/*VLYS*, and *Nv LYS* according to Lowe *et al*. (1981). Five measurements are made per section in each of the 5–10 individuals per sample. The stereological formulae include a correction factor for particles with an average diameter smaller than the section thickness (Lowe *et al*., 1981).

For this reason the average diameter of at least 90 lysosomes must be directly measured at the light microscope with the aid of a graded eyepiece or similar device (or directly by the image analysis system):

VvLYS ($\mu m^3/\mu m^3$) = K×AA;
SvLYS ($\mu m^2/\mu m^3$) = (4/t)×AA;
S/VLYS (μm^{-1}) = 4/(t×K) ; and
NvLYS ($\mu m^2/\mu m^3$) = (4×AA×n) / (t×π×Σ Yi2)

being

AA = x/m and K = (2/(3×t))(ΣYi3/ΣYi2);

and where "*x*" is the number of hits on digestive cell lysosomes, "*m*" is the number of hits on digestive cells (lysosomes included), "*t*" is the section thickness (i.e. 8 μm), "*n*" is the number of lysosomes whose diameter has been measured; and "Y" are lysosomal diameters (Y1, Y2, … Y90 for n=90).

Lysosomal structural changes test parameters can be tested using analysis of variance. *VvLYS* and *Nv LYS* data may need to be logarithmically transformed previous to the statistical analyses because the variance within individuals may depend on the mean. Parametric tests for multiple comparisons between paired means (e.g. Duncan's test) can be further applied to detect significant ($P<0.05$) differences between means.

In general terms, lysosomes become enlarged under stress conditions, which are reflected as increased in *VvLYS* and *SvLYS* values, concomitant with lowered *S/VLYS* values (Cajaraville *et al*., 1995; Marigómez *et al*., 2005). In certain cases, lysosomal enlargement is accompanied by increased *Nv LYS*, (increased numbers of lysosomes relative to digestive cell cytoplasm) but reductions in *Nv LYS* have also been reported. On the other hand, exposure to pollutants may also elicit an intricate response that includes different phases (Marigómez and BayBay-Villacorta, 2003): (a) transient lysosomal enlargement; (b) transient lysosomal size reduction; and finally (c) lysosomal enlargement after long-term exposure. Overall, reference values for these lysosomal parameters vary with season but *VvLYS* >0.002 $\mu m^3/\mu m^3$ and S/VLYS >5 may be indicative of the existence of a degraded health status in mussels that correlates with e.g. the degree of exposure to pollutants.

10. Inflammation

Inflammation affects all tissues and organs and is particularly obvious in mussels that have been adversely affected by contaminants (Auffret, 1988; Crouch, 1985). While this may be true, it is important to remember that the presence of pathogens can also result in a host immune response (but not always) manifested as inflammation. Inflammation is observed as either diffuse, focal or both in appearance throughout the vesicular connective tissue and at varying degrees of severity.

Haemocytic infiltration is generally characterized by the infiltration of granulocytes possessing an eosinophilic cytoplasm into the connective tissues. Care should be taken not to confuse this with normal circulating haemocytes that are often situated around the stomach and intestine. Heavy diffuse inflammation will appear as a marked increase in the number of circulating haemocytes situated throughout the majority of connective tissues and in between organs such as the digestive diverticula and gonad. Haemocytic infiltration of the visceral mass in bivalves is generally considered to be indicative of stress, unrecognized injury or sub-microscopic agents in bivalves. Haemocytic infiltration could be interpreted as a repair process following tissue damage, albeit pathological effects could be exerted through acting as space occupying lesions. Its presence has been suggested as a qualitative or quantitative index of stress, indicative of a loss of condition. Previous studies have reported haemocytic infiltration in response to starvation and spawning stress, shell damage, and exposure to pollutants.

Brown cell (BC) aggregates (foci) are generally small and posses varying quantities of the pigment lipofuscin and are often seen in elevated numbers in mussels from contaminated environments. Consisting of serous cells, these phagocytes are mostly found within the connective tissue and posses the ability to physically remove endocytosed matter across epithelia via diapedesis. These cells are responsible for the metabolism of metal ions and can be found within the gills, which is an important organ for metal ion exchange (Marigomez *et al.*, 2002). The occurrence of BC aggregates (foci) has been considered an indicator of stress caused by xenobiotics, as well as with age and reproductive stress. BC aggregates are also observed within the gonad follicles following spawning, which is a normal event.

Large foci of inflammation termed granulocytomas (consists of granulocytes), have previously been seen in mussels of both laboratory and field studies designed to monitor the effects of contaminants. Granulocytomas represent an inflammatory response to an irritant or pollutant, resulting in vascular occlusions. They are believed to result from chronic exposure to domestic and industrial waste products and have been reported in bivalves subjected to the impact of oil, chlorinated pesticides and heavy metals. Granulocytomas are also associated with pathogens therefore it is important to look for any indication of infection in affected individuals. These lesions can be seen at varying degrees of severity from singular foci to large numbers affecting the majority of the connective tissues. Granulocytomas can vary largely in size. In mussels, the maximum size of a known parasitically induced granulocytoma is 400 μm, however granulocytomas of unknown aetiology can be over 800 μm (up to 1500 μm).

The following table describes a semi-quantitative scoring index for inflammation.

Stage	Response	Description
0	None	No inflammatory foci can be seen within tissues. Brown cell foci rare.
1	Low	Small numbers of inflammatory foci occupying ≤ 10% of the vesicular connective tissue (approximately 20 small foci) within standardized tissue cross section. Brown cell foci rare.
2	Elevated	Increased numbers and/or size of inflammatory foci occupying between 10% and 50% of vesicular connective tissue. Foci may displace other structures. Areas of diffuse haemocyte infiltration may also be present. Increased numbers of Brown cell foci predominately within the vesicular connective tissue, stomach and digestive gland epithelium.
3	High	Significant inflammatory response - numerous and/or large inflammatory foci (possibly with granulocytoma present) occupying ≥ 75% of vesicular connective tissue. Widespread diffuse haemocytic infiltration may be present. Increased numbers of Brown cell foci predominately within the vesicular connective tissue, stomach and digestive gland epithelium. Increased pigment density.

11. Assessment criteria

Several parameters have been identified as suitable for the development of assessment criteria. Other histological parameters can also be measured using histopathology, although many of these fluctuate showing clear seasonal cycles (Bignell *et al.*, 2008). As such the development of assessment criteria is not deemed appropriate. Nonetheless, the collection of these data can provide additional information on the health and physiology of the mussel. Parameters include reproductive markers such as adipogranular cells, gonadal apoptosis, atresia, hermaphroditism and intersex. All health parameters are described in full detail in the ICES TIMES document (Bignell *et al.*, 2011).

The thresholds identified here have been determined using data collected as part of previous studies (Cajaraville *et al.*, 1992; Marigomez *et al.*, 2004; Marigomez *et al.*, 2005; Marigomez *et al.*, 2006; Bignell *et al.*, 2008). It must be stressed that these thresholds are preliminary and will require further review as part of a holistic assessment of these histological parameters.

BIOLOGICAL EFFECT	QUALIFYING COMMENTS	BACKGROUND	ELEVATED	HIGH
Mussel histopathology	VVbas: Cell type composition of digestive gland epithelium (quantitative)	<0.12 µm3/µm3	0.12–0.18 µm3/µm3	>0.18 µm3/µm3
	MLR/MET: Digestive tubule epithelial atrophy and thinning (quantitative)	<0.7 µm/µm	1.2–1.6 µm/µm	>1.6 µm/µm
	VVLYS & S/VLYS: Lysosomal enlargement* (quantitative)	VvLYS <0.0002 µm3/µm3 S/VLYS > 4 µm2/µm3	0.0002–0.0004 µm3/µm3 S/VLYS < 4 µm2/µm3	V>0.0004 µm3/µm3 S/VLYS <<4 µm2/µm3
	Digestive tubule epithelial atrophy and thinning (semi-quantitative)	STAGE ≤1 (Mode)	STAGES 2–3 (Mode)	STAGE 4 (Mode)
	Inflammation (semi-quantitative)	STAGE ≤1 (Mode)	STAGE 2 (Mode)	STAGE 3 (Mode)

12. References

Auffret, M. 1988. Histopathological changes related to chemical contamination in *Mytilus edulis* from field and experimental conditions. Mar Ecol Prog Ser 46: 101–107.

Bayne, B.L. 1976. Watch on mussels. Mar Pollut Bull 7(12): 217–128.

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

Bignell, J.P., Caraville, M., Feist, S.W., Villalba, A., Marigomez, I. (in prep) Histopathology of mussels *Mytilus* sp. for health assessment in biological effects monitoring. ICES Techniques in Marine Environmental Sciences No. XX.

Cajaraville MP, Marigómez JA, Angulo E. 1989. A stereological survey of lysosomal structure alterations in *Littorina littorea* exposed to 1-naphthol. Comp Biochem Physiol 93C: 231–237.

Cajaraville MP, Díez G, Marigómez IA, Angulo E. 1990. Responses of basophilic cells of the digestive gland of mussels to petroleum hydrocarbon exposure. Diseases Aquatic Organisms 9, 221–228.

Cajaraville, M.P., Marigómez, I., Díez, G., Angulo, E. 1992. Comparative effects of the water accommodated fraction of three oils on mussels. 2. Quantitative alterations in the structure of the digestive tubules. Comp. Biochem. Physiol. 102 (C): 113–123.

Cajaraville MP, Abascal I, Etxeberria M, Marigómez I. 1995. Lysosomes as cellular markers of environmental pollution: time- and dose-dependent responses of the digestive lysosomal system of mussels after petroleum hydrocarbon exposure. Environ Toxicol Water Qual 10:18.

Couch J. 1984. Atrophy of diverticular epithelium as an indicator of environmental irritants in the oyster *Crassostrea virginica*. Mar. Environ. Res. 14: 525–526.

Couch JA. 1985. Prospective study of infectious and noninfectious diseases in oysters and fishes in three Gulf of Mexico estuaries Diseases of Aquatic Organisms 1: 59–82.

Domouhtsidou GP, Dimitriadis VK. 2001. Lysosomal and lipid alterations in the digestive gland of mussels, *Mytilus galloprovincialis* (L) as biomarkers of environmental stress. Environ Pollut;115:123–137.

Garmendia L, Izagirre U, Cajaraville MP and Marigomez I. 2011a. Application of a battery of biomarkers in mussel digestive gland to assess long-term effects of the Prestige oil spill in Galicia and the Bay of Biscay: Lysosomal responses, J. Environ. Monit., DOI: 10.1039/c0em00409j.

Garmendia L, Soto M, Vicario U, Kim Y, Cajaraville MP and Marigómez I. 2011b. Application of a battery of biomarkers in mussel digestive gland to assess long-term effects of the Prestige oil spill in Galicia and the Bay of Biscay: Tissue-level biomarkers and histopathology, J. Environ. Monit., DOI: 10.1039/c0em00410c.

Goldberg, E.D., Bowen, V.T., Farrington, J.W., Harvey, G., Martin, J.H., Parker, P.L., Risebrough, R.W., Robertson, W., Schneider, E., Gamble, E. 1978. The Mussel Watch Environ Conserv 5 (2): 101–125.

Granmo, A. 1995. Mussels as a tool in impact assessment. Phuket Mar Biol Cent Spec Publ, 15: 215–220.

Hines, A. Oladiran , G., Bignell, J., Stentiford, G.D., Viant, M.R. 2007a. Direct sampling of organisms from the field and knowledge of their phenotype: Key recommendations for environmental metabolomics. Environ Sci Tech 41: 3375–3381.

Izagirre U, Ramos R, Marigómez I. 2008. Natural variability in size and membrane stability of lysosomes in mussel digestive cells: season and tidal zonation. Mar Ecol Prog Ser 372; 105–117.

Izagirre U, Marigómez I. Lysosomal enlargement and lysosomal membrane destabilisation in mussel digestive cells measured by an integrative index. 2009a. Environ Pollut; 157: 1544–1553.

Izagirre U, Angulo E, Wade SC, ap Gwynn I, Marigómez I. 2009b. β-Glucuronidase and hexosaminidase are marker enzymes for different compartments of the endo-lysosomal system in mussel digestive cells. Cell and Tissue Research 335: 441–454.

Kim Y, Ashton-Alcox A, Powell EN. 2006. Histological techniques for marine bivalve molluscs: update NOAA Technical Memorandum NOS NCCOS 27. p. 76.

Kluytmans, J.H., Brands, F., Zandee, D.I. 1988. Interactions of Cadmium with the Reproductive Cycle of *Mytilus edulis L.* Mar Environ Res 24: 189–192.

Langton RW. 1975. Synchrony in the digestive diverticula of *Mytilus edulis* L. J. Mar Biol Ass UK 55:221–229.

Lowe, D.M., and Pipe, R.K. 1987. Mortality and quantitative aspects of storage cell utilization in mussels, *Mytilus edulis*, following exposure to diesel oil hydrocarbons. Mar Environ Res 22: 243–251.

Lowe, D.M., Moore, M.N., Clarke, K.R. 1981. Effects of oil on digestive cells in mussels: Quantitative alterations in cellular and lysosomal structure. Aquatic Toxicology 1, Issue 3–4, 213–226.

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

Lowe, D.M., Clarke, K.R. 1989. Contaminant-induced changes in the structure of the digestive epithelium of *Mytilus edulis*. Aquatic Toxicology 15, Issue 4, 345–358.

Marigómez I, Cajaraville MP, Angulo E. 1990. Histopathology of the digestive gland-gonad complex of the marine prosobranch *Littorina littorea* (L) exposed to cadmium. Dis. Aquat. Toxicol. 9: 229–238.

Marigómez I, Soto M, Etxeberria M, Angulo E. 1993. Effects of size, sex, reproduction, and trematode infestation on the quantitative structure of digestive tubules in stressed winkles. Zool. Jb. Anat. 123: 319–336.

Marigómez I, Cajaraville MP, Soto M, Lekube X. 1998. Cell-type replacement, a successful strategy of molluscs to adapt to chronic exposure to pollutants. Cuad. Invest. Biol. 20: 255–258.

Marigómez I, Soto M, Cajaraville MP, Angulo E, Giamberini L. 2002. Cellular and subcellular distribution of metals in molluscs. Micros. Res. Tech. 56: 358–392.

Marigómez, I., Baybay-Villacorta, L. 2003. Pollutant-specific and general lysosomal responses in digestive cells of mussels exposed to model organic chemicals. Aquatic Toxicology 64, Issue 3, Pages 235–257.

Marigómez, I., Soto, M., Orbea, A., Cancio, I., Cajaraville, M.P. 2004. Chapter 14: Biomonitoring of environmental pollution in the Basque coast using molecular, cellular and tissue-level biomarkers: an integrative approach. In: Borja, A., Collins, M. (Eds.), Oceanography and marine environment in the Basque Country. Elsevier Oceanography Series N. 70, Elsevier, Amsterdam, pp. 335–364.

Marigómez, I., Izagirre, U., Lekube, .X. 2005. Lysosomal enlargement in digestive cells of mussels exposed to cadmium, benzo[a]pyrene and their combination. Comp Biochem Physiol 2005a; 141:188–193.

Marigómez, I., Soto, M., Cancio, I., Orbea, A., Garmendia, L., and Cajaraville, M.P. 2006. Cell and tissue biomarkers in mussel, and histopathology in hake and anchovy from Bay of Biscay after the Prestige oil spill (Monitoring Campaign). Mar Pollut Bull 53: 287–304.

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

Moore MN. 1985. Cellular responses to pollutants. Mar Pollut Bull 16:134–139.

Moore MN. 1988. Cytochemical responses of the lysosomal system and NADPH-ferrihemoprotein reductase in mollusc digestive cells to environmental and experimental exposure to xenobiotics. Mar Ecol Prog Ser 46:81–89.

Rasmussen L.P.D., Hage E., Karlog O. 1985. Light and electronmicroscopic studies of the acute and long-term toxic effects of N-nitrosodipropylamine and N-methylnitrosurea on the marine mussel *Mytilus edulis*. Mar Biol 1985; 85: 55–65.

Robinson WE. 1983. Assessment of bivalve intracellular digestion based on direct measurement. J. Moll. Stud. 49: 1–8.

Salazar, M.H., Salazar, S.M. 1995. In-situ bioassays using transplanted mussels: I. Estimating chemical exposure and bioeffects with bioaccumulation and growth. ASTM Special Technical Publication, (1218), pp. 216–241.

Soto M, Zaldibar B, Cancio I, Taylor MG, Turner M, Morgan JA, Marigómez I. 2002. Subcellular distribution of cadmium and its cellular ligands in mussel digestive gland cells as revealed by combined autometallography and X-ray microprobe analysis. Histochem. J. 34: 273–280.

Stentiford, G.D., Viant, M.R., Ward, D.G., Johnson, P.J., Martin, A., Wei, W., Cooper, H.J., Lyons, B.P., Feist, S.W. 2005. Liver tumours in wild flatfish: a histopathological, proteomic, and metabolomic study. Omics: A Journal of Integrative Biology, 9(3): 281–299.

Sunila, I. 1984. Copper- and cadmium- induced histological changes in the mantle of *Myituls edulis* L. (Bivalvia). Limnologica, 15(2), 523–527.

Vega MM, Marigómez I, Angulo E. 1989. Quantitative alterations in the structure of digestive cells of *Littorina littorea* on exposure to cadmium. Mar. Biol. 103: 547–553.

Ward D.G., Wei W, Cheng Y, Billingham LJ, Martin A, Johnson, PJ, Lyons BP, Feist SW, Stentiford G.D. 2006. Plasma proteome analysis reveals the geographical origin and liver tumour status of dab (*Limanda limanda*) from UK marine waters. Environ Sci Tech 40: 4031–4036.

Weibel E.R. 1979. Stereological methods. Vol 1. Acad Press, London, 415 pp.

Widdows, J., and Donkin, P. 1992. Mussels and environmental contaminants: Bioaccumulation and physiological aspects. In: Gosling, The mussel *Mytilus*: Ecology, physiology, genetics and culture. Elsevier, Amsterdam, p 383–464.

Winstead JT. 1995. Digestive tubule atrophy in Eastern oysters, *Crassostrea virginica* (Gmelin, 1791), exposed to salinity and starvation stress. J. Shellfish Res. 14: 105–111.

Zaldibar B, Cancio I, Marigómez I. 2007. Reversible alterations in epithelial cell turnover in digestive gland of winkles (*Littorina littorea*) exposed to cadmium and their implications for biomarkers measurements. Aquat. Toxicol. 81: 183–196.

Zorita I, Ortiz-Zarragoitia M, Soto M, Cajaraville MP. 2006. Biomarkers in mussels from a copper site gradient (Visnes, Norway): An integrated biochemical, histochemical and histological study. Aquat Toxicol ;78S:S109–S116.

Zorita I, Apraiz I, Ortiz-Zarragoitia M, Orbea A, Cancio I, Soto M, Marigómez I, Cajaraville MP. 2007. Assessment of biological effects of environmental pollution along the NW Mediterranean Sea using mussels as sentinel organisms. Environ Pollut. 148:236–250.

Annex 9: Background document: Micronucleus assay as a tool for assessing cytogenetic/DNA damage in marine organisms

Authors: Janina Baršienė[1] and Brett Lyons[2]

The primary authors acknowledge contributions from Aleksandras Rybakovas[1], Concepción Martínez-Gómez[5], Steinar Sanni[3], Steven Brooks[4], Beatriz Fernández[5], and Thomas Maes[6]

[1]Institute of Ecology of Nature Research Center (NRC), Vilnius, Lithuania
[2] Centre for Environment Fisheries and Aquaculture Science (Cefas), Weymouth, UK
[3] International Research Institute of Stavanger, (IRIS)–Akvamiljø, Randaberg, Norway
[4]Norwegian Institute for Water Research (NIVA), Oslo, Norway
[5] Instituto Espeañol de Oceanografía (IEO), Centro Oceanográfico de Murcia, Spain
[6] Centre for Environment Fisheries and Aquaculture Science (Cefas), Lowestoft, UK

Background

1. Micronuclei (MN) consist of acentric fragments of chromosomes or whole chromosomes which are not incorporated into daughter nuclei at anaphase. These small nuclei can be formed as a consequence of the lagging of a whole chromosome (aneugenic event) or acentric chromosome fragments (clastogenic event) (Heddle, 1973; Schmid, 1975). A micronucleus (MN) arises in cell divisions due to spindle apparatus malfunction, the lack or damage of centromere or chromosomal aberrations (Fenech, 2000).

Clastogens induce MN by breaking the double helix of DNA, thereby forming acentric fragments that are unable to adhere to the spindle fibers and integrate in the daughter nuclei, and are thus left out during mitosis. Aneuploidogenic agents are chemicals that prevent the formation of the spindle apparatus during mitosis which can generate not only whole chromatids that are left out of the nuclei, thus forming MN, but also can form multinucleated cells in which each nucleus would contain a different number of chromosomes (Serrano-García and Montero-Montoya, 2001). Thus, the scoring during interphase provides a measure of genotoxicity both in the field and also specifically through genotoxic compound exposure in the laboratory due to clastogens and/or aneugens (Al-Sabti and Metcalfe, 1995; Heddle *et al*., 1991). In addition, there are direct indications that MN additionally may be formed via a nuclear budding mechanizm in the interphase of cell division. The formation of such type MN reflects in an unequal capacity of the organisms to expel damaged, amplified, failed replicated or improperly condensed DNA, chromosome fragments without telomeres and centromeres from the nucleus (Lindberg *et al*., 2007).

2. The micronuclei assay involves the scoring of the cells which contain one or more micronuclei in the cytoplasm (Schmid, 1975). The assay was first developed as a routine in vivo mutagenicity assay for detecting chromosomal mutations in mammalian studies (Boller and Schmid, 1970; Heddle, 1973). Hooftman, de Raat (1982) were the first to successfully apply the assay to aquatic species when they demonstrated the induction of micronuclei in erythrocytes of the eastern mudminnow (*Umbra pygmaea*) following waterborne exposure to the known mutagen ethyl methanosulphate (EMS). Since these initial experiments, other studies have validated the detection of micronuclei as a suitable biomarker of genotoxicity in a wide range of both vertebrate and invertebrate species (for review see Chaudhary *et al*., 2006; Udroiu *et al*., 2006; Bolognesi and Hayashi, 2011). In fish most studies have utilized circulating erythrocytes

(blood) cells but can also be sampled from a number of tissues, such as liver, kidney, gill or fin epithelium (Archipchuk, Garanko, 2005; Baršienė *et al*., 2006a; Rybakovas *et al*., 2009).

3. Environmental genotoxicity levels in organisms from North Sea, Mediterranean and northern Atlantic have been described in indigenous fish and mussel species inhabiting reference and contaminated sites (Wrisberg *et al*., 1992; Bresler *et al*., 1999; Baršienė *et al*., 2004, 2008a, 2010a; Bagni *et al*., 2005; Bolognesi *et al*., 2006b; Magni *et al*., 2006; Fernandez *et al*., 2011). Concerns about the environmental genotoxicity in an oil and gas industrial areas of the North Sea were raised when comparatively high levels of micronuclei incidences were detected in mussels *Mytilus edulis* and Atlantic cod *Gadus morhua* caged closely to the oil platforms (Hylland *et al*., 2008). Increased environmental genotoxicity and cytotoxicity has been described in an offshore Ekofisk oil extraction field (Rybakovas *et al*., 2009). The Water Column Monitoring Programme indicated increased genotoxicity in caged mussels in sites that were close to the Ekofisk oil platform indicating the ability to pinpoint source discharges with genotoxic endpoints in caged mussels (Baršienė, IRIS WCM Reports 2006, 2008; Brooks *et al*., 2011). Significant MN elevation in fish and mussels was found after exposure to the crude oil extracted from the North Sea (Baršienė *et al*., 2006a; Bolognesi *et al*., 2006; Baršienė, Andreikėnaitė, 2007; Andreikėnaitė, 2010) and from arctic zones (Baršienė *et al*., unpublished data).

4. The frequency of the observed micronuclei may be considered as a suitable index of accumulated genetic damage during the cell lifespan providing a time integrated response of an organism's exposure to contaminant mixtures. Depending on the lifespan of each cell type and on their mitotic rate in a particular tissue, the micronuclei frequency may provide early warning signs of cumulative stress (Bolognesi and Hayashi, 2011). The exposure of caged mussels in the Genoa harbour, heavily polluted by aromatic hydrocarbons showed a continuous increase of micronuclei in mussel gill cells reaching a plateau after a month of caging (Bolognesi *et al*., 2004). After 30 days caging of mussels at the Cecina estuary in Tyrrhenian coast, twofold increase of MN incidences in gill cells has been observed (Nigro *et al*., 2006). The gradient-related increase in MN was found in haemocytes of mussels and liver erythrocytes of Atlantic cod caged for 5–6 months at Norwegian oil platforms in the North Sea (Hylland *et al*., 2008, Brooks *et al*., 2011). Furthermore recovery was detected in the Haven oilship sinking zone using the MN test in caged mussels ten years after the oil spill (Bolognesi *et al*., 2006b). In this respect, increase in micronuclei frequency represents a time integrated response to cumulative stress.

Short description of methodology

5. Target species

Micronuclei frequency test has generally been applied to organisms where other biological effects, techniques and contaminant levels are well documented. That is the case for mussels and for certain demersal fish species (as European flounder, dab, Atlantic cod or red mullet), which are routinely used in biomonitoring programmes and assess contamination along western European marine waters (see Table 1). However, the MN assay may be adapted for alternative sentinel species using site-specific monitoring criteria.

When selecting an indicator fish species, consideration must be given to its karyotype as many teleosts are characterized by an elevated number of small chromosomes (Udroiu *et al*., 2006). Thus, in certain cases micronuclei formed after exposure to clastogenic contaminants will be very small and hard to detect by light microscopy. This can be addressed to a certain extent by using fluorescent staining. After selecting target/suitable species, researchers should also ensure that other factors including age, sex, temperature and diet are similar between the sample groups. If conducting transplantation studies, consideration needs to be given to the cellular turnover rate of the tissue being examined to ensure sufficient cells have gone through cell division. For example, if using blood the regularities of erythropoiesis should be known prior to sampling.

In general, indigenous, ecologically and economically important fish and mollusc species could serve as indicator species for biomonitoring of environmental genotoxicity levels, for screening of genotoxins distribution or for assessments of genotoxicity effects from contaminant spills or effluent discharges. For monitoring in deep waters in northern latitudes (deeper than 1000 m), the fish Arctic rockling *Omogadus argentus* and amphipods *Eurythenes gryllus* are suitable species. In equatorial regions of the Atlantic, indicator fish species *Brachydetrius aurectus, Synoglosus senegalensis, Cynoponticus ferox* are available for the MN analysis (Baršienė IRIS reports for Deepvann and Anquilla reports).

6. Target tissues

The majority of studies to date have used haemolymph and gill cells of molluscs and peripheral blood cells of fish for the MN analysis (Bolognesi, Hayashi, 2011). There are other studies (albeit limited) available describing the use of blood cells of fish in other tissues, such as liver, kidney and gills (Baršienė *et al.*, 2006a; Rybakovas *et al.*, 2009), and also other cells (fin cells) (Archipchuk, Garanko, 2005) The application of the MN assay to blood samples of fish is particularly attractive as the method is non-destructive, easy to undertake and results in an easy quantifiable number of cells present on the blood smears for microscopic analysis. However, studies must be undertaken to assess the suitability of any species or cell type analysed. For example it is known that Atlantic cod have very low levels of MN in blood erythrocytes in specimens from reference sites, or control groups in laboratory exposures to crude oil. Furthermore it has been shown that MN induction in cod blood erythrocytes and erythrocytes from different haemopoetic tissues (liver, kidney, gill and spleen) differ significantly after three weeks exposure to Stafjord B crude oil. In multiple laboratory exposures (108 exposure groups of cod), developing liver and kidney erythrocytes were proved to be the most sensitive endpoint and most suitable approach for the assessment of oil pollution in the northern Atlantic and North Sea (Baršienė *et al.*, 2005b, 2006a). Liver as a target organ can also be used in *in situ* exposures with turbot and halibut (caging or laboratory) (Baršienė, IRIS reports on BioSea, PROOF, WCM projects).

7. Sample and cell scoring size

The detected MN frequency in fish erythrocytes is approximately 6–10 times lower than in mussels and clams. The large inter-individual variability associated to the low baseline frequency for this biomarker confirming the need for the scoring of a consistent number of cells in an adequate number of animals for each study point. Sampling size in most of studies conducted with mollusc species have been scoring 1000–2000 cells per animal (Izquierdo *et al.*, 2003; Hagger *et al.*, 2005; Bolognesi *et al.*, 1996, 2004, 2006a; Magni *et al.*, 2006; Baršienė *et al.*, 2006a, 2006b, 2008b, 2010a, 2010b; Kopecka *et al.*, 2006; Nigro *et al.*, 2006; Schiedek *et al.*, 2006; Francioni *et al.*, 2007; Siu *et al.*, 2008; Koukouzika and Dimitriadis, 2005, 2008) and previous reviews have suggested that when using fish erythrocytes at least 2000–4000 cells should be scored per animal (Udroiu *et al*., 2006; Bolognesi *et al*., 2006). Previously scorings of 5000–10 000 fish erythrocytes where used for a MN analysis (Baršienė *et al*., 2004). Since 2009–2010, the frequency of MN in fish from the North and Baltic seas was mostly scored in 4000 cells. In stressful heavily polluted zones, the scoring of 5000–10 000 cells in fish is still recommended.

Mussel sampling size in MN assays range from 5 to 20 mussels per site as reported in the literature (Venier and Zampieron, 1997; Bolognesi *et al*., 2004; Baršienė *et al*., 2004, 2006e, 2008a, 2008b; Francioni *et al*., 2007; Siu *et al*., 2008). Evidence suggests that a sample size of ten specimens per site is enough for the assessment of environmental genotoxicity levels and evaluation of the existence of genetic risk zones. In heavily polluted sites, MN analysis in 15–20 specimens is recommended, due to higher individual variation of the MN frequency. MN analysis in more than 20 mussel or fish specimens shows only a minor change of the MN means (Figure 1 in Fang *et al*., 2009; Baršienė *et al*., unpublished results).

8. MN identification criteria

Most of the studies have been performed using diagnostic criteria for micronuclei identification developed by several authors (Heddle *et al*., 1973, 1991; Carrasco *et al*., 1990; Al-Sabti and Metcalfe, 1995; Fenech, 2000; Fenech *et al*., 2003):

- The size of MN is smaller than 1/3 of the main nucleus.
- Micronuclei are round- or ovoid-shaped, non-refractive chromatin bodies located in the cytoplasm of the cell and can therefore be distinguished from artefacts such as staining particles.
- Micronuclei are not connected to the main nuclei and the micronuclear boundary should be distinguishable from the nuclear boundary.

After sampling and cell smears preparation, slides should be coded. To minimize technical variation, the blind scoring of micronuclei should be performed without knowledge of the origin of the samples. Only cells with intact cellular and nuclear membrane can be scored. Particles with colour intensity higher than that of the main nuclei were not counted as MN. The area to be scored should first be examined under low magnification to select the part of the slide showing the highest quality (good staining, non overlapping cells). Scoring of micronuclei should then be undertaken at 1000x magnification.

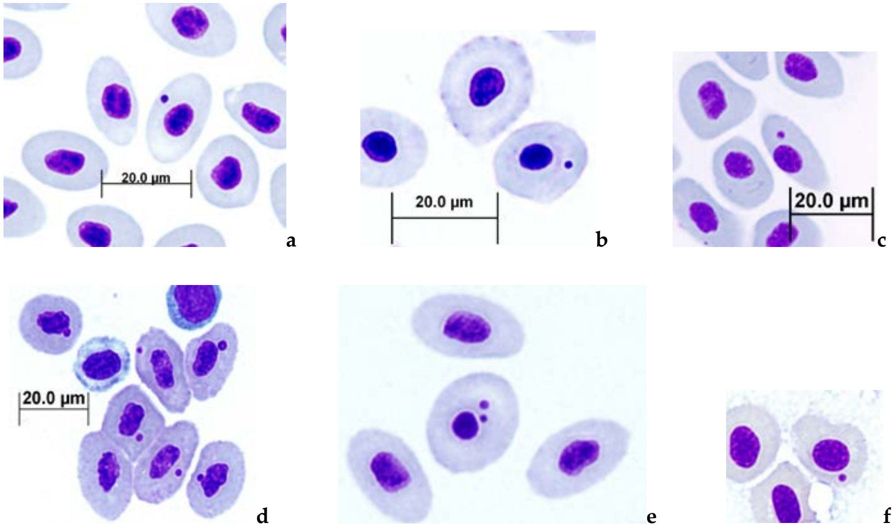


Image A. Micronuclei in blood erythrocytes of *Platichthys flesus* (a), *Limanda limanda* (b), *Zoarces viviparous* (c), *Clupea harrengus* (d), two MN in *Limanda limanda* (e) and MN liver erythrocytes of *Gadus morhua* (f). Images from NRC database.

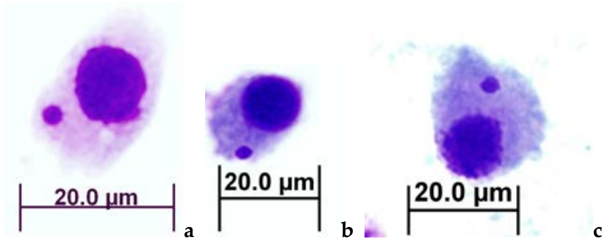


Image B. Micronuclei in gill cells of *Mytilus edulis* (a), *Macoma baltica* (b) and in haemocyte of *Chlamys islandica* (c). Images from NRC database.

Confounding factors

9. Earlier studies on MN formation in mussels have disclosed a significant influence of environmental and physiological factors (Dixon *et al*., 2002). Therefore, the role of the confounding factors should be considered prior to the application of MN assay in biomonitoring programmes, as well as in description of genetic risk zones, or ecosystem health assessments.

Water temperature

MN induction is a cell cycle-related process and depends on water temperature, which is a confounding factor for the mitotic activity in poikilotherm animals. Several studies have demonstrated that baseline frequencies of MN in mussels are related to water temperature (Brunetti *et al*., 1988, 1992; Kopecka *et al*., 2006). Baseline frequencies of MN are regarded as the incidence of MN observed in the absence of environmental risk or before exposure to genotoxins (Fenech, 1993). In fish MN frequencies showed also seasonal differences in relation to water temperature with lower MN levels in winter than in autumn (Rybakovas *et al*., 2009). This was assumed to be an effect of higher mitotic activity and MN formation due to high water temperatures in autumn (Brunetti *et al*., 1988). Additionally, it has been reported that increases in water temperature (4–37ºC) can increase the ability of genotoxic compounds to damage DNA (Buschini *et al*., 2003).

Types of cells

MN may be seen in any type of cell, both somatic and germinal and thus the micronucleus test can be carried out in any active tissue. Nevertheless there are some limitations using different types of cells, for example, agranular and granular haemocytes in mussels. There are also differences between MN induction level in mussel haemolymph and gill cells, mainly because gills are primary targets for the action of contaminants. The anatomical architecture of the spleen in fish does not allow erythrocytes removal in the spleen (Udroiu *et al*., 2006) like mammals do.

Salinity

The influence of salinity on the formation of MN was observed in mussels from the Danish coast located in the transitional zone between the Baltic and North Sea. No relationship between salinity and MN frequencies in mussels could be found for mussels from the North Sea (Karmsund zone), Wismar Bay and Lithuanian coast. Similar results were found for *Macoma balthica* from the Baltic Sea – from Gulfs of Bothnia, Finland, Riga and Lithuanian EZ (Baršienė *et al*., unpublished data).

Size

Because the linear regression analysis of animal's length and induction of MN shows that the size could be a confounding factor, sampling of organisms with similar sizes should take place (Baršienė *et al*., unpublished data). It should also be noted that size is not always indicative of age and therefore age could also potentially affect the response of genotoxicity in the fish.

Diet

Results have shown that MN formation was not influenced in mussels who were maintained under simple laboratory conditions without feeding (Baršienė *et al*., 2006e).

Ecological relevance

10. Markers of genotoxic effects reflect damage to genetic material of organisms and thus get a lot of attention (Moore *et al*., 2004). Different methods have been developed for the detection of both double- and single-strand breaks of DNA, DNA-adducts, micronuclei formation and chromosome aberrations. The assessment of chemical induced genetic damage has been widely utilized to predict the genotoxic, mutagenic and carcinogenic potency of a range of substances, however these investigations have mainly been restricted to humans or mammals (Siu *et al*., 2004). Micronucleus formation indicates chromosomal breaks, known to result in teratogenesis (effects on offspring) in mammals. There is however limited knowledge of relationships between micronucleus formation and effects on offspring in aquatic organisms. With a growing concern over the presence of genotoxins in the aquatic media, the application of cytogenetic assays on ecologically relevant species offers the chance to perform early tests on health in relation to exposure to contaminants.

Applicability across the OSPAR maritime area

Large-scale and long-term studies took place from 2001 to 2010 at the Nature Research Center (NRC, Lithuania) on micronuclei (MN) and other abnormal nuclear formations in different fish and bivalve species inhabiting various sites of the North Sea, Baltic Sea, Atlantic Ocean and Barents Sea. These studies revealed the relevance of environmental genotoxicity levels in ecosystem assessments. Nature Research Center established a large database on MN and other nuclear abnormalities in 13 fish species from the North Sea, Barents Seas and Atlantic Ocean, in eight fish species and in mussels, scallops and clams *Macoma balthica* from the Baltic Sea. Fish and bivalve species were collected from 85 sites in the North Sea and Atlantic and from 117 coastal and offshore sites in the Baltic (Figures 1 and 2). Monitoring of MN and other nuclear abnormalities levels was performed (2–8 times) in many sites of the North and Baltic Seas. Data on MN levels in organisms inhabiting deep-sea and arctic zones are also available (Table 1).

The validation of the MN assay was done with indigenous and cultured mussels *M. edulis,* Atlantic cod, turbot, halibut and long rough dab in multiple laboratory exposures to crude oil from the North Sea and Barents Sea, to produced water discharged from the oil platforms and to other contaminants. Additional active monitoring using mussels and Atlantic cod took place in the Ekofisk, Statfjord, Troll oil platform, oil refinery zones, some northern Atlantic sites as well as in sites heavily polluted by copper or PAHs.

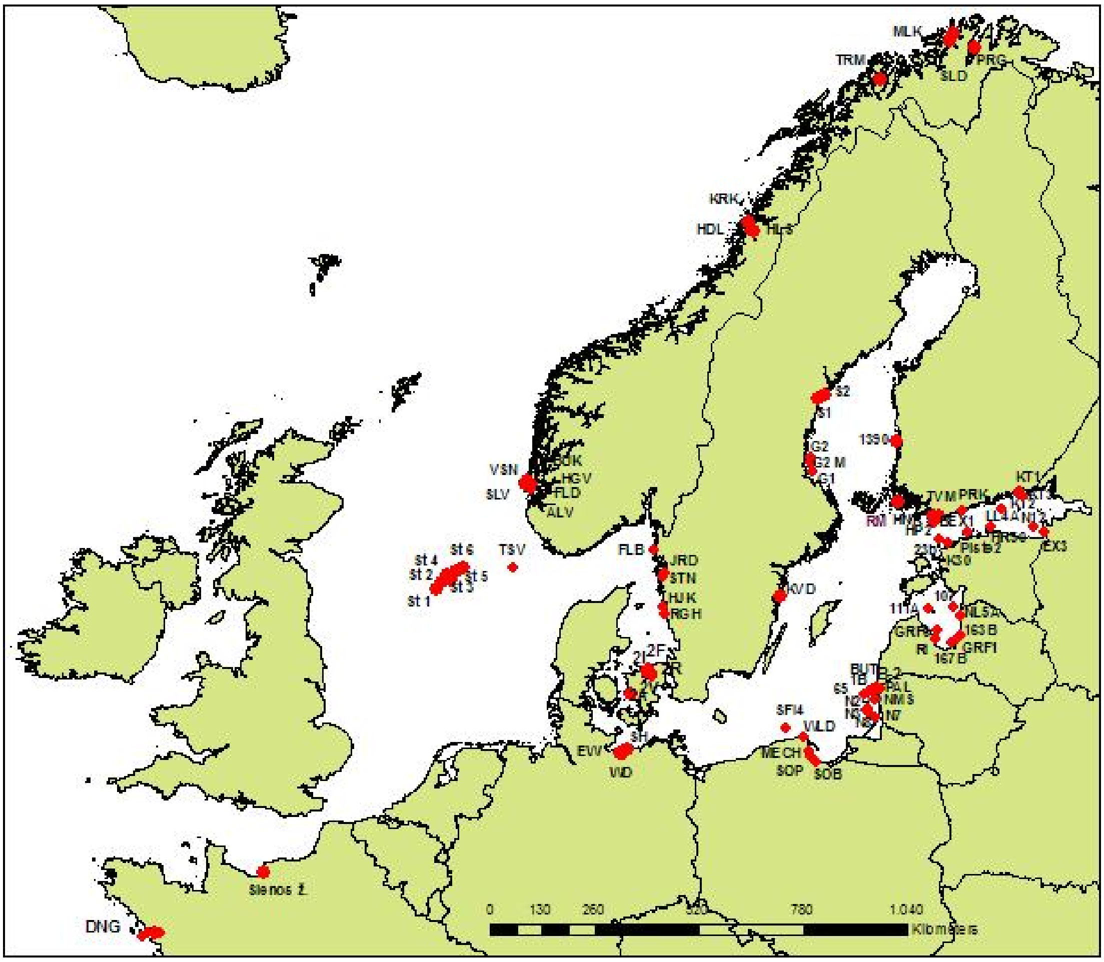


Figure 1. Sampling stations of bivalve molluscs for the micronuclei studies (NRC, Lithuania).

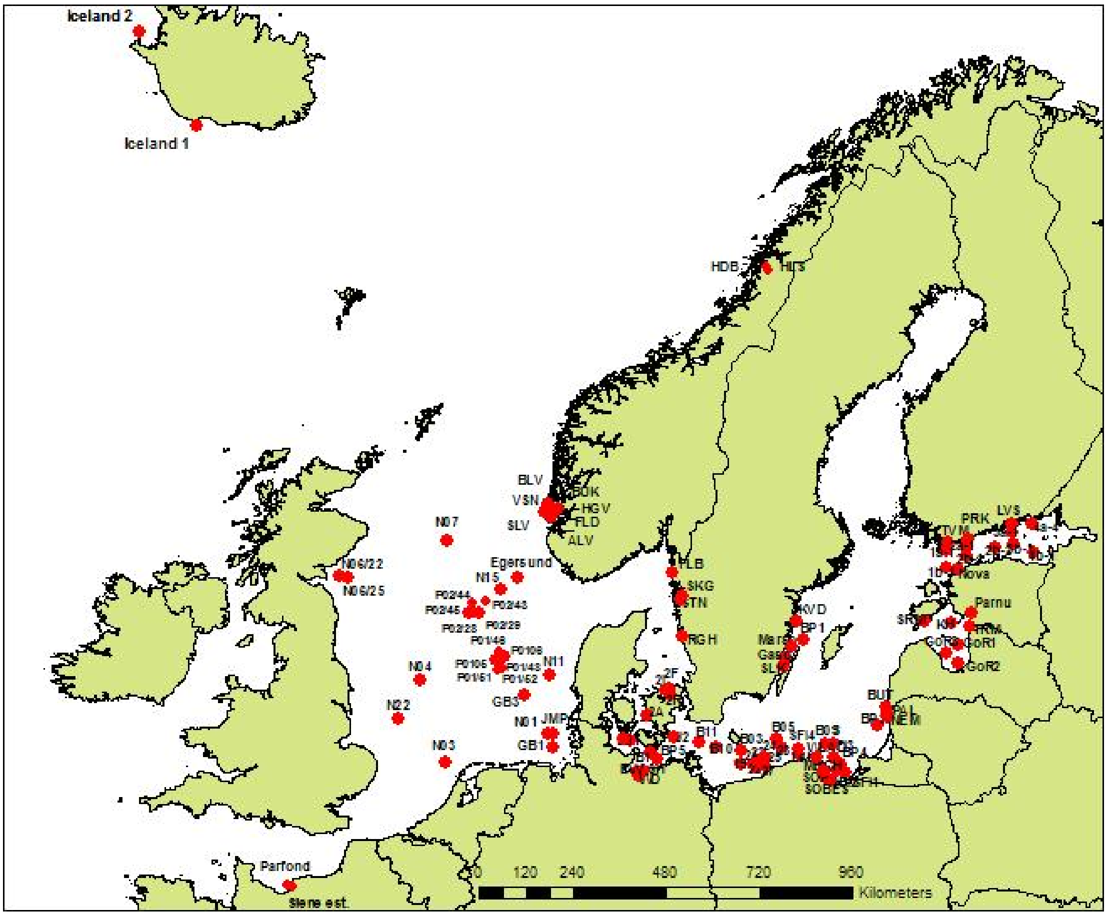


Figure 2. Sampling stations of fish species used for the micronuclei studies (NRC, Lithuania).

Background responses

Baseline or background frequency of MN can be defined as incidence of MN observed in the absence of environmental risk or before exposure to genotoxins (Fenech, 1993). As mentioned above, several studies have demonstrated that MN baseline frequencies depend on water temperature. In fish, MN frequencies lower than 0.05‰ (the Baltic Sea) and lower than 0.1‰ (the North Sea) has been suggested by Rybakovas *et al*. (2009) as a reference level in the peripheral blood erythrocytes of the flatfish flounder (*Platichthys flesus*) and dab (*Limanda limanda*) and also cod (*Gadus morhua*)- after analysing fishes from 12 offshore sites in the Baltic Sea (479 specimens) and 11 sites in the North Sea (291 specimens). For unpolluted sites in the Mediterranean Sea, baseline MN levels in gills of *M. galloprovincialis* have been set depending on water temperature to 1% at temperatures below 15ºC, 2% between 15 and 20ºC, and 3% above 20ºC (Brunetti *et al*., 1992).

The frequencies of micronuclei in marine species sampled from field reference sites are summarized in Table 1. Additionally, the frequencies of MN in blood erythrocytes of fish and in gill cells of mussels deployed to the uncontaminated sites are shown. (Table 2).

Table 1. The reference levels of micronuclei (MN/1000 cells) in European marine species *in situ*.

Species	Tissue	Location	Response MN/1000 cells	Reference
Mytilus galloprovincialis	Gills	Adriatic and Tyrrhenian Sea	1.0-at 15°C 2.0-at 15–20°C 3.0-at above 20°C	Brunetti *et al*., 1992
M. galloprovincialis	Haemolymph	Mediterranean coast	4.2± 0.7	Burgeot *et al*., 1996
M. galloprovincialis	Gills	La Spiezia Gulf, Ligurian Sea	3.0± 2.0	Bolognesi *et al*., 1996
M. galloprovincialis	Gills, Haemolymph	Venice Lagoon	0,73–1,42	Dolcetti and Venier, 2002
M. galloprovincialis	Haemolymph	Strymonikos gulf, Mediterranean Sea	0.30; 1.30	Dailianis *et al*., 2003
M. edulis	Gills	Gijon coast, Spain	1.42	Izquierdo *et al*., 2003
M. galloprovincialis	Gills	Strymonikos gulf, Mediterranean Sea	1.30	Dailianis *et al*., 2003
M. galloprovincialis	Haemolymph	Venice lagoon	0.44	Pampanin *et al*., 2005
M. galloprovincialis	Gills	Tyrrhenian Sea	5.4	Nigro *et al*., 2006
M. galloprovincialis	Gills	Gulf of Oristano, Mediterranean Sea	2.94–4.70	Magni *et al* 2006
M. galloprovincialis	Haemolymph	Adriatic Sea	1.0–1.5	Klobučar *et al*., 2008
M. galloprovincialis	Haemolymph	Adriatic Sea	1.38–1.75	Pavlica *et al*., 2008
M. galloprovincialis	Gills	Gulf of Patras	≈ 2.0	Pytharopoulou *et al*., 2008
M. galloprovincialis	Gills	Algerian coast	0.0–1.18	Taleb *et al*., 2009
M.galloprovincialis	Haemolymph	Algerian coast	1.6–2.47	Taleb *et al*., 2009
M galloprovincialis	Gills	Western Mediterranean	1.9–2.1	Fernandez *et al*., 2011
M. edulis	Haemolymph	Langesundfjord (Norway, rock)	0.90	Wrisberg *et al*., 1992
M. edulis	Haemolymph	Store Belt (Denmark)	0.89	Wrisberg *et al*., 1992
M. edulis	Gills	North Sea (Norwegian coast and Karmsund fjords)	1.05± 0.32	Baršienė *et al*., 2004
M. edulis	Gills	North Sea (Goteborg coast)	0.71± 0.12	Baršienė *et al*., 2008a
M. edulis	Haemolymph	North Sea	1.24± 0.37	Brooks *et al*., 2011
M. edulis	Gills	Baltic Sea	0.37± 0.09	Baršienė *et al*., 2006b
M. trossulus	Gills	Baltic Sea	2.07± 0.32	Baršienė *et al*., 2006b; Kopecka *et al*., 2006
Macoma baltica	Gills	Baltic Sea	0.53–1.28	Baršienė *et al*., 2008b, unpublished data

Species	Tissue	Location	Response MN/1000 cells	Reference
.M baltica	Gills	Stockholm archipelago	0.4	Smolarz, Berger, 2009
Limanda limanda	Blood, kidney erythrocytes	North Sea	0.02± 0.01	Rybakovas *et al*., 2009
Platychthys flesus	Blood erythrocytes	Atlantic Ocean	0.06± 0.04	Baršienė *et al*., unpublished data
P. flesus	Blood erythrocytes	North Sea	0.04± 0.03	Baršienė *et al*., 2008a;
P. flesus	Blood erythrocytes	Baltic Sea	0.15± 0.03	Baršienė *et al*., 2004
P. flesus	Blood erythrocytes	Baltic Sea	0.0± 0.0	Kohler, Ellesat, 2008
P. flesus	Blood erythrocytes	Baltic Sea	0.08± 0.02	Napierska *et al*., 2009
P. flesus	Blood erythrocytes	UK estuaries	0.27–0.66	Lyons, unpublished
Zoarces viviparus	Blood erythrocytes	Baltic Sea	0.02± 0.02	Baršienė *et al*., unpublished data
Gadus morhua	Blood, kidney erythrocytes	North Sea	0.03± 0.02	Rybakovas *et al*., 2009
G. morhua	Blood, kidney erythrocytes	Baltic Sea	0.03± 0.02	Rybakovas *et al*., 2009
Clupea harrengus	Blood erythrocytes	Baltic Sea	0.03± 0.03	Baršienė *et al*., unpublished data
Symphodus melops	Blood erythrocytes	North Sea	0.08± 0.04	Baršienė *et al*., 2004
Scophthalmus maximus	Blood erythrocytes	Baltic Sea	0.10± 0.04	Baršienė *et al*., unpublished data
Perca fluviatilis	Blood erythrocytes	Baltic Sea	0.06± 0.02	Baršienė *et al*., 2005a; Baršienė *et al*., unpublished data
Mugil cephalus	Blood erythrocytes	Mediterranean coast, Turkey	0.82–2.07	Cavas, Ergene-Gozukara, 2005
M. cephalus	Gill cells	Mediterranean coast, Turkey	1.84–2.91	Cavas, Ergene-Gozukara, 2005
Mullus barbatus	Blood erythrocytes	La Spiezia Gulf (Italy)	0.33 [a]	Bolognesi, 2006a
Dicentrarchus labrax	Blood erythrocytes	La Spiezia Gulf (Italy)	0.75 [a]	Bolognesi, 2006a
Pagellus mormyrus	Blood erythrocytes	La Spiezia Gulf (Italy)	0.4 [a]	Bolognesi, 2006a
Sargus sargus	Blood erythrocytes	La Spiezia Gulf (Italy)	0.25 [a]	Bolognesi, 2006a
Seriola dumerili	Blood erythrocytes	La Spiezia Gulf (Italy)	0.38 [a]	Bolognesi, 2006a
Serranus cabrilla	Blood erythrocytes	La Spiezia Gulf (Italy)	0.0 [a]	Bolognesi, 2006a

Species	Tissue	Location	Response MN/1000 cells	Reference
Sparus auratus	Blood erythrocytes	La Spiezia Gulf (Italy)	0.12 [a]	Bolognesi, 2006a
Sphyraena sphyraena	Blood erythrocytes	La Spiezia Gulf (Italy)	0.25 [a]	Bolognesi, 2006a
Trachurus trachurus	Blood erythrocytes	La Spiezia Gulf (Italy)	0.25 [a]	Bolognesi, 2006a
Mugil cephalus	Blood erythrocytes	Mediterranean Goksu Delte, Turkey	1.26± 0.40	Ergene *et al*., 2007
Mullus barbatus	Blood erythrocytes	Western Mediterranean-Spain	0.10–0.16	Martínez-Gómez, 2010.
Dicentrarchus labrax	Blood erythrocytes	Eastern Adriatic Sea	1.25 ± 1.97	Strunjak-Perovic *et al*., 2009

[a] – number of micronuclei per 1000 studied erythrocytes.

Note: It is important to ensure that the data are normally distributed (e.g. Kolmogorov–Smirnov test) if the standard deviation is to be used to calculate MN frequency percentiles of the distribution, as this assumes that the data are normally distributed, which may not be the case.

Table 2. The reference levels of micronuclei (MN/1000 cells) in European marine organisms after caging in uncontaminated/reference sites *in situ*.

Species	Tissue	Location/exposure time	Response MN/1000 cells	Reference
Mytilus galloprovincialis	Gills	Ligurian coast/ 30 days	1.78± 1.04 [a]	Bolognesi *et al*., 2004
M. galloprovincialis	Gills	Gulf of Patras/1 month	2.3–2.5	Kalpaxis *et al*., 2004
M. galloprovincialis	Gills	Haven oil spill area/ 30 days	3.7± 1.62 [a]	Bolognesi *et al*., 2006b
M. galloprovincialis	Gills	Cecina estuary/ 4 weeks	5.4	Nigro *et al*., 2006
M. galloprovincialis	Haemolymph	Adriatic Sea/1 month	1.0	Gorbi *et al*., 2008
M. galloprovincialis	Haemolymph	Tyrrhenian coast/ 1 month	0.27	Bocchetti *et al*., 2008
M. galloprovincialis	Haemolymph	Algerian coast/ 1 month	1.6–2.47	Taleb *et al*., 2009
M. galloprovincialis	Gills	Algerian coast/ 1 month	0.0–1.18	Taleb *et al*., 2009
Mytilus edulis	Gills	Visnes copper site (Norway)/3 weeks	1.87± 0.43	Baršienė *et al*., 2006d
M. edulis	Gills	Karmsund (Norway)/4 weeks	1.40± 0.29	Baršienė *et al*., unpublished data
M. edulis	Heamolymph	North Sea, oil platforms (Norway)/6 weeks	2.13± 0.48	Hylland *et al*., 2008
M. edulis	Heamolymph	Seiland site (Norway)/5.5 months	2.60± 0.21	Baršienė *et al*., unpublished data
M. edulis	Heamolymph	Ekofisk oil platform, North Sea/6 week	1.24± 0.37 (2006) 3.34± 0.28 (2008) 2.78± 0.50 (2009)	Brooks *et al*., 2011
M. edulis	Heamolymph	Oil refinery (France, 2004)/	3.20± 0.36	Baršienė *et al*., unpublished data
M. edulis	Heamolymph	Oil refinery (France, 2006)/	2.34± 0.37	Baršienė *et al*., unpublished data
M. edulis	Heamolymph	Oil refinery (Mongstad, 2007)/100 days	2.90± 0.40	Baršienė *et al*., unpublished data
M. edulis	Heamolymph	Sea Empress clean reference area (90 days)	0.75 ± 0.46	Lyons *et al*., 1998
M. edulis	Heamolymph	Sea Empress clean reference area (110 days)	0.81±0.36	Lyons *et al*., 1998
Crassostrea gigas		Haven oil spill area/30 days	1.49± 0.79 [a]	Bolognesi *et al*., 2006b

Species	Tissue	Location/exposure time	Response MN/1000 cells	Reference
Gadus morhua	Liver erythrocytes	North Sea, oil platforms (Norway)/5 weeks	0.12± 0.05	Hylland *et al*., 2008
G. morhua	Liver erythrocytes	North Sea, oil platforms (Norway)/6 weeks	0.27± 0.13	Baršienė *et al*., unpublished data
Boops boops		Haven oil spill area/ 30 days	0.6± 0.7 [a]	Bolognesi *et al*., 2006b
Mulus barbatus		Haven oil spill area/ 30 days	0.7± 0.6 [a]	Bolognesi *et al*., 2006b
Uranoscopus scaber		Haven oil spill area/ 30 days	1.1± 0.5 [a]	Bolognesi *et al*., 2006b

[a] – number of micronuclei per 1000 studied cells

Additionally, the range of variation of the frequency of MN in blood erythrocytes of fish and gill cells of *M. galloprovincialis* is displayed in Table 3.

Table 3. The range of MN frequency fish (blood, liver, kidney erythrocytes), in mussels, clams, scallops, amphipods (haemolymph, gill and mantle cells) from different sites of the Atlantic Ocean, North Sea, Baltic Sea and Mediterranean Sea.

Species	Number of sites studied	Tissue	MN frequency range, ‰	Reference
Mytilus edulis	3	Haemolymph	0.89–2.87	Wrisberg *et al.*, 1992
M. edulis	2	Haemolymph	0.90–2.32	Wrisberg *et al.*, 1992
M. edulis	3	Mantle	≈ 3-7[a]	Bresler *et al.*, 1999
M. edulis	60	Gills, haemolymph	0.37–7.20	Baršienė *et al.*, 2004, 2006b, 2006c, 2008b, 2010a; Schiedek *et al.*, 2006
Mytilus trossulus	5	Gills	2.07–6.70	Baršienė *et al.*, 2006b, Kopecka *et al.*, 2006
M. galloprovincialis	13	Gills	1.8–24	Brunetti *et al.*, 1988; Scarpato *et al.*, 1990; Bolognesi *et al.*, 2004; Nigro *et al.*, 2006
M. galloprovincialis	3	Gills	2–12	Kalpaxis *et al.*, 2004
M. galloprovincialis	5	Haemolymph	1.38–6.50	Pavlica *et al.*, 2008
M. galloprovincialis	3	Gills	1.2–11.8	Taleb *et al.*, 2009
M. galloprovincialis		Gills	0–22	Fernandez *et al.*, 2011
Macoma balthica	29	Gills	0,53–11.23	Baršienė *et al.*, 2008b; Baršienė *et al.*, unpublished data
Chlamys islandica	3	Haemolymph	3.50 to 5.83	Baršienė *et al.*, unpublished data
Eurythenes gryllus	2	Haemolymph	0.35–0.52	Baršienė *et al.*, unpublished data
Limanda limanda	3	Blood	≈ 2–5b	Bresler *et al.*, 1999
L. limanda	26	Blood, kidney	0.02–1.22	Rybakovas *et al.*, 2009; Baršienė *et al.*, unpublished data
Platychthys flesus	3	Blood	≈ 2–6 [b]	Bresler *et al.*, 1999
P. flesus	53	Blood, kidney	0.08–1.45	Baršienė *et al.*, 2004, 2005a, 2008a; Napierska *et al.*, 2009; Baršienė *et al.*, unpublished data
Zoarces viviparus	40	Blood	0.02–0.81	Baršienė *et al.*, 2005a, Baršienė *et al.*, unpublished data
Gadus morhua	19	Liver, blood,	0.0–0.64	Rybakovas *et al.*, 2009; Baršienė *et al.*, 2010a
Symphodus melops	9	Blood	0.07–0.65	Baršienė *et al.*, 2004, 2008a
Clupea harrengus	32	Blood	0.03–0.92	Baršienė *et al.*, unpublished data
Melanogrammus aeglefinus	3	Liver	0.06–0.75	Baršienė *et al.*, unpublished data
Scophthalmus maximus	4	Blood, liver, kidney	0.10–0.93	Baršienė *et al.*, unpublished data
Perca fluviatilis	14	Blood	0.06–1.15	Baršienė *et al.*, 2005a; Baršienė *et al.*, unpublished data

Species	Number of sites studied	Tissue	MN frequency range, ‰	Reference
Brachydetrius aurestus	3	Liver	0.28–0.85	Baršienė *et al*., unpublished data
Synoglossus senegalensis	2	Liver	0.33–0.45	Baršienė *et al*., unpublished data
Cynoponticus ferox	2	Liver	0.13–0.96	Baršienė *et al*., unpublished data
Rhinobatos irvinei	1	Liver	0,50	Baršienė *et al*., unpublished data
Omogadus argentus	2	Liver	0.23–0.47	Baršienė *et al*., unpublished data

[a] – **Frequency of micronuclei in cells**

[b] – **Frequency of micronuclei in erythrocytes**

Assessment criteria

Assessment Criteria (AC) have been established by using data available from studies of molluscs and fish in the North Sea, northern Atlantic (NRC database) and Mediterranean area (Table 4). The background/threshold level of micronuclei incidences is calculated as the empirical 90% percentile (P90). Until more data becomes available, values should be interpreted from existing national datasets. It should be noted that these values are provisional and require further validation when data becomes available from the ICES database.

The 90% percentile (P90) separates the upper 10% of all values in the group from the lower 90%. The rationale for this decision was that elevated MN frequency would lie above the P90 percentile, whereas the majority of values below P90 belong to unexposed, weakly medium exposed or non-responding adapted individuals. P90 values were calculated for those stations/areas which were considered 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).

ACs in bivalves *Mytilus edulis*, *Mytilus trossulus*, *Macoma balthica* and *Chlamys islandica* (data from MN analysis in 4371 specimens), in fish *Limanda limanda*, *Zoarces viviparus*, *Platichthys flesus*, *Symphodus melops*, *Gadus morhua*, *Clupea harrengus* and *Melogrammus aeglefinus* (data from MN analysis in 4659 specimens) from the North Sea, Baltic Sea and northern Atlantic have been calculated using NRC (Lithuania) databases using data from five or more reference locations (Table 1).

ACs for mussel *Mytilus galloprovincialis* and fish red mullet (*Mullus barbatus*) have been estimated using available data from the Spanish Institute of Oceanography (IEO, Spain). This dataset was obtained using *M. galloprovincialis* from reference stations along the northern Iberian shelf in spring 2003 namely Cadaqués and Medas Islands. In the case of red mullet, background values were derived from the results obtained in Almeria and Málaga areas (SE Spain). Because significant sexual differences were not observed in red mullet, data of both genders were considered.

Table 4. Assessment criteria of MN frequency levels in different bivalve mollusc and fish species. BR =Background response; ER = Elevated response; N = number of specimens analysed.

Species	Size (cm)	Temperature (°C)	Regional Area	Tissue	BR	ER	N
Mytilus edulis	3–4	11–17	Atlantic-North Sea	Haemolymph, gills	<2.51	>2.51	1280
M. edulis	1.5–3	8–18	Baltic Sea	Gills	<2.50	>2.50	1810
M. edulis caged for 4–6 weeks	3–4	7–9	North Sea	Haemolymph	<4.1	> 4.1	44
M. edulis caged for 4–6 weeks	3–4	9–16	North Sea	Haemolymph	<4.06	> 4.06	656
M. trossulus	2–3	3–15	Baltic Sea	Gills	<4.50	> 4.50	230
Macoma balthica	1–3	13–18	Baltic Sea	Gills	<2.90	> 2.90	330
M. galloprovincialis	3–4	13	Western Mediterranean	Gills	<3.87	>3.87	12
Chlamys islandica	4–5	2–4	North Sea	Haemolymph	<4.5	> 4.5	65
Zoarces viviparous	17–30	15–17	North Sea	Erythrocytes	<0.28	>0.28	226
Zoarces viviparous	15–32	7–17	Baltic Sea	Erythrocytes	<0.38	>0.38	824
Limanda limanda	19–24	8–17	North Sea	Erythrocytes	<0.52	>0.52	544
Limanda limanda	18–25	8–17	Baltic Sea	Erythrocytes	<0.49	>0.49	117
Platichtys flesus	20–28	15–17	Atlantic-North Sea	Erythrocytes	<0.33	>0.33	62
Platichtys flesus	17–39	10–17	Baltic Sea coastal	Erythrocytes	<0.29	>0.29	828
Platichtys flesus	18–40	6–18	Baltic Sea offshore	Erythrocytes	<0.23	>0.23	970
Symphodus melops	12–21	13–15	Atlantic-North Sea	Erythrocytes	<0.36	>0.36	158
Gadus morhua	20–48	13–15	Atlantic-North Sea	Erythrocytes	<0.38	>0.38	340
Gadus morhua	20–48	13–15	Baltic Sea	Erythrocytes	<0.38	>0.38	50
Clupea harengus	19–25	5–10	Atlantic-North Sea	Erythrocytes	<0.32	>0.32	60
Clupea harengus	16–29	6–18	Baltic Sea	Erythrocytes	<0.39	>0.39	450
Melogrammus aeglefinus	27–44	8–14	North Sea	Erythrocytes	<0.30	>0.30	30
Mullus barbatus	12–18	17	Western Mediterranean	Erythrocytes	<0.32	>0.32	64

Quality Assurance

The micronucleus test showed to be a useful in vivo assay for genotoxicity testing. However, many aspects of its protocol need to be refined, knowledge of confounding factors should be improved and interspecies differences need further investigation. In 2009 an inter-laboratory comparison exercise was organized within the framework of the MED POL programme using *M. galloprovincialis species.* The results are expected by mid-2011.

Intercalibration of MN analysis in fish was done between experts from NRC and Caspian Akvamiljo laboratories, as well as between NRC experts and the University of Aveiro, Portugal (Santos *et al*., 2010). It is recommended that these relatively simple interlaboratory collaborations are expanded to include material from all the commonly used bio-indicator species in 2011/2012.

Scientific potential

MN analysis in different marine and freshwater species of bivalves and fish is carried out in many laboratories of European countries: Italy, Portugal, Spain, Turkey, Lithuania, UK, Greece, Germany, Poland, Croatia, Estonia, Russia, Norway and Ukraine. There are single laboratories in Hungary, Algeria and Egypt. Highly qualified expert groups work in Italy, Lithuania, Spain, Turkey, Portugal, UK and are able to perform analysis in both groups of animals; both in invertebrates and vertebrates.

References

Al-Sabati K, Metcalfe CD. Fish micronuclei for assessing genotoxicity in water. Mutation Research 1995; 323: 121–35.

Andreikėnaitė L. 2010. Genotoxic and cytotoxic effects of contaminants discharged from the oil platforms in fish and mussels. Summary of doctoral dissertation, 38p., Vilnius.

Arkhipchuk V.V., Garanko N.N. 2005. Using the nucleolar biomarker and the micronucleus test on *in vivo* fish fin cells. Ecotoxicology and Environmental Safety. Vol. 62. P. 42–52.

Bagni G., Baussant T., Jonsson G., Baršienė J., Mascini M. 2005. Electrochemical device for the rapid detection of genotoxic compounds in fish bile samples. Analytical Letters, 38: 2639–2652.

Baršienė J., Rybakovas A. 2006e. Cytogenetic and cytotoxic effects in gill cells of the blue mussel *Mytilus edulis* from the Baltic coast and after 1–3-day maintenance in laboratory. Acta Zoologica Lituanica, 16: 191–197.

Baršienė J., Andreikėnaitė L. 2007. Induction of micronuclei and other nuclear abnormalities in blue mussels exposed to crude oil from the North Sea. Ekologija, 53, No 3: 9–15.

Baršienė J., Andreikėnaitė L., Garnaga G., Rybakovas A. 2008b. Genotoxic and cytotoxic effects in bivalve molluscs *Macoma balthica* and *Mytilus edulis* from the Baltic Sea. Ekologija, 54: 44–50.

Baršienė J., Bjornstad A., Andreikėnaitė L. 2010b. Induction of micronuclei and other nuclear abnormalities in blue mussels *Mytilus edulis* after 1-, 2-, 4- and 8-day treatment with crude oil from the North Sea. Ekologija, 56, 3–4: 116–123.

Baršienė J., Bjornstad A., Rybakovas A., Šyvokienė J., Andreikėnaitė L. 2010a. Environmental genotoxicity and cytotoxicity studies in mussels and fish inhabiting northern Atlantic zones impacted by aluminum industry. Ekologija, 56, No 3–4: 124–131.

Baršienė J., Dedonytė V., Rybakovas A., Andersen O-K. 2005b. Induction of micronuclei in Atlantic cod (*Gadus morhua*) and turbot (*Scophthalmus* maximus) after treatment with bisphenol A, diallyl phthalate and tetrabromodiphenyl ether-47. Ekologija No 4, 1–7.

Baršienė J., Dedonytė V., Rybakovas A., Andreikėnaitė L., Andersen O.K. 2006a. Investigation of micronuclei and other nuclear abnormalities in peripheral blood and kidney of marine fish treated with crude oil. Aquatic Toxicology, 78, Suppl. 1: S99–104.

Baršienė J., Dedonytė V., Rybakovas A., Broeg K., Forlin L. , Gercken J., Kopecka J., Balk L. 2005a. Environmental mutagenesis in different zones of the Baltic Sea. Acta Zoologica Lituanica, 15: 90–95.

Baršienė J., Lazutka J., Šyvokienė J., Dedonytė V., Rybakovas A., Bjornstad A., Andersen O.K. 2004. Analysis of micronuclei in blue mussels and fish from the Baltic and the North Seas. Environmental Toxicology, 19: 365–371.

Baršienė J., Lehtonen K., Koehler A., Broeg K., Vourinen P.J., Lang T., Pempkowiak J., Šyvokienė J., Dedonytė V., Rybakovas A., Repečka R., Vountisjarvi H., Kopecka J. 2006c. Biomarker responses in flounder (*Platichthys flesus*) and mussel (*Mytilus edulis*) in the Klaipėda-Būtingė area (Baltic Sea). Marine Pollution Bulletin, 53: 422–436.

Baršienė J., Rybakovas A., Andreikėnaitė L. 2006d. Cytogenetic damage in marine organisms as a biomarker of environmental pollution. Metals in the Environment, 3rd Intern. Conf. in Lithuania. Vilnius 2006a, 156–157.

Baršienė J., Rybakovas A., Forlin L., Šyvokienė J. 2008a. Environmental genotoxicity studies in mussels and fish from the Goteborg area of the North Sea. Acta Zool. Lituanica, 18: 240–247.

Baršienė J., Schiedek D., Rybakovas A., Šyvokienė J., Kopecka J., Forlin L. 2006b. Cytogenetic and cytotoxic effects in gill cells of the blue mussel *Mytilus spp.* from different zones of the Baltic Sea. Marine Pollution Bulletin, 53: 469–478.

Bocchetti R., Fattorini D., Pisanelli B., Macchia S., Oliviero L., Pilato F., Pellegrini D., Regoli F. 2008. Contaminant accumulation and biomarker responses in caged mussels, *Mytilus galloprovincialis,* to evaluate bioavailability and toxicological effects of remobilized chemicals during dredging and disposal operations in harbour areas. Aquatic Toxicology, 89: 257–266.

Boller K., Schmid W. 1970. Chemische Mutagenese beim Sauger. Das Knochermark des Chinesischen Hamsters als *in vivo*-Test system. Haematologische Befunde nach Behandlung mit Trenimon. Humangenetik, 11, 35–54.

Bolognesi C., Perrone E., Roggieri P., Pampanin D.M., Sciutto A. 2006a. Assessment of micronuclei induction in peripheral erythrocytes of fish exposed to xenobiotics under controlled conditions. Aquatic Toxicology, 78, Suppl. 1: S93–S98.

Bolognesi C., Perrone E., Roggieri P., Sciutto A. 2006b. Bio-indicators in monitoring long-term genotoxic impact of oil spill: Haven case study. Marine Environ. Res., 62: S287–S291.Bolognesi C., Hayashi M. 2011. Micronucleus assay in aquatic animals. Mutagenesis, 26: 205–213.

Bolognesi C. and Hayashi M. 2011. Micronucleus assay in aquatic animals. Mutagenesis, 26: 205–213.

Bolognesi C., Frenzilli G., Lasagna C., Perrone E., Roggieri P. 2004. Genotoxicity biomarkers in *Mytilus galloprovincialis*: wild versus caged mussels. Mutation Research. 552: 153–162.

Bolognesi C., Rabboni R. and Roggieri P. 1996. Genotoxicity biomarkers in *M. Galloprovincialis* as indicators of marine pollutants. Comparative Biochemistry and Physiology 113C (2): 319–323.

Bolognesi C., Rabboni R. and Roggieri P. 1996. Genotoxicity biomarkers in *M. galloprovincialis* as indicators of marine pollutants. Comparative Biochemistry and Physiology, 113C, No 2: 319–323.

Bresler V., Bissinger V., Abelson A., Dizer H., Sturm A., Kratke R., Fischelson L., Hansen P.-D. 1999. Marine molluscs and fish as biomarkers of pollution stress in littoral regions of the Red Sea, Mediterranean Sea and North Sea. Helgoland Marine Research, 53: 219–243.

Brooks S.J., Harman C., Grung M., Farmen E., Ruus A., Vingen S., Godal B.F., Baršienė J., Andreikėnaitė L., Skarphedinsdottir H., Liewenborg B., Sundt R.C. 2011. Water column monitoring of the biological effects of produced water from the Ekofisk offshore oil installation from 2006 to 2009. Journal of Toxicology and Environmental Health, Part A, 74: 582–604.

Brunetti R., Gabriele M., Valerio P. and Fumagalli O. 1992. The micronucleus test: temporal pattern of the base-line frequency in *Mytilus galloprovincialis* Lmk. Marine Ecol. Prog. 83: 75–78.

Brunetti R., Majone F., Gola I and Beltrame C. 1988. The micronucleus test: examples of application to marine ecology. Marine Ecol. Prog., 44:65–68.

Buschini A, Carboni P, Martino A, Poli P, Rossi C. 2003. Effects of temperature on baseline and genotoxicant-induced DNA damage in haemocytes of *Dreissena polymorpha*. Mutatation Research, 537: 81–92.

Burgeot, T., Woll, S. and F. Galgani. 1996. Evaluation of the micronucleus test on *Mytilus galloprovincialis* for monitoring applications along French coasts. Marine Pollution Bulletin, Volume 32, No. 1: 39–46.

Martínez-Gómez, C. Navarrro, B. Fernández, J. Valdés and J. Benedicto. 2010. Evidence of genotoxic effects in red mullet populations from the Mediterranean coast of Spain. ICES Annual Science Conference. Nantes (France), 18–25 September 2010. Proceeding book, ICES CM 2010:/F31.

Carrasco K.R., Tilbury K.L., Myers M.S. 1990. Assessment of the piscine micronucleus test as an *in situ* biological indicator of chemical contaminant effects. Canadian Journal Fishery and Aquatic Sciences, 47: 2123–2136.

Çavaş T., Ergene-Gözükara S. 2005. Micronucleus test in fish cells: a bioassay for *in situ* monitoring of genotoxic pollution in the marine environment. Environmental and Molecular Mutagenesis, 46: 64–70.

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

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

Dixon D.R., Dixon L.R.J., Shillito B., Gwynn J.P. 2002. Background and induced levels of DNA damage in Pacific deep-sea vent palychaetes: the case for avoidance. Cah. Biol. Marine 43: 333–336.

Dolcetti L. and Venier, P. 2002. Susceptibility to genetic damage and cell types in Mediterranean mussels. Marine Environmental Research, 54: 487–491.

Environmental genotoxicity and cytotoxicity studies in mussels and fish inhabiting northern Atlantic zones impacted by aluminum industry. Ekologija, 56, No 3–4: 124–131.

Ergene S., Çavaş T., Celik A., Koleli N., Kaya F., Karahan A. 2007. Monitoring of nuclear abnormalities in peripheral erythrocytes of three fish species from the Goksu Delta (Turkey): genotoxic damage in relation to water pollution. Ecotoxicology, 16: 385–391.

Fang J.K.H., Wu R.S.S., Yip C.K.M., Shin P.K.S. 2009. Power analysis for biomarkers in mussels for use in coastal pollution monitoring. Marine Pollution Bulletin, 58: 1152–1158.

Fenech M. 2000. The in vitro micronucleus technique. Mutation Research 455:81–95.

Fenech, M. 1993. The Cytokinesis-Block Micronucleus Technique: A Detailed Description of the Method and Its Application to Genotoxicity Studies in Human Populations. Mutation Research, 285, 35–44.

Fenech M., Chang W.P., Kirsch-Volders M., Holland N., Bonassi S., Zeiger E. 2003. HUMN project: detailed description of the scoring criteria for the cytokinesis-block micronucleus assay using isolated human lymphocyte cultures. Mutation Research, 534: 65–75.

Fernández B, Campillo JA, Martínez-Gómez C, Benedicto J. 2011. Micronuclei and other nuclear abnormalities in mussels (*Mytilos galloprovincialis*) as biomarkers of cyto-genotoxic pollution in Mediterranean waters. Environmental and Molecular Mutagenesis. *Accepted 25 January 2011.* DOI 10.1002/em.20646.

Francioni E., Wagener A. de L.R., Scofield A. de L., Depledge M.H., Cavalier B., Sette C.B. Carvalhosa L., Lozinsky C., Mariath R. 2007. Polycyclic aromatic hydrocarbon in intertidal mussel *Perna perna*: Space–time observations, source investigation and *genotoxicity.* Science of the Total Environment, Vol. 372 Issue 2/3: 515–531.

Gorbi S., Lamberti C.V., Notti A., Benedetti M., Fattorini D., Moltedo G., Regoli F. 2008. An Ecotoxicological protocol with caged mussels, *Mytilus galloprovincialis*, for monitoring the impact of an offshore platform in the Adriatic Sea. Marine Environmental Research, 65: 34–49.

Hagger J.A., Depledge M.H., Galloway T.S. 2005. Toxicity of tributyltin in the marine mollusc *Mytilus edulis*. Marine Pollution Bulletin, 51:811–816.

Heddle J.A. 1973. A rapid *in vivo* test for chromosomal damage. Mutation Research, 18: 187–190.

Heddle J.A., Cimino M.C., Hayashi M., Romagna F., Shelby M.D., Tucker J.D., Vanparys Ph. and MacGregor J.T. 1991. Micronuclei as an index of cytogenetic damage: past, present, and future. Environmental and Molecular Mutagenesis, 18: 277–291.

Hooftman R.N., Raat W.K. 1982. Induction of nuclear anomalies (micronuclei) in the peripheral blood erythrocytes of an eastern mudminnow *Umbra pygmaea* by ethyl methanesulphonate. Mutation Research, 104: 147–152.

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

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

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

Klobučar G.I.V., Štambuk A., Hylland K., Pavlica M. 2008. Detection of DNA damage in haemocytes of *Mytilus galloprovincialis* in the coastal ecosystems of Kaštela and Trogir bays, Croatia. Sciences of the Total Environment, 405: 330–337.

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

Koukouzika N., Dimitriadis V.K. 2005. Multiple biomarker comparison in *Mytilus galloprovincialis* from the Greece coast: "Lysosomal membrane stability, neutral red retention, micronucleus frequency and stress on stress". Ecotoxicology, 14:449–463.

Koukouzika N., Dimitriadis V.K. 2008. Aspects of the usefulness of five marine pollution biomarkers, with emphasis on MN and lipid content. Marine Pollution Bulletin, 56:941–949.

Lindberg H.K., Wang X., Järventaus H., Falck G.C.-M., Norppa H., Fenech M. 2007. Origin of nuclear buds and micronuclei in normal and folate-deprived human lymphocytes. Mutation Research, 617: 33–45.

Lyons, B.P. 1998. Methods for detecting genotoxic exposure in the aquatic environment. PhD thesis. Centre for Molecular Genetics and Toxicology, University of Wales Swansea.

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

Moore MN, Depledge MH, Readman JW, Leonard P. 2004a. An integrated biomarker-based strategy for ecotoxicological evaluation of risk in environmental management. Mutation Research, 552:247–268.

Napierska D., Baršienė, J., Mulkiewicz E., Podolska M., Rybakovas A. 2009. Biomarker responses in flounder *Platichthys flesus* from the Polish coastal area of the Baltic Sea and application in biomonitoring. Ecotoxicology, 18: 846–859.

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

Pampanin D.M., Marangon I., Volpato E., Campesan G., Nasci C. 2005. Stress biomarkers and alkali-labile phosphates level in mussels (*Mytilus galloprovincialis*) collected in the urban area of Venice (Venice Lagoon, Italy). Environmental Pollution, 136: 103–107.

Pavlica M., Podrug M., Štambuk A., Cvjetko P., Klobučar G.I.V. 2008. Seasonal Variability in Micronuclei Induction in Haemocytes of Mussels along the Eastern Adriatic Coast. Polish Journal of Environmental Studies, 17, No. 5: 765–771.

Pytharopoulou S., Sazaklib E., Grintzalisc K., Georgiouc C.D., Leotsinidis M., Kalpaxis D.L. 2008. Translational responses of *Mytilus galloprovincialis* to environmental pollution: Integrating the responses to oxidative stress and other biomarker responses into a general stress index. Aquatic Toxicology, 89: 18–27.

Rybakovas A., Baršienė J., Lang T. 2009. Environmental genotoxicity and cytotoxicity in the offshore zones of the Baltic and North Seas. Marine Environmental Research, 68: 246–256.

Santos M.M., Sole M., Lima D., Hambach B., Ferreira A.M., Reis-Henriques M.A. 2010. Validating a multi-biomarker approach with the shanny *Lipophrys pholis* to monitor oil spills in European marine ecosystems. Chemosphere, 81, 6, 685–691.

Scarpato R., Migliore L., Barale R. 1990. The micronucleus assay in *Anodonta cygnea* for the detection of drinking water mutagenecity. Mutation Research, 245: 231–237.

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

Schmid W. 1975. The micronucleus test. Mutation Research, 31:9–15.

Serrano-García L., Montero-Montoya R. 2001. Micronuclei and chromatid buds are the result of related genotoxic events. Environmental and Molecular Mutagenenesis 38:38–45.

Siu, W. H. L., J. Cao, R. W. Jack, R. S. S. Wu, B. J. Richardson, L. Xu and P. K. S. Lam. 2004. Application of the comet and micronucleus assays to the detection of B a P genotoxicity in haemocytes of the greenlipped mussel (*Perna viridis*). Aquatic Toxicology, 66: 381–392.

Smolarz K., Berger A. 2009. Long-term toxicity of hexabromocyclododecane (HBCDD) to the benthic clam *Macoma balthica* (L.) from the Baltic Sea. Aquatic Toxicology, 95: 239–247.

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

Taleb Z.M., Benali I., Gherras H., Ykhlef-Allal A., Bachir-Bouiadjra B., Amiard J.C., Boutiba Z. 2009. Biomonitoring of environmental pollution on the Algerian west coast using caged mussels *Mytilus galloprovincialis*. Oceanologia, 51 (1): 63–84.

Udroiu I. 2006. The micronucleus test in piscine erythrocytes. Aquatic Toxicology, 79: 201–204.

Santos M.M., Sole M., Lima D., Hambach B., Ferreira A.M., Reis-Henriques M.A. 2010. Validating a multi-biomarker approach with the shanny *Lipophrys pholis* to monitor oil spills in European marine ecosystems. Chemosphere, 81: 685–691.

Venier P. and Zampieron C. 2005. Evidence of genetic damage in grass gobies and mussels from the Venice lagoon. Environment International, 31: 1053–1064.

Wrisberg M.N., Bilbo C.M., Spliid H. 1992. Induction of micronuclei in hemocytes of *Mytilus edulis* and statistical analysis. Ecotoxicology and Environmental Safety, 23: 191–205.

Annex 10: Background document: Comet assay as a method for assessing DNA damage in aquatic organisms

Contributed by Brett Lyons (UK) with some additions from Steinar Sanni (NO) and Ketil Hylland (NO).

Background

The analysis of modified or damaged DNA has been shown to be a highly suitable method for assessing exposure to genotoxic contaminants in aquatic environments. In general, the methods developed are sensitive to a range of contaminant concentrations, applicable to a wide range of species and have the advantage of detecting and quantifying exposure to genotoxins without a detailed knowledge of the contaminants present. The Single Cell Gel Electrophoresis (SCGE) or comet assay was first applied to ecotoxicology over 15 years ago, and has since become one of the most widely used tests for detecting DNA strand breaks in aquatic animals[1-5]. The comet assay has many advantages over other methods commonly used to assess genotoxic exposure, including (1) genotoxic damage can be detected in most eukaryotic cell types at the single cell level; (2) only a small number of cells are required; (3) it is a rapid and sensitive technique; (3) due to the nature of DNA strand break formation it provides an early warning response of genotoxic exposure; (4) sites of oxidative damage can be identified using enzymatic pretreatment.

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

Confounding factors: Protocols, cell types and target organs

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

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

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

A protocol has recently been developed for conserving fish erythrocytes sampled in the field for subsequent Comet analysis (Hylland *et al*., in prep), which will make the assay more directly applicable for monitoring purposes.

Ecological relevance

Marine invertebrates (bivalves)

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

Mussels have also been used extensively in the field as part of transplantation studies[25-27]. The use of indigenous organisms is often hampered by the absence of a suitable sentinel species, or if present, the genotoxic responses obtained may be influenced by local physiological adaptations. Furthermore the use of transplanted organisms also offers advantages over indigenous species, such as ensuring genetic homogeneity, developmental/reproductive status and controlling the precise exposure window. Validation studies have been under taken with the comet assay to assess the time course variations in DNA damage following field transplantation experiments[25, 26]. It was observed that within the first seven days following transplantation the level of DNA damage can fluctuate, which is likely to be caused by manipulation disturbance, then after two weeks the level reaches a plateau. Such data suggests that transplantation experiments lasting less than two weeks may give spurious results, with the levels of DNA damage detected attributable to artefacts associated with the sampling procedure rather than genotoxic exposure. Studies conducted in a coastal area of Denmark, impacted by a disused chemical site, have also highlighted that the levels of DNA damage in mussels can be affected by seasonal variations in baseline levels[25]. Such results are likely to be influenced by the seasonal variations, which are known to exist for a range of physiological and reproductive processes in mussels[28, 29].

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

Marine vertebrates (fish)

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

Quality assurance

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

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

Members of WGBEC strongly supported the development of an intercalibration exercise for Comet in both blue mussel and fish. Ketil Hylland (NO) will take the initiative to generate samples for such an exercise using both types of cells. Samples will be distributed immersed in lysis buffer. This activity is currently scheduled for 2012.

Background responses and assessment criteria

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

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

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

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

In addition to the above, there was a recent study as part of ICON in which dab (*Limanda limanda*) were collected from the North Sea and in Icelandic waters (Skei *et al*., in prep). Ninety percentiles from the reference location support a value of 4–5% tail DNA as a BAC assessment criterion for this species.

In laboratory experiments with Atlantic cod (*Gadus morhua*) a Comet assay value of 4.9 % tail DNA was measured in the control group (Sanni, unpubl). The water was supplied continuously from a non-polluted source at 78 meters depth of a North Sea coastal location outside Stavanger (Norway). In this experiment, dose dependent increases in Comet values were observed with increasing exposure concentrations of produced water but the range of concentrations in the study were not large enough to be able to establish EAC Comet values corresponding to critical mortality values for larval stages of cod. At the highest exposure in the experiment, the Comet value was 8.4% tail DNA, hence the EAC Comet value (based on toxicity experiments) can be expected found at a higher level than this.

In a similar experiment with blue mussel (*Mytulis edulis*) haemocytes, the Comet level in the controls were 7% tail DNA, while the Comet value corresponding to the exposure level of a dispersed North Sea crude oil critical for mussel larval mortality was 14% tail DNA[48].

From the above, it would appear that a preliminary BAC for Comet analyses of dab and Atlantic cod erythrocytes could be set at 5%. There is not sufficient data to provide an EAC at this time.

For mussel haemocytes, available data suggest a BAC of 10%. One study has been able to determine an effect level that could be used to derive an EAC (14%), but this needs to be supported by further studies.

There is a requirement for a standardized protocol for the main species used in monitoring programmes (dab, flounder, cod, blue mussel), including minimum acceptable reporting criteria (cellular toxicity, +/- control etc) and a decision about reporting format (tail moment, % DNA in tail). There is furthermore a need for QA and intercalibration exercises (will be initiated by WGBEC members) and further evaluation of the suggested assessment criteria.

References

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

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

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

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

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

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

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

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

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

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

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

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

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

[14] N.P. Singh, D.B. Danner, R.R. Tice, M.T. McCoy, G.D. Collins and E.L. Schneider. Abundant alkali-sensitive sites in DNA of human and mouse sperm, *Exp. Cell. Res.*, 1989, 184, 461–470.

[15] G. Frenzilli, V. Scarcelli, F. Taddei and M. Nigro. Adaption of SCGE as a candidate for monitoring marine ecosystems, *Neoplasma*, 1999, 46, 6–7.

[16] M. Moretti, M. Villarini, G. Scassellati-Sforzolini, A.M. Santroni, D. Fedeli and G. Falcioni. Extent of DNA damage in density-separated trout erythrocytes assessed by the 'comet' assay, *Mutat. Res.* 1998, 397, 353–360.

[17] G. Wirzinger, L. Weltje, J. Gercken and H. Sordyl. Genotoxic damage in field-collected three-spined sticklebacks (*Gasterosteus aculeatus* L.): a suitable biomonitoring tool? *Mutat. Res.*, 2007, 628, 19–30.

[18] B.L.Bayne. Watch on mussels. *Mar. Pollut. Bull.*, 1976 7(12), 217–128.

[19] R. Seed. Ecology. In: B.L. Bayne, Marine Mussels, their ecology and physiology, International Biological Programme (13–65), Cambridge University Press, 1976, 13–65.

[20] M.H. Salazar and S.M. Salazar. *In situ* bioassays using transplanted mussels: I. Estimating chemical exposure and bioeffects with bioaccumulation and growth. In Hughes JS, Biddinger GR, Mones E, Eds. Environmental Toxicology and Risk Assessment – Third Volume. Philadelphia: American Society for Testing and Materials STP, 1995 2118, 216–241.

[21] B. Perez-Cadahia, B. Laffon, E. Pasaro and J. Mendez. Evaluation of PAH bioaccumulation and DNA damage in mussels (*Mytilus galloprovincialis*) exposed to spilled Prestige crude oil, *Comp. Biochem. Physiol. C: Toxicol. Pharmacol.*, 2004, 138, 453–460.

[22] B. Laffon, T. Rabade, E. Pasaro and J. Mendez. Monitoring of the impact of Prestige oil spill on *Mytilus galloprovincialis* from Galician coast, *Environ. Int.*,2006 32, 342–348.

[23] I.C. Taban, R.K. Bechmann, S. Torgrimsen, T. Baussant and S. Sanni. Detection of DNA damage in mussels and sea urchins exposed to crude oil using comet assay, *Mar. Environ. Res.*, 2004, 58, 701–705.

[24] J. Rank, K. Jensen and P.H. Jespersen. Monitoring DNA damage in indigenous blue mussels (*Mytilus edulis*) sampled from coastal sites in Denmark, *Mutat. Res./Genet. Toxicol. Environ. Mutagen.*, 2005, 585, 33–42.

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

[26] F. Regoli, G. Frenzilli, R. Bocchetti, F. Annarumma, V. Scarcelli, D. Fattorini and M. Nigro. Time-course variations of oxyradical metabolism, DNA integrity and lysosomal stability in mussels, *Mytilus galloprovincialis*, during a field translocation experiment, *Aquat. Toxicol.*, 2004, 68, 167–178.

[27] M. Nigro, A. Falleni, I. Del Barga, V. Scarcelli, P. Lucchesi, F. Regoli and G. Frenzilli. Cellular biomarkers for monitoring estuarine environments: transplanted versus native mussels, *Aquat. Toxicol.*, 2006, 77, 339–347.

[28] A. Hines, G.S. Oladirana, J.P. Bignell, G.S. Stentiford, M.R. Viant. Direct sampling of organisms from the field and knowledge of their phenotype: key recommendations for environmental metabolomics, *Environ. Sci. Tech.*, 2007, 41, 3375–3381.

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

[30] H.P. Halldorsson, G. Ericson and J. Svavarsson. DNA strand breakage in mussels (*Mytilus edulis* L.) deployed in intertidal and subtidal zone in Reykjavik harbour, *Mar. Environ. Res.*, 2004, 58, 763–767.

[31] S.W. Feist, T. Lang, G.D. Stentiford, A. Köhler, A. Biological effects of contaminants: use of liver pathology of the European flatfish dab (*Limanda limanda* L.) and flounder (*Platichthys flesus* L.) for monitoring. *ICES Techniques in Marine Environmental Sciences*, 2004 38. 42 pp.

[32] JAMP guidelines for general biological effects monitoring. *Joint Assessment and Monitoring Programme*. Oslo and Paris Commissions, 1998, 38 pp.

[33] F. Regoli, G.W. Winston, S. Gorbi, G. Frenzilli, M. Nigro, I. Corsi and S. Focardi. Integrating enzymatic responses to organic chemical exposure with total oxyradical absorbing capacity and DNA damage in the European eel *Anguilla anguilla*: toward development of a more holistic biomarker assessment, *Environ. Toxicol. Chem.*, 2003, 22, 2120–2129.

[34] E. Mamaca, R.K. Bechmann, S. Torgrimsen, E. Aas, A. Bjornstad, T. Baussant and S.L. Floch. The neutral red lysosomal retention assay and Comet assay on haemolymph cells from mussels (*Mytilus edulis*) and fish (*Symphodus melops*) exposed to styrene, *Aquat. Toxicol.*, 2005, 75, 191–201.

[35] M.G.J. Hartl, M. Kilemade, D. Sheehan, C. Mothersill, J. O'Halloran, N.M. O'Brien and F.N.A.M. van Pelt. Hepatic biomarkers of sediment-associated pollution in juvenile turbot, *Scophthalmus maximus* L, *Mar. Environ. Res.*, 2007, 64, 191–208.

[36] G. Frenzilli, V. Scarcelli, I. Del Barga, M. Nigro, L. Forlin, C. Bolognesi and J. Sturve. DNA damage in eelpout (*Zoarces viviparus*) from Goteborg harbour, *Mutat. Res.*, 2004, 552, 187–195.

[37] F. Akcha, G. Leday and A. Pfohl-Leszkowicz. Potential value of the comet assay and DNA adduct measurement in dab (*Limanda limanda*) for assessment of *in situ* exposure to genotoxic compounds, *Mutat. Res.*,2003, 534, 21–32.

[38] F. Akcha, F. Vincent Hubert and A. Pfohl-Leszkowicz. Measurement of DNA adducts and strand breaks in dab (*Limanda limanda*) collected in the field: effects of biotic (age, sex) and abiotic (sampling site and period) factors on the extent of DNA damage, *Mutat. Res.*, 2004, 552, 197–207.

[39] B.P. Lyons, G.D. Stentiford, J. Bignell, F. Goodsir, D.B. Sivyer, M.J. Devlin, D. Lowe, A. Beesley, C.K. Pascoe, M.N. Moore and E. Garnacho. A biological effects monitoring survey of Cardigan Bay using flatfish histopathology, cellular biomarkers and sediment bioassays: findings of the Prince Madog Prize 2003, *Mar. Environ. Res.*, 2006 62, S342–S346.

[40] Jha, A.N., Dogra, Y., Turner, A., Millward, G.E. 2005. Impact of low doses of tritium on the marine mussel, *Mytilus edulis*: Genotoxic effects and tissue-specific bioconcentration. Mut. Res.Genet. Toxicol. Environ. Mutagen 586 (1), pp. 47–57.

[41] Hagger, J.A., Depledge, M.H., Galloway, T.S. 2005. Toxicity of tributyltin in the marine mollusc *Mytilus edulis*. Mar. Poll. Bull. 51 (8–12), pp. 811–816.

[42] Emmanouil, C., Smart, D.J., Hodges, N.J., Chipman, J.K. 2006. Oxidative damage produced by Cr(VI) and repair in mussel (*Mytilus edulis* L.) gill. Mar. Environ. Res. 62 (SUPPL. 1), pp. S292–S296.

[43] Mitchelmore, C.L., Birmelin, C., Livingstone, D.R., Chipman, J.K. 1998. Detection of DNA Strand Breaks in Isolated Mussel (*Mytilus edulis* L.) Digestive Gland Cells Using the ''Comet'' Assay. *Ecotox and Environ. Saf.* 41, 51Ð58 (1998).

[44] Canty, M.N., Hutchinson, T.H., Brown, R.J., Jones, M.B., Jha, A.N. 2009. Linking genotoxic responses with cytotoxic and behavioural or physiological consequences: Differential sensitivity of echinoderms (*Asterias rubens*) and marine molluscs (*Mytilus edulis*). Aquat. Toxicol., 94 (1), pp. 68–76.

[45] Rank J. 2009. Intersex in Littorina littorea and DNA damage in *Mytilus edulis* as indicators of harbour pollution Ecotox. Environ. Saf., 72 (4), pp. 1271–1277.

[46] Rank J. Lehtonen, K.K., Strand, J., Laursen, M. 2007. DNA damage, acetylcholinesterase activity and lysosomal stability in native and transplanted mussels (*Mytilus edulis*) in areas close to coastal chemical dumping sites in Denmark Aquat.Toxicol., 84 (1), pp. 50–61.

[47] Woo, S., Kim, S., Yum, S., Yim, U.H., Lee, T.K. 2006. Comet assay for the detection of genotoxicity in blood cells of flounder (*Paralichthys olivaceus*) exposed to sediments and polycyclic aromatic hydrocarbons Mar. Poll. Bull. 52 (12), pp. 1768–1775.

[48] Baussant, T. *et al*. 2010.

Annex 11: Background document: Sediment seawater elutriate and pore-water bioassays with early developmental stages of marine invertebrates

Bivalve D-larva

Sea-urchin pluteus larva

Background

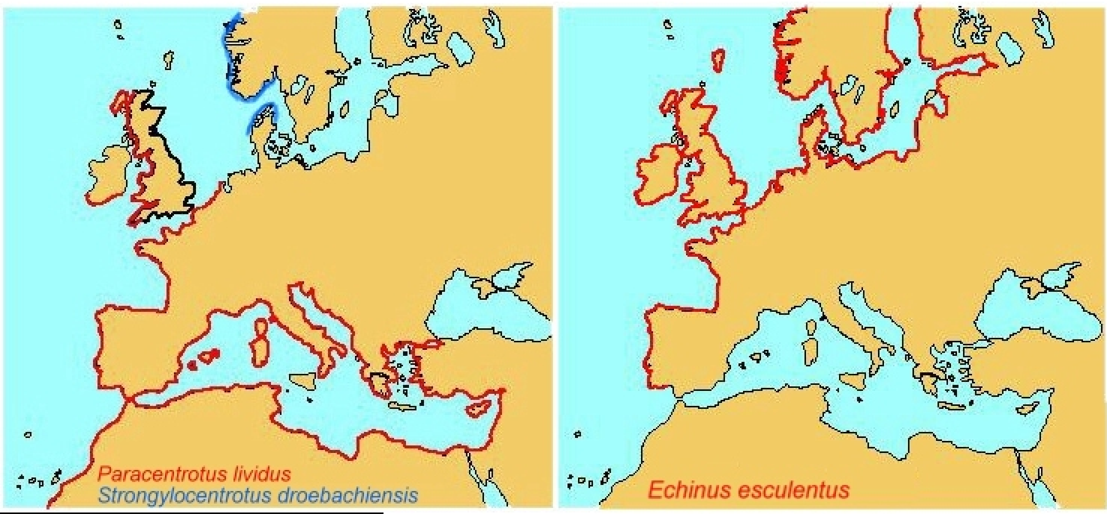
1. The embryogenesis and early larval development of marine invertebrates have been frequently used as a rapid, sensitive, cost-effective biological tool for the assessment of seawater, sediment elutriates and pore-water quality. Early developmental stages are generally more sensitive than adults and 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. . 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; Saco-Álvarez *et al*., 2010). 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 hours incubation at 18 to 24ºC (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, size increase.

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 EC_{50} values of 18 priority pollutants to bivalve vs. sea-urchin embryos reflected a correlation coefficient $r^2=0.96$ ($p<0.01$) and a slope $b=1.00$ (Beiras and Bellas, 2008). 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* and *Echinus sculentus*, extend the applicability of the assay with indigenous species to Northern countries (see figures).

4. Within bivalves, *Crassostrea gigas* and, in the US *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 of a broader range of salinities, including estuarine waters down to 20 ppt (His and 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: size increase (Saco-Álvarez *et al*., 2010).



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. The toxicity of sediment can be assessed by either obtaining an elutriate from the sediment (by mixing with control seawater) or by directly obtaining the interstitial pore-water from the sediment and undertaking toxicity tests on these aqueous solutions using water column (pelagic) organisms. 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 seawater

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. Generally, the embryo-larval bioassay showed higher sensitivity than the amphipod bioassay to polluted sediments (Becker *et al.*, 1990; Long *et al.*, 1990; Carr and Chapman, 1992). However, similar sensitivities have also been reported (Williams *et al.*, 1986). However the differences in estimates of toxicity using different organisms can be large, and different tests indicate may reflect different patterns or mechanizms of toxicity (Long *et al.*, 1996). Therefore, comparisons of different sediment toxicity tests must be conducted using samples representing a broad range of types of pollution in order to evaluate the comparability of the different tests.

Confounding factors

8. In order to avoid false positives, water quality values 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. In the case of sea urchins Saco-Álvarez *et al.* (2010) described the optimal range for salinity from 31 to 35, and from 7.0 to 8.5 for pH.

9. The presence of the toxic substances such as unionized ammonia and H_2S has been identified as the main sources of false positives in sediment elutriate toxicity testing where the objective is to investigate responses to chemical contaminants (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. For NH_3 Saco-Álvarez *et al.* (2010) obtained an EC_{10} of 68.4 μg/L and a NOEC/LOEC of 40/80 μg/L using *Paracentrotus lividus*.

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, 20ºC (48 h) is recommended for mussels and *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.

Endpoints measured

12. The endpoint 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 at this stage, but may be considered as a field for future research.

13. In sea-urchins normal larvae should exhibit four fully formed arms (two longer post-oral arms and two shorter oral arms) and a regular 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. An alternative endpoint for the sea-urchin test was recently proposed by Saco-Álvarez *et al*. (2010), who measure the size increase in 48 h. This avoids lengthy and subjective microscopical inspection, speeding up test readings, makes automatic reading feasible, and allows a more than twofold increase in sensitivity compared to the classical morphological endpoint.

Assessment criteria

14. Discrete approach: ICES (2008) currently recommends classification of the toxicity of a liquid sample as "elevated" when embryo abnormalities are >20% for bivalves and >10% for sea urchins, and "high concern" when they are >50% for both invertebrates.

15. Generally speaking, an elutriate can be classified as toxic when it induces a statistically significant reduction in the endpoint (either normal morphology or size increase) compared to the elutriate from the reference site, for a confidence level of 95%. Percentages of response 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 physico-chemical and microbiological properties of an elutriate, some of which may affect the response.

16. Continuous approach: Once identified as polluted, the toxicity of any sediment elutriate that causes a marked inhibition in normal development can be quantified by serial dilution with reference seawater, and calculation of the toxic units (TU). $TU = 100 / ED_{50}$, where ED_{50} 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: $P' = (P - Pc) \times 100 / (100 - Pc)$; where P and P' are the raw and corrected abnormality percentages, and Pc is the control abnormality. Once corrected, percentages must be arcsine trans-

formed 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 (Saco-Álvarez *et al.*, 2010).

17. In the case of the sea-urchin test Durán and Beiras (2010) developed quantitative assessment criteria for the size increase endpoint on the basis of the distribution of results from sites not significantly different to reference. The methodology to obtain BAC and EAC values followed OSPAR (2009). The resulting **BAC** value was PNR=0.702, which means a 30% decrease in growth (size increase) in the tested population.

18. A more detailed evaluation of the results from the sea-urchin test can be obtained by pooling the results from sites not significantly different to reference in a first dataset, and pooling toxic sites in a second dataset. Taking different percentiles from those distributions the following environmental assessment criteria (EAC) for Percent Net Response (PNR) and Toxic Units (TU) data were obtained.

19. A BAC of 22 was set for mussel larvae (Table 1).

20. EAC-values for both assays were retained at 50% as recommended earlier by ICES, either mortality (mussel embryo) or reduced growth (sea urchin embryo).

Table 1. Background response for mussel embryo bioassays (mortality); data from IEO-Vigo.

Average	90-percentile	median	10-percentile	n
14.7	22.3	8	3.2	38

Quality assurance

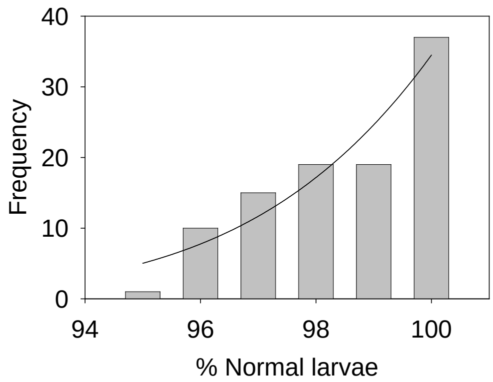
21. Sediment manipulations during sampling, storage and testing, and quality of the test organisms have been often identified as the main sources of variability in sediment toxicity bioassays. Concerning the first point, sediments intended to toxicity testing should not be frozen but stored under refrigeration 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 glassfibre 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.

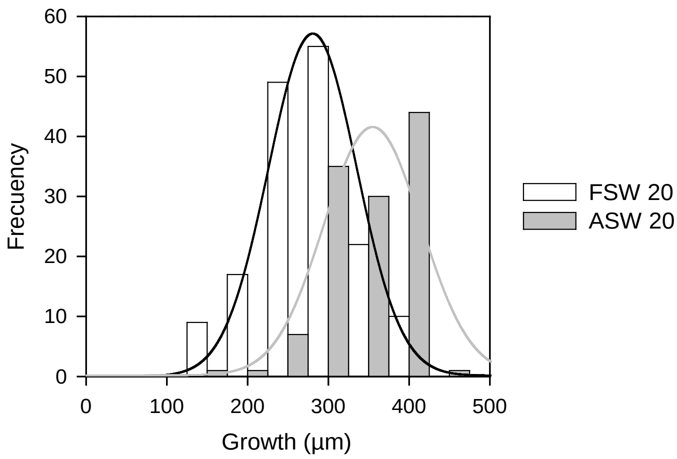
22. Concerning the effect of homogeneous biological material, interlaboratory comparisons carried out following strict protocols are necessary. In these intercalibrations it would be desirable that not only different populations of a certain species, but also different species (oysters, mussels, clams,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

23. 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 (see also AFNOR 2009). Preliminary results of background response levels for *Mytilus* embryo bioassays are shown in Table 1 below. Taking as acceptability criteria the 10th percentile of the distribution of all controls with natural filtered seawater (FSW) throughout several years during the natural spawning season (April, May and June), a minimum of 68% normal D-larvae in controls is required. Nevertheless if the bioassay is carried out outside the spawning season, failure to reach the acceptability criteria is likely to occur, and a compromise between sensitivity and feasibility must be reached.

24. In the case of the *P. lividus* normal larval development, the distribution of the endpoints measured (percentage of normal larvae, and size increase) in controls with natural filtered seawater (FSW) and artificial seawater (ASW), throughout several years of tests conducted at 20ºC for 48h, was the following (Saco-Álvarez *et al*., 2010):





25. From these data, and taking the 5th percentile as the acceptability criteria, a test is correct when mean response in the control exceeds 91% embryogenesis success and 218 µm size increase in FSW (natural filtered seawater) or 253 µm in ASW (artificial seawater).

26. 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 using a chart of the reference toxicant EC_{50} historical values.

References

AFNOR. 2009. Qualité de l'eau. Bioindicateur de la toxicité potentielle de milieux aqueux. Determination de la toxicité potentielle d'echantillons aqueux sur le developpement embryo-larvaire de bivalve. Norme expérimentale XP T90-382. Septembre 2009. AFNOR.

ASTM. Standard guide for conducting static acute toxicity tests with echinoid embryos. ASTM. (E 1563-95), 962–980. 1995. Annual Book of ASTM Standards. American Society for Testing and Materials.

Becker DS, Bilyard GR, Ginn T.C. 1990. Comparisons between sediment bioassay and alterations of benthic macroinvertebrate assemblages at a marine superfund site: Commencement Bay, Washington. Environmental Toxicology and chemistry 9:669–685.

Beiras, R. 2001. Comparison of methods to obtain a liquid phase in marine sediment toxicity bioassays with *Paracentrotus lividus* sea urchin embryos. Arch. Environ. Contam. Toxicol. 42, 23–28.

Beiras R, Bellas J. 2008. Inhibition of embryo development of the *Mytilus galloprovincialis* marine mussel by organic pollutants; assessment of risk for its extensive culture in the Galician Rias. Aquaculture 277:208–212.

Beiras R, Saco-Álvarez L, Fernández N, González JJ. 2005. Bivalve and sea-urchin embryogenesis bioassay as biological tool for monitoring coastal pollution in Europe. SETAC Europe 15th annual meeting, Lille.

Cardwell RD, Woelke CE, Carr MI, Sanborn EW. 1976. Sediment and elutriate toxicity to oyster larvae. In: Proceedings of the speciality conference on Dredging and its environmental effects. Mobice, Alabama. American society of Civil Engineers, New York.

Carr RS, Chapman DC. 1992. Comparison of solid-phase and pore-water approaches for assessing the quality of marine and estuarine sediments. Chemistry and Ecology. 7:19–30.

Carr RS. 1998. Marine and estuarine porewater toxicity testing. In: Wells PG, Lee K, Blaise CB (eds) Microscale testing in aquatic toxicology. Advances, techniques, and practice. CRC Press, Boca Raton, p 523–538.

Durán I, Beiras, R. 2010. Assessment criteria for using the Sea-urchin Embryo Test with sediment elutriates as a tool to classify the Ecotoxicological status of marine water bodies. Environ. Toxicol. Chem. 29: 1192–1198.

His, E., Beiras, R. 1995. Monitoring fresh and brackish water quality around shellfish farming areas with a bivalve embryo and larva simplified bioassay method. Oceanologica Acta, 18, 591–595.

His, E., Seaman, M.N.L., Beiras, R. 1997. A simplification of the bivalve embryogenesis and larval development bioassay method for water quality assessment. Wat. Res. 31(2): 351–355.

His, E., Beiras, R., Seaman, M.N.L. 1999. The assessment of marine pollution - Bioassays with bivalve embryos and larvae. In: Southward, A.J., Tyler, P.A., Young, C.M. (Eds.) Advances in Marine Biology Vol. 37, Academic Press, San Diego, pp. 1–178.

ICES. 2008. Report of the Fourth ICES/OSPAR Workshop on Integrated Monitoring of Contaminants and their Effects in Coastal and Open Sea Areas (WKIMON IV). ICES CM2008/ACOM 49. ICES Advisory Committee. Copenhagen, 71 pp.

Knezovich, J.P., Steichen, D.J., Jelinski, J.A., Anderson, S.L. 1996. Sulphide tolerance of four marine species used to evaluate sediment and pore-water toxicity. Bull Environ Contam Toxicol 57, 450–457.

Kobayashi N. 1995. Bioassay data for marine pollution using echinoderms. In: Cheremisinoff PN (ed) Encyclopedia of environmental control technology. Gulf Publ. Co., Houston, p 539–609.

Long ER, Buchman MF, Bay SM, Breteler RJ, Carr RS, Chapman PM, Hose JE, Lissner AL, Scott J, Wolfe DA. 1990. Comparative evaluation of five toxicity tests with sediments from San Francisco Bay and Tomales Bay, California. Environmental Toxicology and chemistry 9:1193–1214.

Long ER, Robertson A, Wolfe DA, Hameedi J, Sloane GM. 1996. Estimates of the spatial extent of sediment toxicity in major US estuaries. Environmental Science and Technology 30:3585–3592.

Matthiessen P, Bifield S, Jarret F, Kirby MF, Law RJ, McMinn WR, Sheahan DA, Thain JE, Whale GF. 1998. An assessment of sediment toxicity in the river Tyne estuary, UK by means of bioassays. Marine Environmental Research 45:1–15.

OSPAR. 2009. Agreement on CEMP assessment criteria for the QSR2010, Agreement number: 2009-2. OSPAR Commission.

Quiniou F, His E, Delesmont R, Caisey X. 2005. Bio-indicateur de la toxicité potentielle de milieux aqueux: bio-essei "developpement embryo-larvaire de bivalve". Ed. Ifremer, 24 pp.

Saco-Álvarez L, Durán I, Lorenzo JI, Beiras, R. 2010. Methodological basis for the optimization of a marine sea-urchin embryo test (SET) for the ecological assessment of coastal water quality. Ecotoxicol. Environ. Saf. 73: 491–499.

Thain, J. E. 1991. Biological effects of contaminants: oyster (*Crassostrea gigas*) embryo bioassay. ICES. (No. 11). 1991. Copenhagen, ICES. Techniques in marine environmental sciences.

Williams LG, Chapman PM, Ginn TC. 1986. A comparative evaluation of marine sediment toxicity using bacteria luminiscence, oyster embryo and amphipod sediment bioassays. Marine Environmental Research 19:225–249.

Annex 12: Background document: Sediment seawater elutriate and pore-water bioassays with copepods (*Tisbe*, *Acartia*), mysids (*Siriella, Praunus*), and decapod larvae (*Palaemon*)

Tisbe battagliai

Siriella armata

Palaemon elegans

Background

1. The toxicity of sediment can be assessed either through the exposure of test organisms to whole sediment, or through the exposure of pelagic organisms to sediment seawater elutriates or to pore-waters. In tests with elutriates or pore waters, crustaceans, and particularly early life stages, have been found to be several orders of magnitude more sensitive to insecticides than echinoderms and molluscs (Ramamoorthy and 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 and Williams, 1989; UNEP, 1989). The endpoint recorded may be mortality or motility after 48 to 96h incubation in the test samples at 20ºC 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 48 h *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. Because these organisms undergo direct development in short time periods, they are suitable for life cycle assessments. Some European mysids such as *Neomysis* (for brakish waters), *Praunus* (Garnacho *et al.*, 2000; Mclusky and Hagerman, 1987) and *Siriella* (Pérez and Beiras, 2010) 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 spe-

cies (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 because 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 H_2S, 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 foodweb in coastal and oceanic ecosystems.

Assessment criteria

8. ICES (2008) currently recommends classification of the toxicity of a seawater sample as "elevated" when *Tisbe* mortality is >10% and "high concern" when it is >50%.

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 glassfibre 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, and protocols should include a period of acclimation to avoid sharp changes. Results of background response levels for *Tisbe* bioassays are shown in Table 1, resulting in a BAC of 5.0.

Table 1. Preliminary results of background response levels for *Tisbe* bioassays (mortality) – data from Cefas.

Average	210-percentile	median	90-percentile	n
1.3	0.0	0.0	5.0	28

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

Bellas J, Beiras R, Mariño-Balsa JC, Fernández N. 2005. Toxicity of organic compounds to marine invertebrate embryos and larvae: a comparison between the sea urchin embryogenesis bioassay and alternative test species. Ecotoxicology, 14: 337–353.

Cheung YH, Neller A, Chu KH, Tam NFY, Wong CK, Wong YS, Wong MH. 1997. Assessment of sediment toxicity using different trophic organisms. Arch. Environ. Contam. Toxicol. 32: 260–267.

Garnacho E, Peck LS, Tyler PAA. 2000. Variations between winter and summer in the toxicity of copper to a population of the mysid *Praunus flexuosus*. Mar Biol 137:631–636.

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.

ICES. 2008. Report of the Fourth ICES/OSPAR Workshop on Integrated Monitoring of Contaminants and their Effects in Coastal and Open Sea Areas (WKIMON IV). ICES CM2008/ACOM 49. ICES Advisory Committee. Copenhagen, 71 pp.

Mariño-Balsa, J.C., Poza, E., Vázquez, E., Beiras, R. 2000. Comparative toxicity of dissolved metals to early larval stages of *Palaemon serratus*, *Maja squinado* and *Homarus gammarus* (Crustacea: Decapoda). Arch. Environ. Contam. Toxicol. 39, 345–351.

Matthiessen P, Bifield S, Jarret F, Kirby MF, Law RJ, McMinn WR, Sheahan DA, Thain JE, Whale GF. 1998. An assessment of sediment toxicity in the river Tyne estuary, UK by means of bioassays. Marine Environmental Research 45:1–15.

Mclusky DH, Hagerman L. 1987. The toxicity of chromium, nickel and zinc: effects of salinity and temperature, and the osmoregulatory consequences in the mysid *Praunus flexuosus*. Aquat Toxicol 10:225–238.

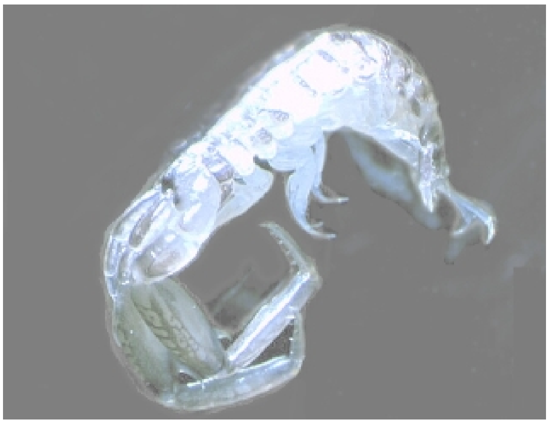
Pérez S, Beiras R. 2010. The mysid *Siriella armata* as a model organism in marine ecotoxicology: comparative acute toxicity sensitivity with *Daphnia magna*. Ecotoxicology 19:196–206.

Ramamoorthy S and Baddaloo EG. 1995. Handbook of chemical toxicity profiles of biological species Vol I Aquatic species, Lewis Publishers, Boca Raton.

UNEP, FAO, and IAEA. Estimation of the toxicity of pollutants to marine phytoplanktonic and zooplanktonic organisms. (44), -24. 1989. Reference methods for marine pollution studies.

US-EPA. 2002 Short-term methods for estimating the chronic toxicity of effluents and receiving waters to marine and estuarine organisms. EPA-821-R-02-014. US Environmental Protection Agency, Washington.

Annex 13: Background document: Whole sediment bioassays with amphipods (*Corophium* sp) and *Arenicola marina*



Corophium multisetosum

Background

1. The toxicity of sediment can be assessed either through the exposure of pelagic organisms to sediment seawater elutriates or to pore-waters, or through the exposure of test organisms to whole sediment. The *Rhepoxynius abronius* amphipod test is commonly used in North America to evaluate the quality of whole sediments intended for dredging or dumping, and very detailed protocols are available (Swartz *et al*., 1985; ASTM, 1992). The endpoint is survival after ten days incubation in the whole sediment at 20ºC. 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 and 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 and Thain, 2001).

3. Other sediment dwelling species from different taxa (polychaetes, echinoderms, bivalves) may be also suitable after methodological standardization and sensitivity comparisons with amphipods. Furthermore, *Corophium* is not tolerant of 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 ten day exposure (Bat and Raffaelli, 1998), and 28-days growth (Nipper and 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 H_2S 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. According to USEPA (1998) a sediment sample is classified as toxic when it induces an amphipod mortality 20% higher than control and the difference is statistically significant. Similarly, ICES (2008) currently recommends classification as "elevated" when *Corophium* mortality is >30% and "high concern" when it is >60%. For *Arenicola* these benchmarks go down to >10% for "elevated" and >50% for "high concern" (ICES, 2008).

8. ANOVA and *a posteriori* Dunnett's test allows comparison to control and classification of sampling sites into homogeneous groups according to their toxicity. Mortality data must be arcsine transformed prior to analysis. When data from different test rounds were pooled together for statistical analysis, mortalities 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 mortality percentages, and Pc is the control mortality. In this case no control treatment is available and Tukey's rather than Dunnett's *post hoc* test is preferred. 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 to 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. Because 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 criterion 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 and Thain (2001) raise this threshold to 15%. Results of background response levels for *Corophium* and *Arenicola* bioassays are shown in Table 1. All laboratories show a 90th percentile for mortality higher than 10% and most above the recommended 15%, indicating that special care must be taken in avoiding any damage to the individuals during collection, maintenance and transfer into the experimental beakers.

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.

Table 1. Background response levels for whole sediment bioassays (mortality); the median 90-percentile, i.e. BAC, is 18.4%.

Test	lab	Average	10-percentile	median	90-percentile	n
Corophium	RIKZ	12.3	6.6	10.5	19.3	4
Corophium	Cefas	9.5	0.0	6.7	20.0	21
Corophium	IEOV	7.7	5.6	6.3	10.8	5
Corophium	AZTI	10.4	4.8	10.8	17.4	27

Test	lab	Average	10-percentilemedian90-percentile	n		
Arenicola	Cefas	4.7	0.0	0.0	13.3	20

References

ASTM. Standard guide for conducting 10-day static sediment toxicity tests with marine and estuarine amphipods. ASTM. (E 1367-92), 1138–1163. 1992.

Bat L, Raffaelli D. 1998. Sediment toxicity testing: a bioassay approach using the amphipod *Corophium volutator* and the polychaete *Arenicola marina*. J Exp Mar Biol Ecol 226: 217–239.

Casado-Martínez MC, Beiras R, Belzunce MJ, González-Castromil MA, Marín-Guirao L, Postma JF, Riba I, DelValls TA. 2006. Interlaboratory assessment of marine bioassays to evaluate the environmental quality of coastal sediments in Spain. IV. Whole sediment toxicity test using crustacean amphipods. Ciencias Marinas 32(1B):

De Witt TH, Ditsworth GR, Swartz RC. 1988. Effects of natural sediment features on survival of the phoxocephalid amphipod Rhepoxynius abronius Mar Environ Res 25: 99–124.

ICES. 2008. Report of the Fourth ICES/OSPAR Workshop on Integrated Monitoring of Contaminants and their Effects in Coastal and Open Sea Areas (WKIMON IV). ICES CM2008/ACOM 49. ICES Advisory Committee. Copenhagen, 71 pp.

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

Phillips BM, Anderson BS, Hunt JW. 1997. Measurement and distribution of interstitial and overlying water ammonia and hydrogen sulphide in sediment toxicity tests. Mar Environ Res 44(2): 117–126.

Swartz RC, de Ben WA, Jones JKP, Lamberson JO, Cole FA. 1985. Phoxocephalid amphipod bioassay for marine sediment toxicity. In: Aquatic toxicology Hazard Assessment: 7th Symposium, ASTM STP 854, pp. 284–307. Ed. by RD Cardwell, R Purdy and RC Bahner. American Society for Testing Materials, Philadelphia.

Roddie B, Thain J. 2001. Biological effects of contaminants: *Corophium* sp. sediment bioassay and toxicity test. ICES Techniques in Marine Environmental Sciences, No. 29, 21 pp.

USEPA. 1994. Methods for assessing the toxicity of sediment-associated contaminants with estuarine and marine amphipods. United States Environmental Protection Agency, EPA/600/R-94/025.

USEPA. 1998. Evaluation of dredge material proposed for discharge in waters of the US. Testing manual. United States Environmental Protection Agency, EPA/823/F/98/005.

Annex 14: Background Document (revised): DNA adducts

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

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

Of the techniques currently available for the detection of DNA adducts the most sensitive method for the detection of a wide range of compounds chemically bound to DNA is the ^{32}P-postlabelling assay (Gupta *et al.*, 1982). The method possesses a number of advantages that make it suitable for the assessment of DNA adduct induced by environmental genotoxins (for a review see Beach and Gupta, 1992; Phillips, 1997; Phillips, 2005). The technique is applicable to any tissue sample from which DNA can be isolated and is also extremely sensitive, capable of detecting one adducted nucleotide in 10^9–10^{10} undamaged nucleotides from 5–10 μg DNA. In addition, providing the adduct is amenable to the labelling reaction and subsequent thin layer chromatography, its prior characterization is not required. It is this last feature that makes the assay particularly appropriate to aquatic biomonitoring, because it is suitable for the analysis of the diverse array of adducts induced by complex mixtures of environmental chemicals. It is important to note that ^{32}P-postlabelling is only semi-quantitative as not all DNA adducts are labelled with the same efficiency and the various enrichment and chromatograph steps involved will preferentially select certain adducts. However, the assays sensitivity, coupled with the assays ability to detect a wide range of 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 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 demonstrated that at certain contaminated sites the prevalence of DNA adducts are associated with the prevalence of toxicopathetic lesions including foci of cellular alteration and neoplasia (for review see Reichert *et al.*, 1998).

Studies from North America and Europe suggest that DNA adduct levels are not markedly influenced by factors such as age, sex, season or dietary status, which are known to confound the interpretation of other biomarkers (e.g. EROD). However, validation of any biomarker, including DNA adducts in a species of interest is essential to ensure against any unforeseen species-specific responses (Reichert *et al.,* 1999). While there is no evidence to suggest that environmental factors such as salinity and temperature significantly affect the formation of DNA adducts these factors should always be considered, as it is known that cellular detoxification systems (e.g. Cyp1A) are influenced by changes in environmental variables (Sleiderink *et al.,* 1995).

Species selection and target tissue

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

Of the affected organs, liver is the most commonly studied when fish are used as sentinel organisms. Field data infers a chemical aetiology for many of the commonly observed hepatic lesions seen in wild fish collected from contaminated areas. Laboratory data supporting this association stems from biochemical and molecular studies which have revealed the liver to be the major site for contaminant detoxification pathways (e.g. cytochrome P-450-mediated biotransformation enzyme systems). Furthermore, contaminant metabolisms 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-postabelling assay is that it is not tissue specific and assuming sufficient DNA can be extracted it can be applied in a fit-for-purpose manner in any tissue of choice. To this end it has been used successfully in a range of tissues (both invertebrate and vertebrate), including liver, intestine, gill, brain, gonad and digestive 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 10^9–10^{10} bases (Reichert *et al.*, 1999). Following the adduct enrichment step, the 3′-monophosphates are radiolabelled at the 5′-hydroxyl using ^{32}P-ATP and T4-polynucleotide kinase to form 3′,(^{32}P)5′-bisphosphates. Separation of the ^{32}P-labeled adducts is accomplished by multidimensional high-resolution anion exchange thin-layer chromatography. Autoradiography is then used to locate the 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-postlabelling 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) [2] 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.

---

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

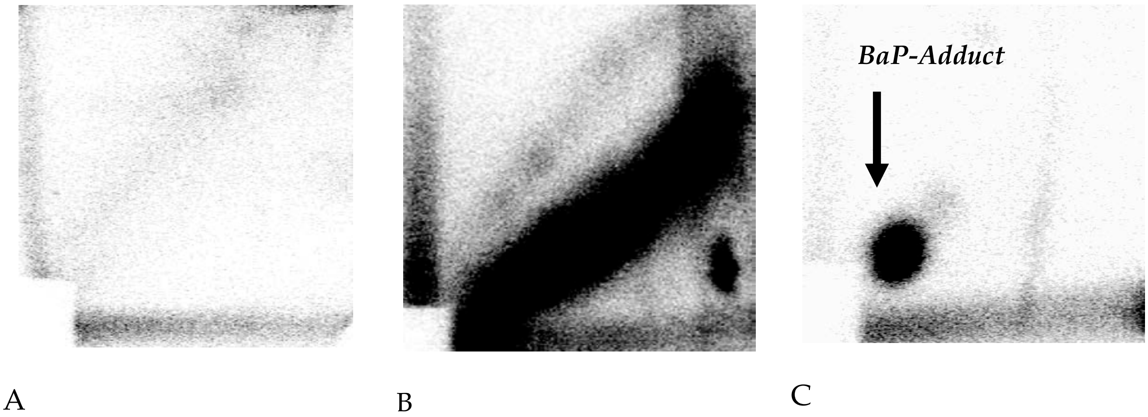


Figure 1. Representative hepatic DNA adducts profiles produced following ^{32}P-postlabelling. (A) DNA adduct profile obtained from a site with a low level of PAH contamination. A faint DRZ is visible, indicating a low level of DNA adducts representative of a clean reference location. (B) DNA adduct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

Determination of threshold level of significant effects 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 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 Northeast 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:

- ^{32}P-postlabelling studies should be conducted using internationally agreed protocols incorporating appropriate positive and negative control samples (Phillips and Castegnaro, 1999; Reichert *et al.*, 1999).
- All studies need to include supporting environmental data to confirm the contaminant load at the reference location and where possible supporting

biomarker and histopathological data to confirm health status of the individual.

- While the assay ^{32}P-postlabelling can be applied to any species deemed fit for purpose, it should only be applied to those species where there is sufficient background information available on life-history traits and behaviour (e.g. migration).

Derivation of assessment criteria

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

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 p.13 above for method of estimation). This value is also indicative for flatfish (to be verified).

Summary of assessment criteria.

Biological Effect	Qualifying comments	Background Response Range	Elevated Response Range	High and Cause for Concern Response
DNA adducts; nm adducts / mol DNA	Dab	≤ 1.0	> 1.0	(> 6)
	Flounder	≤ 1.0	> 1.0	(> 6)
	Cod	≤ 1.6	> 1.6	> 6
	Haddock	≤ 3.0	> 3.0	(> 6)

Concluding remarks

- *DNA adducts as biomarkers of genotoxic exposure*. DNA adducts provide a measure of biologically active contaminant to have reached a critical cellular target (DNA). They are persistent and therefore considered a 'cumulative index' of exposure to genotoxins and a significant body of research demonstrates their importance in the initiation and progression of carcinogenesis induced by important environmental contaminants (e.g. PAHs).
- *Safety considerations when conducting the 32p-postlabelling assay.* The ^{32}P-postlabeling assay uses large amounts of ^{32}P, which is an energetic beta emitter. This requires specialist laboratories may limit the use of the assay to a few appropriately equipped research groups.*Applicability across OSPAR maritime area*. DNA adducts have been applied in a wide range of species across the whole OSPAR maritime area including blue mussels, *Mytilus* sp, perch (*Perca fluviatilis*), dab (*Limanda limanda*), European flounder (*Platichthys flesus*), eelpout (*Zoarces viviparous*) and cod (*Gadus morhua*). A particular advantage of the ^{32}P-postblabelling assay is that it is not spe-

cies-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.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Annex 15: Background Document: *In vitro* DR-Luc/DR-CALUX® bioassay for screening of dioxin-like compounds in marine and estuarine sediments

Executive summary

Applicability across the OSPAR maritime area. The *in vitro* DR-Luc assay (also called DR-CALUX®, a trademark of BDS, NL, hereafter generally referred to as DR-Luc), is a rapid, extremely sensitive and cost-effective tool for screening marine and estuarine sediments for dioxin-like compounds including congeners of polychlorinated dibenzo-p-dioxins (PCDDs), dibenzo- furans (PCDFs) and chlorinated biphenyls (PCBs). The DR-Luc assay is available for immediate deployment within the OSPAR JAMP CEMP. The DR-Luc assay has been recommended by ICES and is of sufficient standing in terms of methodological development and application for uptake across the whole OSPAR area.

Quality assurance. QA procedures are in place and interlaboratory performance studies are organized frequently, but there remains a need for QA within international programmes such as BEQUALM. The methodology for DR-Luc and related extraction protocols are well developed and available through ICES TIMES series documents. DR-Luc data can be submitted to the ICES database for subsequent assessment, as appropriate, by ICES/OSPAR.

Influence of environmental variables. In general, there is little influence of environmental variables on the test conditions and bioassay response; the use of extracts will reduce any disturbing factors. Sediments should be sampled according to guidelines for chemical analysis to take account of OC content and particle size.

Thresholds and assessment tools. Three assessment classes were derived for DR-Luc based on silica clean up / 24 h exposure; a background response <10 pg TEQ g^{-1} dry wt; an elevated response (warning level) of >10–<40 pg TEQ g^{-1} dry wt and; a high and cause for concern response of >40 pg TEQ g^{-1} dry wt.

Synergism between CEMP/MSFD and WFD. The DR-Luc bioassay can be immediately applied in offshore and coastal sediments and is equally suitable for estuarine and freshwater sediments (see further also BG document on water *in vivo* bioassays). As such, the use of DR-Luc can play a role in linking the MSFD with the WFD.

Background

Dioxin levels in the marine environment have declined significantly in the past two decades due to reductions in emissions from man-made sources (Rappe, 1996; Aylward and Hays, 2002). However, degradation in the environment is slow and therefore dioxin-like compounds from past releases are expected to remain in the environment for many decades. The term 'dioxin-like compounds' refers to a group of structurally similar congeners known as polychlorinated dibenzo-p-dioxins (PCDDs), polychlorinated dibenzofurans (PCDFs) and some polychlorinated biphenyls (PCBs) (see also OSPAR Background Document on dioxins; OSPAR Commission, 2007). Dioxin-like compounds are unintentionally released by-products of the combustion of chlorinated compounds in the environment. In addition, there are a number of other compounds that exhibit dioxin-like properties, such as polybrominated biphenyls (PBBs) and polycyclic aromatic hydrocarbons (PAHs).

In the past two decades, there has been growing environmental concern regarding dioxins, and other compounds that have dioxin-like properties. The major concerns with dioxin-like compounds are their effects upon wildlife and human health due to their resistance to degradation and ability to be bioaccumulated (Van den Berg *et al.*, 1998). They have also been shown to produce a wide variety of toxic and biochemical effects via aryl hydrocarbon receptor (AhR)-mediated signalling pathways (Mandal, 2005). The effects on laboratory animals and wildlife include developmental and reproductive effects, immunotoxicity, neurotoxicity and carcinogenesis (for more details and references, see OSPAR Commission, 2007). Animals at particular risk are fish-eating top predators, such as otters (Murk *et al.*, 1998), seals (Vos *et al.*, 2000) and birds (Bosveld, 1995; Henshel, 1998). The effects of dioxin-like compounds in humans include high acute toxicity, skin lesions, developmental and reproductive abnormalities, and probably cancer (WHO, 2000; Aylward *et al.*, 2003; Heilier *et al.*, 2005). It has been shown that aquatic organisms can ingest dioxin-like compounds that have been flushed into surface water from land, providing a potential pathway into the food chain (Leonards *et al.*, 2008).

Dioxin-like compounds share (at least initially) a common mode of action by binding to the aryl hydrocarbon (Ah) receptor, which mediates and interacts with a series of biological processes including cell division and growth and homeostatic functions (Puga *et al.*, 2005; Stevens *et al.*, 2009). Of 75 PCDD congeners, only seven have been have been identified as having dioxin-like toxicity (Liem and Zorge, 2005) and only ten of the 135 PCDFs are thought to have dioxin-like toxicity (Aarts and Palmer, 2002). For PCBs, only twelve of the 209 congeners are thought to have dioxin-like toxicity (Liem and Zorge, 2005). The Ah receptor or dioxin receptor based *in vitro* assay DR-Luc (also known as DR-CALUX® (Dioxin Response Chemical Activated LUciferase gene eXpression, a trademark of BDS, Amsterdam, The Netherlands) is considered to be the most useful *in vitro* bioassay technique for screening for dioxin-like compounds. However the induction of CYP1A/ EROD in fish liver (see OSPAR background document on CYP1A/EROD activity) and chronic *in vivo* bioassays (Foekema *et al.*, 2008) may also be relevant. 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, and so are applicable in a generic way. The use of extracts will minimize the influence of environmental variables and reduce any disturbing factors. Sediments should be sampled according to guidelines for chemical analysis to take account of OC content and particle size.

DR-Luc as bioassay for dioxin-like compounds

The DR-Luc is a reporter gene assay that has been developed by Wageningen University (Aarts *et al.*, 1995; Murk *et al.*, 1996) and is distributed as DR-CALUX® by Bio Detection System (BDS, NL). This system incorporates a reporter firefly gene into a cultured Rat H4IIE hepatoma cell line. Exposed to dioxin-like compounds, this system produces the enzyme luciferase, which reacts with luciferin and emits light of a characteristic wavelength with intensity proportional with the dioxin concentration. The mode of action of Ah receptor- mediated action is illustrated and further explained in Figure 1.

The DR-Luc is a highly sensitive reporter gene assay, allowing detection of 1 pM TCDD (Murk *et al.*, 1996). As such the DR-Luc assay for dioxin-like substances is much cheaper and faster than the conventional chemical HRGC-MS4 methods.

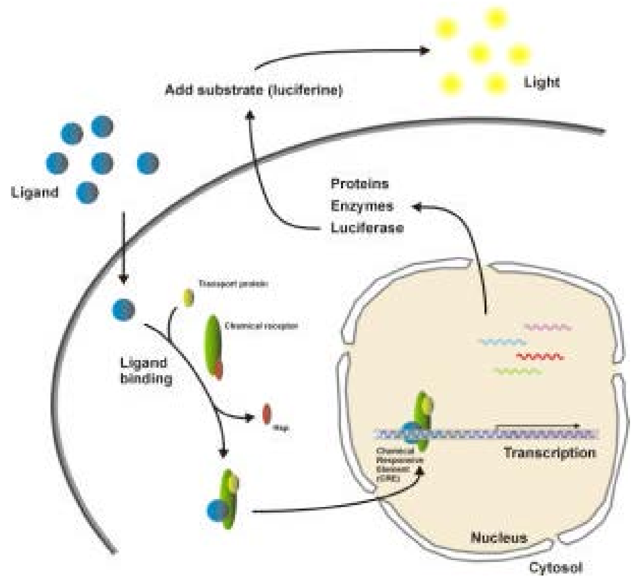


Figure 1. Activation of the Ah-receptor mediated luciferase gene in the DR-Luc bioassay (figure by RIKZ/BDS, 2006). Following activation of the receptor, the ligand–Ah receptor complex translocates to the nucleus of the cell, where it binds to specific DNA sequence, the so called DREs. The binding of the ligand-Ah receptor complex to the DREs results in changes in the expression of DR-Luc associated genes (e.g. cytochrome P4501A1). These changes in gene expression result in the disturbance of normal cell physiology. Following exposure of the cells to dioxin or dioxin-like compounds, the enzyme luciferase is produced. Addition of the substratum luciferin to lysed cells results in light production. The amount of light produced is recorded in a luminometer and is interpolated on the amount of 2,3,7,8-TCDD toxic equivalents standard curve to which the genetically modified H4IIE cells were exposed.

The response of DR-Luc is a measure of toxic potency and usually 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 HRGCMS (see OSPAR Commission, 2007).

Applicability of *in vitro* DR-Luc bioassay across the OSPAR maritime area

The DR-Luc assay is a suitable screening method for dioxins and dioxin-like-PCBs in feed and food (for example, a survey in the Netherlands to control the dioxin levels in eel (Hoogenboom *et al*., 2006)), risk assessment and management of saline and freshwater whole effluents (e.g. Oris and Klaine, 2000; Power, 2004), and for dredged material (Stronkhorst *et al*., 2002; 2003; Schipper *et al*., 2010).

The DR-Luc assay is widely recognized within Europe to be an efficient way to assess sediment quality (e.g. Hurst *et al*., 2004; Stronkhorst *et al*., 2003; Houtman *et al*., 2004, 2006; Legler *et al*., 2006a,b; Van den Brink and Kater, 2006; Sanctorum *et al*., 2007; Schipper *et al*., 2009, 2010; Hamers *et al*., 2010). Bioassays are also applied on national level by several countries (ICES, 2010). 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).

From the above studies, it was concluded that the method could be useful as screening method associated with a specific action level, because if the bioassay results are below the action level, it is most likely that results by the chemical method also would have been below. Good correlations were usually observed between DR-Luc/CALUX bioassay results obtained on marine biological matrices and results obtained from use of advanced chemical methods (Windal *et al.*, 2002; Hoogenboom, 2004). An intra- and interlaboratory study using CALUX for analysis of dioxins and dioxin-like chemicals in dredged sediments also concluded that the tool was accurate and reliable for monitoring of coastal sediments (Besselink *et al.*, 2004).

The uptake of other *in vitro* reporter gene bioassays that can be applied together with DR-Luc in a test battery, such as *in vitro* bioassays for endocrine disruption (ER-Luc, YES, YAS) and for immunotoxic and neurotoxic compounds (Hamers *et al.*, 2010), as well as general toxicity (e.g. Microtox SPT assay), should also be encouraged.

Introduction of DR-Luc bioassays to the CEMP and status of quality assurance

The DR-Luc assay is proposed in the OSPAR JAMP guidelines as a suitable specific biological effect method for monitoring of PCBs, polychlorinated dibenzodioxins and furans, and also as a suitable method for general biological effect monitoring. In addition, the DR-Luc assay can be used in Toxicity Reduction Evaluation (TRE), Toxicity Identification Evaluation (TIE), and Effect-Directed Analysis (EDA) procedures (Burgess, 2000) as well as sediment toxicity profiling (Hamers *et al.*, 2010).

A number of papers have been published describing the validation of the DR-Luc bioassay and describing the correlation between DR-Luc and HRGCMS derived 2,3,7,8-TCDD TEQs (Van den Berg *et al.*, 1998; Stronkhorst *et al.*, 2002; Besselink *et al.*, 2003; Van Loco *et al.*, 2004). It has been shown that frequent participation in interlaboratory exercises improves performance (De Boer *et al.*, 1996; Besselink *et al.*, 2004), but there remains a need for QA to be established as routine within international programmes such as BEQUALM.

The protocol for the DR-Luc assay including methods for sediment extraction is available in the ICES Techniques in Marine Environmental Sciences series on Biological Effects of Contaminants (Schipper *et al.*, 2011 submitted).

Synergism between CEMP, MSFD and WFD

Though *in vitro* DR-Luc and other bioassays are not included as ecological quality elements in the monitoring for the Water Framework Directive (WFD) (WFD CIS, 2003), it is generally accepted that they will be able to contribute to investigative monitoring and the Pressures and Impacts/Risk Assessment process (this is especially true for chronic water and sediment bioassays). Further chemical analysis can be combined with water bioassays at smaller interval time points for the purposes of trend monitoring. In this way, bioassays can be used as a partial replacement for chemical analysis of priority and/or other relevant substances and prioritizing locations for further chemical analysis. This "bioanalysis approach" can lead to more cost-efficient and cost-effective monitoring and would put the precautionary principle called for in the WFD into practice. Pilot studies carried out in the Netherlands to explore these possibilities have had promising results (e.g. Maas *et al.*, 2005). It can be concluded that clear opportunities exist for synergism between the CEMP or the MSFD and WFD for the application of DR-Luc bioassay in coastal and estuarine areas. In addition to being a cost-effective technique, the DR-Luc will strengthen the monitoring capacity for dioxin like compounds and better understand the status of dioxin pollution in marine environment.

Thresholds and assessment tools

Three assessment classes were derived for DR-Luc based on silica clean up / 24 h exposure; a background response <10 pg TEQ g^{-1} dry wt; an elevated response (warning level) of (>10–<40 pg TEQ g^{-1} dry wt and; a high and cause for concern response of >40(pg TEQ g^{-1} dry wt. These AC are based on datasets and experience from the UK, Belgium and The Netherlands. It is advised that these AC should be further refined as more data will become available.

Derivation of AC for DR-Luc

The most conservative criteria for dioxin contaminated sediments are from Canada(4 pg TEQ g^{-1})(AEA Technology, 1999) and from the US (2.5 pg TEQ g^{-1})(Thain et al., 2006) (Table 1). These criteria are "screening levels" which, if exceeded, trigger further investigation at a particular site. Exceeding a screening level does not immediately imply that there is a health risk. Any risk will be relative to the exposure assumed in the derivation of the guideline and the exposure likely in the actual situation. In some international guidelines concerning the regulation of dioxins, sediments are divided pragmatically into 'clean' and polluted locations on the basis of existing measurements of *in vitro* bioassays, as with the DR-Luc/DR-CALUX (Stronkhorst *et al.*, 2002). The expected serious chronic effect levels are the average maximum found at locations assumed to be 'clean'. For example, DR-CALUX measurements showed in Dutch surface sediments (Stronkhorst *et al.*, 2002; Klamer *et al.*, 2005) from major Dutch "clean" offshore sites up to 70 miles offshore, with values at three offshore sites below 10 pg g^{-1} (6.9 and 8 respectively). Based on this a background response level has been derived of <10 pg TEQ g^{-1} dry wt. In the analysis of dioxins and dioxin-like chemicals in sediments, ranges of TEQs in dredged sediments from rivers in the coastal zone were 12–70 pg TEQ g^{-1} dw, and on average 24 pg TEQ g^{-1} dw (Schipper *et al.*, 2010). In several studies from the Dutch and Belgium coastal zone, a range of TEQ values was observed between 9 and 27 pg TEQ g^{-1} dw, (Klamer *et al.*, 2005) and 10–42 pg TEQ g^{-1} dw sediment (Sanctorum *et al.*, 2007). The level of serious concern is then the average maximum found at locations assumed to be 'clean': >40 pg TEQ g^{-1} dry wt. The elevated response has been derived as warning level of >10–< 40 pg TEQ g^{-1} dry wt.

Table 1. International dioxin guidelines (TCDD TEQ) in sediments (dry weight basis).

Country	Maximum allowable Concentration-dry weight basis	Comments	Reference
Vietnam	150 pg/g TEQ	Dioxin heavily contaminated sites (sediments)	Hatfield consultants, 2009
USA	2.5 pg/g TEQ	Protection level	Thain et al., 2006
Canada	4 pg/g TEQ	Protection of ecological receptors	AEA Technology, 1999
Germany	5–10 pg/g TEQ	Protection of human receptors	AEA Technology, 1999
Netherlands	50 pg/g TEQ	Target value	Stronkhorst, 2002

Conclusions

- DR-Luc/DR-CALUX® *in vitro* bioassays for dioxin-like compounds 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 and continuation of QA should be by BEQUALM. Therefore, bioassay data can be submitted to the ICES database for subsequent assessment, as appropriate, by ICES/OSPAR.
- The range of *in vitro* bioassays needs to be expanded to include estrogenic and androgenic compounds, as well as neurotoxic and immunotoxic compounds and cell-based general toxicity assays.
- Appropriate protocols for DR-Luc and associated extraction methods are available through the ICES TIMES Series.
- Assessment criteria for the DR-Luc bioassay are available.
- It is recommended that OSPAR lists the DR-Luc/DR-CALUX® bioassay as a Category-II-rated method in the JAMP CEMP programme and integrated monitoring scheme.

References

Aarts J.M.M.J.G., Denison M.S., Cox M.A., Schalk C.M.A, Garrison P.A., Tullis K., de Haan L.H.J., Brouwer A. 1995. Species-specific antagonism of Ah receptor action by 2,2′,5,5′ tetrachloro- and 2,2′,3,3′,4,4′-hexachlorobiphenyl., Eur. J. Pharm Env. Tox. Section 293:463–474.

Aarts, J.M.M.J.G. and Palmer, R. 2002. Established and deduced aromatic hydrocarbon (Ah) receptor (AhR) agonists and their relative potency. Wageningen University, Department of Agrotechnology and Food Sciences, Wageningen.

AEA Technology. 1999. Compilation of EU dioxin exposure and health data, task 7 - Ecotoxicology. Produced for: European Commission DG Environment and UK Dept. of the Environment, Transport and the Regions (DETR). October. Available online:AEA Technology. 1999. Compilation of EU Dioxin Exposure and Health Data. Prepared for the European Commission DG Environment. Oxfordshire, England.

Aylward, L., Kirman, C., Cher, D., Hays, S. 2003. Analysis of Dioxin Cancer Threshold. Journal article by; Environmental Health Perspectives 111: 1145–1147.

Aylward, L.L. and Hays, S.M. 2002. Temporal trends in human TCDD body burden: Decreases over three decades and implications for exposure levels. Journal of Exposure Analysis and Environmental Epidemiology 12: 319–28.

Besselink, H., Jonas, A., Pijnappels, M., Swinkels, A., Fjellanger, K. and Brouwer, A. 2003. Comparison of DR Calux® and HRGCMS-derived TEQs: introduction of conversion factors. *Organohalogen Compounds*, 60, pp 203–206.

Besselink, H.T., Schipper, C., Klamer, H., Leonards, P., Verhaar, H., Felzel, E.C., Murk, A.J., Thain, J., Hosoe, K., Schoeters, G., Legler, J. and Brouwer, A. 2004. Intra- and interlaboratory calibration of the DR-Calux bioassay for the analysis of dioxins and dioxin-like chemicals in sediments. *Environmental Toxicology and Chemistry, 23*(12), 2781–2789.

Bosveld, A.T.C., Graneder, J., Murk, A.J., Brouwer, A., Kampen, M. van, Evers, E.H.G. and Berg, M. van den. 1995. Effects of PCBs, PCDDs, and PCDFs in common tern (*Stern hirundo*) breeding in estuarine and coastal colonies in the Netherlands and Belgium. Environ. Toxicol. Chem. 14: 99–115.

Burgess, R.M. 2000. Characterizing and identifying toxicants in marine waters: a review of marine Toxicity Identification Evaluations (TIEs). International Journal of Environment and Pollution 13: 2–33.

De Boer J. and Wells, D.E. 1996. The 1994 QUASIMEME laboratory-performance studies: chlorobiphenyls and organochlorine pesticides in fish and sediment. Mar. Pollut. Bull. 32: 654–666.

Foekema, E.M., Deerenberg, C., Murk, A.J. 2008. Prolonged ELS test with the marine flatfish sole (*Solea solea*) shows delayed toxic effects of previous exposure to PCB 126. Aquat. Toxicol. 90: 197–203.

Hamers, T., Leonards, P.E.G.. Legler, J., Vethaak, A.D., Schipper, C.A. 2010. Toxicity Profiling: an effect-based integrative tool for for site-specific sediment quality assessment. Integrated Environmental Assessment and Management 6: 761–773.

Hatfield Consultants. 2009. Comprehensive assessment of dioxin contamination in Da Nang airport, Vietnam: environmental levels, human exposure and options for mitigating impact. DANDI-II 1450.3, Hatfield Consultants Partnership.

Heilier, J.F., Nackers, F., Verougstraete, V., Tonglet, R., Lison, D. and Donnez, J. 2005. Increased dioxin-like compounds in the serum of women with peritoneal endometriosis and deep endometriotic (adenomyotic) nodules, Fertil. Steril. 84: 305–312.

Henshel, D.S. 1998. Developmental neurotoxic effects of dioxin and dioxin-like compounds on domestic and wild avian species. Environ Toxicol Chem 17: 88–98.

Hoogenboom, R., Bovee, T., Traag, W., Hoogerbrugge, R., Baumann, B., Portier, L., van de Weg, G., de Vries, J. 2006. The use of the DR CALUX bioassay and indicator polychlorinated biphenyls for screening of elevated levels of dioxins and dioxin-like polychlorinated biphenyls in eel. Mol Nutr Food Res. 50(10): 945–57.

Hoogenboom, R. 2002.The combined use of the CALUX bioassay and the HRGC/HRMS method for the detection of novel dioxin sources and new dioxin-like compounds. Environmental Science and Pollution Research. 9(5): 304–306.

Houtman, C.J., Booij, P., Jover, E., Pascual del Rio, D., Swart, K., Velzen, M. van, Vreuls, R., Legler, J., Brouwer, A. & Lamoree, M.H. 2006. Estrogenic and dioxin-like compounds in sediment from Zierikzee harbour identified with CALUX assay-directed fractionation combined with one and two dimensional chromatography analyses. Chemosphere 65: 2244–2252.

Houtman, C.J., Cenijn, P.H., Hamers, T., Lamoree, M.H., Legler, J., Murk, A.J., Brouwer, A. 2004. Toxicological profiling of sediments using *in vitro* bioassays, with emphasis on endocrine disruption. Environ. Toxicol. Chem. 23:32–40.

Hurst, M.R., Balaam, J., Chan-Man, Y.L., Thain, J.E., Thomas, K.V. 2004. Determination of dioxin and dioxin-like compounds in sediments from UK estuaries using a bio-analytical approach: chemical-activated luciferase expression (CALUX) assay Marine Pollution Bulletin 49: 648–658.

ICES. 2010. Report of the Working Group on biological effects of contaminants. Report 2010.

Klamer, H.J.C., Leonards, P.E.G., Lamoree, M.H., Villerius, L.A., Åkerman, J.E., Bakker, J.F. 2005. A chemical and toxicological profile of Dutch North sea surface sediments. Chemosphere 58: 1579–1587.

Legler, J., Bulder, A.S., De Voogt, P., Leonards, P.E.G., Murk, T.A.J. 2006a. Bioanalysis of Estrogenic and dioxin-like activity in field samples, in Estrogens and xenooestrogens in the aquatic environment: an integrated approach for field monitoring and effect assessment, (editors; Vethaak D., Schrap M., De Voogt P., SETAC, ISBN-13:978-1-880611-85-2, pp. 121–149, USA.

Legler, J., Brouwer, A. and Lamoree, M.H. 2006b. Estrogenic and dioxin-like compounds in sediment from Zierikzee harbour identified with CALUX assay-directed fractionation combined with one and two dimensional gas chromatography analyses. Chemosphere, 65, 2244–2252.

Leonards, P.E.G., Van Hattum, B. and Leslie, H. 2008. Assessing the risks of persistent organic pollutants to top predators: a review of approaches. Integrated Environmental Assessment and Management, Vol 4, 386–398.

Liem A.K.D. and J.A. van Zorge. 1995. Dioxins and related compounds: Status and regulatory aspects. Environmental Science and Pollution Research 2: 46–56.

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

Mandal, P.K. 2005. Dioxin: a review of its environmental effects and its aryl hydrocarbon receptor biology Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology 175 (4): 221–230.

Murk, A.J., P.E.G. Leonards, A.S. Bulder, *et al.* 1997. The Calux (chemical activated luciferase expression) assay adapted and validated for measuring TCDD equivalents in blood plasma. *Environm. Toxicol. & Chem.* 16 (8), 1583–9.

Murk, A.J., Legler, J. Denison, M.S., Giesy, J.P., Van de Guchte, C. and Brouwer, A. 1996. Chemical-activated luciferase gene expression (CALUX): A novel in vitro bioassay for Ah receptor active compounds in sediments and pore water. Fund. Appl. Toxicol. 33: 149–160.

Murk, A.J., Leonards, P.E.G., Hattum, B. van, Luit, R., Weiden, M.E.J. van der and Smit, M. 1998. Application of biomarkers for exposure and effect of polyhalogenated aromatic hydrocarbonsin naturally exposed European otters (*Lutra lutra*). Environ. Toxicol. Pharm. 6(2): 91–102.

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

OSPAR Commission. 2007. Hazardous Substances Series OSPAR Background Document on Dioxins, 56 pp.

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

Puga, A., Tomlinson, C.R., Xia, Y. 2005. Ah receptor signals cross-talk with multiple developmental pathways. Biochem Pharmacol. 69(2): 199–207.

Rappe, C. 1996. Sources and environmental concentrations of dioxins and related compounds, Pure and Applied Chemistry 68 (9): 1781–1789.

RIKZ. 2006. RIKZ/SPECIE-07; Determination of 2,3,7,8-TCDD Toxic Equivalents (TCDD-TEQs) in sediments with the DR-CALUX a bioassay by BDS. RIKZ, Middelburg, The Netherlands.

Sanctorum, H., Windal, I., Hanot, V., Goeyens, L., and Baeyens, W. 2007. Dioxin and dioxin-like activity in sediments of the Belgian Coastal area (Southern North Sea). Archives of Env. Cont. and Tox. 52: 317–325.

Schipper C.A., Joost Lahr, Stephen G. George, Peter-Diedrich Hansen, Helena C. da Silva de Assis, Ron van der Oost, John E. Thain, Dave R. Livingstone, Carys L. Mitchelmore, Frederik-Jan van Schooten, Freek Ariese, Albertinka J. Murk, Guy C.M. Grinwis, Hans J.C. Klamer, Belinda J. Kater, Jaap F. Postma, Paul J. van den Brink, D.C. Bert van der Werf, A. Dick Vethaak. 2009. Application of Biomarkers in Flounder (*Platichthys flesus*) and Bioassays in contaminated estuarine sediments, 66, ICES Journal of Marine Science, 66 (10): 2089–2105.

Schipper, C.A., R.M. Burger, R.M., Rietjens, I.M.C.M., Murk, A.T. 2010. Application of bioassays in toxicological hazard, risk and impact assessments of dredged sediments. Marine Pollution Bulletin 60: 2026–2042.

Schipper, C.A., Leonards, P.E.G., Klamer, H.J.C., Thomas, K.V., Vethaak, A.D. 2011. Biological effects of contaminants: Receptor H4IIE-Luciferase (DR-LUC) cell bioassay for screening of dioxins and/or dioxin-like compounds in environmental samples, ICES TIMES (submitted).

Schoeters, G., Legler, J., Brouwer, A. 2009. Intra- and interlaboratory calibration of the DR-CALUX® bioassay Bradfield CA. The aryl hydrocarbon receptor: a perspective on potential roles in the immune system. Immunology. 2009;127(3):299–311.

Stevens EA, Mezrich JD, Bradfield CA. 2009. The aryl hydrocarbon receptor: a perspective on potential roles in the immune system. Immunology.127(3):299–311.

Stronkhorst J., Leonards P., and Murk A.J. 2002. Using the dioxin receptor-CALUX *in vitro* bioassay to screen marine harbour sediments for compound with a dioxin like-mode of action, Environmental Toxicology and Chemistry 21: 2552–2561.

Stronkhorst, J., Ariese, F., van Hattum, B., Postma, J.F., de Kluijver, M., Den Besten, P.J., Bergman, M.J.N., Daan, R., Murk, A.J. and Vethaak, A.D. 2003. Environmental impact and recovery at two dumping sites for dredged material in the North Sea. Environ. Pollut. 124: 17–31.

Thain, J.E., Hurst, M.R. and Thomas, K.V., 2006. Determination of dioxin-like activity in sediments from the East Shetland basin. Organohal. Cpds. 68, 185-188.

Van den Berg, M. Birnbaum, L., Bosveld, A.T., Brunstrom, B., Cook, P., Feeley, M., *et al*. 1998. Toxic equivalency factors (TEFs) for PCBs, PCDDs, PCDFs for humans and wildlife Environ. Health Perspect. 106: 775–792.

Van den Brink, P.J. and Kater, B.J. 2006. Chemical and biological evaluation of sediments from the Wadden Sea, The Netherlands, Ecotoxicology 15: 451–460.

Van Loco, J., Van Leeuwen, S.P.J., Roos, P., Carbonnelle, S., de Boer, J., Goeyens, L., Beernaert, H. 2004. The international validation of bio- and chemical- analytical screening methods for dioxins and dioxin-like PCBs: the DIFFERENCE project rounds 1 and 2 TALANTA 63 (5): 1169–1182.

Vos, J.G., Dybing, E., Greim, H.A., Ladefoged, O., Lambré, C., Tarazona, J.V., Brandt, I. and Vethaak, A.D. 2000. Health Effects of Endocrine-Disrupting Chemicals on Wildlife, with Special Reference to the European Situation. Critical Reviews in Toxicology 30(1): 71–133.

WFD CIS. 2003. Guidance document No. 3; Analysis of Pressures and Impacts in accordance with the Water Framework Directive. www.circa.eu (April 2009).

Windal, I., van Wouwe, N., Eppe, G., Xhrouet, C., Debacker, V., Baeyens, W., de Pauw, E., Goeyens, L. 2002. Validation and interpretation of CALUX as a tool for the estimation of dioxin-like activity in marine detection of novel dioxin sources and new dioxin-like compounds. Environ Sci & Pollut Res 9: 304–306.

World Health Organization. 2000. Interlaboratory Quality Assessment of Levels of PCBs, PCDDs, and PCDFs in Human Milk and Blood Plasma; fourth round of WHO-coordinated study. WHO, Bilthoven.

Annex 16: Background document: Metallothionein (MT) in blue mussels (*Mytilus edulis*, *Mytilus galloprovincialis*)

Ketil Hylland, Norwegian Institute for Water Research, Department of Biology, University of Oslo.

Introduction

Metallothionein (MT) is a low-molecular-weight, cystein-rich protein, metal-binding protein that is found in all vertebrates and most invertebrates. The natural functions of different isoforms of the protein are under discussion and probably vary between species and for tissues within a species. Most forms are involved in metal-sequestration, thereby possibly:

i) regulating cellular processes requiring Zn and/or Cu; and

ii) binding and thus temporarily detoxifying non-essential elements such as Cd and Hg.

In addition, MT has been suggested to be involved in the cellular defence against free radicals (mainly due to the large number of SH-groups). Most of the data available are for liver or hepatopancreas, but there are also some data for gills in both fish and mussels.

In marine fish species, MT concentration in tissues has been found to be most strongly associated with Zn and Cu levels, although Cd may also result in minor increases in areas with metal stress (Hylland *et al*., 2009). Since tissue requirements, and hence concentrations, of essential elements such as Zn and Cu will also be affected by exposure to other contaminants, interpretation of MT in fish as a simple biomarker for metal stress has not been straightforward except in areas with exceptionally high metal levels (predominantly freshwater).

MT in marine invertebrates, particularly mussels, was reviewed recently (Amiard *et al*., 2006). Two main forms of the protein have been identified in blue mussel species, MT-10 and MT-20 (the names reflecting their approximate molecular size). There are a number of genes encoding MT-10 and fewer encoding MT-20 mRNA in *Mytilus edulis* and *Mytilus galloprovincialis* (reviewed in Aceto *et al*., 2011). Gene transcripts of MT-10 and MT-10 intronless genes are orders of magnitude higher than MT-20 under normal metabolism, but the relative increase in MT-20 gene expression under conditions of metal stress is very much higher than that of MT-10 isoforms (Aceto *et al*., 2011).

Methods for quantification

Three main protocols have been used to quantify metallothionein in mussel tissues:

i) the electrochemical differential pulse polarography method (DPP; Olafson and Thompson, 1974);
ii) metal-substitution; and
iii) the spectrophotometric sulphydryl method (Viarengo *et al*., 1997).

In addition, an immunochemical assay has been described, but this has not been used to any extent (Roesijadi *et al*., 1988). The three former methods rely on the content of sulphydryl-groups (SH-groups) in MT and its small size. There has been an international intercalibration of method (iii) through MEDPOL (Viarengo *et al*., 2000) and of fish MT using all three methods within the BEQUALM framework (Hylland, unpublished). Unfortunately, the three methods do not yield the same values when applied to identical samples. Method (i) appears to provide the most reliable values and is the method that has been most extensively validated; method (ii) is sensitive to the affinity of different metals for MT. Cu bound to MT under normal conditions has high affinity and must either be replaced by a metal with even higher affinity, e.g. Ag or Hg, or displaced prior to incubation with e.g. Cd. Method (iii) gives different results to the other methods, resulting in either over- or underestimation. None of the methods are able to separate between MT-10 and MT-20.

Although MT isoforms are thought to be predominantly cytosolic, they have been shown to be present in the nucleus in blue mussels, presumably as part of a regulatory function (Castillo *et al*., 2008). The quantification methods currently used will mainly include cytosolic MT (nuclei will be excluded in the first separation of the work-up process), but this is not thought to be problematic as the total amount in the cell will anyway be dominated by MT present in the cytosol.

An increasing number of studies have quantified mRNA for MT-10 and/or MT-20 (Dondero *et al*., 2005). There appears to be a large increase in MT-20 following metal stress under controlled experimental conditions, whereas increases in MT-10 are less dramatic (Zorita *et al*., 2007a). Similar results have been found in field studies (Aceto *et al*., 2011). MT-20 appears to be more resistant to oxidative stress than does MT-10 (Vergani *et al*., 2007). mRNA is a much more transient response than protein levels, however (as measured by the methods presented above), and there is a need for more knowledge of response dynamics prior to applying the method in a monitoring context.

MT in tissues is most commonly expressed on either a wet weight or dry weight basis (back-calculated), but some authors also express it on the basis of cytosolic protein (the common standard for fish MT). Appropriate factors can be applied to convert from one basis to another, albeit introducing some error.

Concentrations in reference areas

A range of studies have quantified MT using differential pulse polarography in whole mussel, hepatopancreas and/or gill in *M. edulis* (Table 1) or *M. galloprovincialis* (Table 2). A smaller number of studies have been using the sulphydryl method (Table 3). Early analyses using metal-substitution assays will have underestimated MT and have not been included in this overview.

Table 1. Mean concentrations of MT in different tissues of *Mytilus edulis*; expanded from Amiard *et al.* (2006). Some values were read off figures. Values reported on a dry weight basis were recalculated to wet weight using a factor 0.8 (water content; see e.g. Williams, 1970) and from protein-standardized values using a factor 0.08 (assuming 2/3 cytosolic protein and a protein content of 60% of dry wt; Dare & Edwards, 1975).

tissue	original value	factor	MT (µg/g ww)	reference
Whole animal	2.43	0.2	0.49	Bebianno and Langston (1989)
	2.75	0.2	0.55	Bebianno and Langston (1991)
	0.55	1	0.55	Amiard-Triquet *et al.* (1998)
	0.55	1	0.55	Amiard *et al.* (2008)
	0.35	1	0.35	Amiard *et al.* (2008)
Digestive gland	2.25	1	2.25	Amiard *et al.* (1998)
	8.04	0.2	1.61	Bebianno and Langston (1989)
	8	0.2	1.6	Bebianno and Langston (1991)
	8.8	0.2	1.76	Amiard-Triquet *et al.* (1998)
	1.8	1	1.8	Pellerin and Amiard (2009)
	1.6	1	1.6	Geffard *et al.* (2005)
Gills	0.3	1	0.3	Amiard *et al.* (1998)
	2.2	0.2	0.44	Bebianno and Langston (1991)
	1.7	0.2	0.34	Amiard-Triquet *et al.* (1998)
	8	0.08	0.63	Geret *et al.* (2002)
	0.23	1	0.23	Geffard *et al.* (2005)

Table 2. Mean concentrations of MT in different tissues of *Mytilus galloprovincialis*; expanded from Amiard *et al*. (2006). Some values were read off figures. Values reported on a dry weight basis were recalculated to wet weight using a factor 0.8 (water content; see e.g. Williams, 1970) and from protein-standardized values using a factor 0.08 (assuming 2/3 cytosolic protein and a protein content of 60% of dry wt; Dare and Edwards, 1975).

tissue	original	factor	MT (µg/g ww)	reference
Whole animal	12.1	0.2	2.4	Bebianno and Machado (1997)
	1.21	1	1.21	Raspor *et al*. (1999)
	3.21	0.2	0.64	Bebianno and Langston (1992)
	0.5	1	0.5	Mourgaud *et al*. (2002)
Digestive gland	4.09	1	4.09	Raspor *et al*. (1999a)
	2.1	1	2.1	Pavicic *et al*. (1993)
	45	0.08	3.56	Zorita *et al*. (2007)
Gills	0.62	1	0.62	Raspor *et al*. (1999)
	2.35	0.2	0.47	Bebianno *et al*. (1998)

Table 3. Mean concentrations of MT in different tissues of *Mytilus edulis and M. galloprovincialis*. Some values were read off figures. Values reported on a dry weight basis were recalculated to wet weight using a factor 0.8 (water content; see e.g. Williams, 1970) and from protein-standardized values using a factor 0.08 (assuming 2/3 cytosolic protein and a protein content of 60% of dry wt; Dare and Edwards, 1975).

tissue	original	factor	MT (µg/g ww)	reference
M. edulis				
Whole animal	0.04	1	0.04	Brown *et al*. (2004)
Digestive gland	0.11	1	0.11	da Ros *et al*. (2007)
	0.16	1	0.16	Schiedek *et al*. (2006)
M. galloprovincialis				
Whole animal	20	0.08	1.6	Funes *et al*. (2006)
	0.45	1	0.45	Domouthsidou *et al*. (2004)
Digestive gland	0.3	1	0.3	Viarengo *et al*. (2000)
	0.45	1	0.45	Domouthsidou *et al*. (2004)
	0.15	1	0.15	Donnini *et al*. (2008)
Gills	40	1	40	Hamer *et al*. (2008)

Confounding factors

Some studies indicate seasonal variation in MT in mussels with large changes during the spawning period and lower concentrations of the protein, but more stable values in the rest of the year (Geffard *et al*., 2005; Raspor *et al*., 2004; Zorita *et al*., 2007b).

However, other studies have found higher values in autumn (Pellerin and Amiard, 2009). This may be due to different periods of spawning and/or species differences; *M. galloprovincialis* was used in the Mediterranean and *M. edulis* on the French Atlantic coast. A recent study has indicated that *M. galloprovincialis* dominates the Mediterranean/Iberian peninsula and *M. edulis* the French coast, but that there are mixed populations of the two and *M. trossulus* in some areas of northern Europe (Kijewski *et al.*, 2011).

All available data clearly show that there is a strong seasonal dynamic in tissue metal concentration and metallothionein in blue mussels. There appear to be differences between the two species, possibly associated with different spawning periods.

Assessment criteria

The medians or averages from different studies with the three tissues were remarkably similar for *M. edulis*; provisional Background Assessment Criteria (BACs) were constructed using the 90 percentile of averages/medians from literature: Whole body: 0.6 µg/g ww; digestive gland 2.0 µg/g ww and gills 0.6 µg/g ww. These values comprise medians for a full seasonal cycle.

BACs for *M. galloprovincialis* generated in a similar way were: Whole body: 2.0 µg/g ww; digestive gland 3.9 µg/g ww and gills 0.6 µg/g ww. As above, the values are medians for a seasonal cycle.

MT concentrations measured using the sulphydryl method produced results very different to those found using differential pulse polarography; no assessment criteria have been established for this method.

References

Aceto, S. *et al*. 2011. The metallothionein genes of *Mytilus galloprovincialis*: Genomic organization, tissue expression and evolution. *Marine Genomics*, **4**, 61–68.

Amiard, J.-C., Amiard-Triquet, C., Barka, S., Pellerin, J., Rainbow, P.S. 2006. Metallothioneins in aquatic invertebrates: their role in metal detoxification and their use as biomarkers. Aquat. Toxicol. 76, 160–202.

Amiard, J.C., Geffard, A., Amiard-Triquet, C. 1998. La metallothioneine chez la moule *Mytilus edulis* comme biomarqueur de pollution metallique: variabilite entre sites, saisons et organes. J. Rech. Oceanogr. 23, 25–30.

Amiard-Triquet, C., Rainglet, F., Larroux, C., Regoli, F., Hummel, H. 1998. Metallothionein in Arctic bivalves. Ecotox. Environ. Saf. 41, 96–102.

Amiard, J.-C. *et al*. 2008. Influence of field and experimental exposure of mussels (*Mytilus* sp.) to nickel and vanadium on metallothionein concentration. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology*, **147**, 378–385.

Bebianno, M.J. *et al*. 1992. Cadmium induction of metallothionein synthesis in *Mytilus galloprovincialis*. *Comp.Biochem.Physiol.*, **103C**, 79–85.

Bebianno, M.J. *et al*. 1991. Metallothionein induction in *Mytilus edulis* exposed to cadmium. *Mar.Biol.*, **108**, 91–96.

Bebianno, M.J., Langston, W.J. 1989. Quantification of metallothionein in marine invertebrates using differential pulse polarography. Portugal Electrochim. Acta 7, 59–64.

Bebianno, M.J., Langston, W.J. 1992. Cadmium induction of metallothionein synthesis in *Mytilus galloprovincialis*. Comp. Biochem. Physiol. C 103, 79–85.

Bebianno, M.J., Langston, W.J. 1998. Cadmium and metallothionein turnover in different tissues of the gastropod *Littorina littorea*. Talanta 46, 301–313.

Bebianno, M.J., Machado, L.M. 1997. Concentrations of metals and metallothioneins in *Mytilus galloprovincialis* along the south coast of Portugal. Mar. Pollut. Bull. 34, 666–671.

Brown, R.J., Galloway, T.S., Lowe, D., Browne, M.A., Dissanayake, A., Jones, M.B., Depledge, M.H. 2004. Differential sensitivity of three marine invertebrates to copper assessed using multiple biomarkers. Aquat. Toxicol. 66, 267–278.

Castillo, E. del *et al*. 2008. Nuclear and cytosolic distribution of metallothionein in the blue mussel *Mytilus edulis* L. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, **151**, 46–51.

Da Ros, L. *et al*. 2007. Lysosomal responses and metallothionein induction in the blue mussel *Mytilus edulis* from the southwest coast of Iceland. *Environment International*, **33**, 362–369.

Dare, P.J. *et al*. 1975. Seasonal changes in flesh weight and biochemical composition of mussels (*Mytilus edulis* L.) in the Conwy Estuary, North Wales. *Journal of Experimental Marine Biology and Ecology*, **18**, 89–97.

Domouhtsidou, G.P. *et al*. 2004. Lysosomal membrane stability and metallothionein content in *Mytilus galloprovincialis* (L.), as biomarkers: Combination with trace metal concentrations. *Marine Pollution Bulletin*, **48**, 572–586.

Dondero, F. *et al*. 2005. Quantitative PCR analysis of two molluscan metallothionein genes unveils differential expression and regulation. *Gene*, 345, 259–270.

Donnini, F. *et al*. 2007. A biological and geochemical integrated approach to assess the environmental quality of a coastal lagoon (Ravenna, Italy). *Environment International*, **33**, 919–928.

Funes, V. *et al*. 2006. Ecotoxicological effects of metal pollution in two mollusc species from the Spanish South Atlantic littoral. *Environmental Pollution*, **139**, 214–223.

Geffard, A. *et al*. 2005. Do seasonal changes affect metallothionein induction by metals in mussels, *Mytilus edulis*? *Ecotoxicology and Environmental Safety*, **61**, 209–220.

Géret, F. *et al*. 2002. Influence of metal exposure on metallothionein synthesis and lipid peroxidation in two bivalve molluscs: the oyster (*Crassostrea gigas*) and the mussel (*Mytilus edulis*). *Aquat. Living Resour.*, **15**, 61–66.

Hamer, B. *et al*. 2008. Effect of hypoosmotic stress by low salinity acclimation of Mediterranean mussels *Mytilus galloprovincialis* on biological parameters used for pollution assessment. *Aquatic Toxicology*, **89**, 137–151.

Ivankovic, D. *et al*. 2005. Evaluation of the *Mytilus galloprovincialis* Lam. digestive gland metallothionein as a biomarker in a long-term field study: Seasonal and spatial variability. *Marine Pollution Bulletin*, **50**, 1303–1313.

Kijewski, T. *et al*. 2011. Distribution of *Mytilus* taxa in European coastal areas as inferred from molecular markers. *Journal of Sea Research*, **65**, 224–234.

Leignel, V. *et al*. undefined. Isolation and characterization of *Mytilus edulis* metallothionein genes. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology*, **142**, 12–18.

Mackay, E.A. *et al*. 1990. Polymorphism of cadmium-induced mussel metallothionein. *Experientia*, **46**, A36.

Mourgaud, Y. *et al*. 2011. Metallothionein concentration in the mussel *Mytilus galloprovincialis* as a biomarker of response to metal contamination: validation in the field. *Biomarkers*, **7**, 479–490.

Olafson, R.W. *et al*. 1974. Isolation of Heavy Metal Binding Proteins from Marine Vertebrates. *Mar.Biol.*, **28**, 83–86.

Pavicic, J. *et al*. 1989. Determination of Cd-binding proteins similar to metallothioneins in the digestive gland of *Mytilus galloprovincialis* in relation to the preliminary treatment of the sample. *Period.biol.*, **911**, 213–224.

Pavicic, J., Raspor, B., Marticic, D. 1993. Quantitative determination of metallothionein-like proteins in mussels. Methodological approach and field evaluation. Mar. Biol. 115, 435–444.

Pellerin, J. *et al*. 2009. Comparison of bioaccumulation of metals and induction of metallothioneins in two marine bivalves (*Mytilus edulis* and *Mya arenaria*). *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology*, **150,** 186–195.

Raspor, B., Pavicic, J., Kozar, S., Kwokal, Z., Paic, M., Odzak, N., Ujevic, I., Kjakovic, Z. 1999. Assessment of metal exposure of marine edible mussels by means of a biomarker. In: Klaassen, C. (Ed.), Metallothionein IV. Birkhauser, Basel, pp. 629–632.

Raspor, B., Dragun, Z., Erk, M., Ivankovic, D., Pavicic, J. 2004. Is the digestive gland of *Mytilus galloprovincialis* a tissue of choice for estimating cadmium exposure by means of metallothioneins? Sci. Total Environ. 333, 99–108.

Roesijadi, G. *et al*. 1988. Immunochemical quantification of metallothioneins of a marine mollusc. *Can.J.Fish.aquat.Sci.*, **45**, 1257–1263.

Schiedek, D. *et al*. 2006. Biomarker responses as indication of contaminant effects in blue mussel (*Mytilus edulis*) and female eelpout (*Zoarces viviparus*) from the southwestern Baltic Sea. *Marine Pollution Bulletin*, **53**, 387–405.

Soazig, L. *et al*. 2003. Potential use of the levels of the mRNA of a specific metallothionein isoform (MT-20) in mussel (*Mytilus edulis*) as a biomarker of cadmium contamination. *Marine Pollution Bulletin*, **46**, 1450–1455.

Vergani, L. *et al*. 2007. Molecular characterization and function analysis of MT-10 and MT-20 metallothionein isoforms from *Mytilus galloprovincialis*. *Archives of Biochemistry and Biophysics*, **465**, 247–253.

Viarengo, A. *et al*. 1997. A simple spectrophotometric method for metallothionein evaluation in marine organisms: an application to Mediterranean and Antarctic molluscs. *Marine Environmental Research*, **44**, 69–84.

Viarengo, A. *et al*. 2000. Critical evaluation of an intercalibration exercise undertaken in the framework of the MED POL biomonitoring program. *Marine Environmental Research*, **49**, 1–18.

Zorita, I. *et al*. 2007a. Tissue- and cell-specific expression of metallothionein genes in cadmium- and copper-exposed mussels analysed by in situ hybridization and RT-PCR. *Toxicology and Applied Pharmacology*, **220**,186–196.

Zorita, I. *et al*. 2007b. Assessment of biological effects of environmental pollution along the NW Mediterranean Sea using mussels as sentinel organisms. *Environmental Pollution*, **148**, 236–250.

Annex 17: Background Document: Water *in vivo* bioassays

Executive summary

Applicability across the OSPAR maritime area. Water *in vivo* bioassays are available for immediate deployment within the OSPAR JAMP CEMP. These bioassays have been recommended by ICES and are of sufficient standing in terms of methodological development, ease of use and application for uptake across the whole OSPAR area. The preferred method is short-term tests on concentrates of water. This includes both broad-spectrum (acute and short-term chronic) bioassays, (and can be combined with specific *in vitro* bioassays), which can be applied to salt water, brackish water and 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.

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

Quality assurance. QA procedures are in place for most of the (water) bioassays and is provided for by BEQUALM (www.bequalm.org), therefore bioassay data can be submitted to the ICES database for subsequent assessment as appropriate by ICES/OSPAR. A standardized protocol for bioassay extractions is required to ensure consistency of application between laboratories and member states and comparability of reported data for assessment purposes. A protocol for extraction methods for bioassays will become available as ICES TIMES series document in 2011.

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. The background responses for the water bioassays (*Tisbe* sp., *Acartia* sp., sea urchin and bivalve larvae) were 10%, 10%, 10% and 20% mortality (or deformity as appropriate) respectively; the level of serious concern was between two times the background response and 100% mortality, and the warning level between these values.

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

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

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

Recommendations. The sampling strategy and design of water quality monitoring for spatial and temporal monitoring purposes needs to be clearly defined and in particular the role of water concentrates. In this respect there is an important need to develop and validate appropriate protocols for extraction methods and subsequent *in vivo* (and *in vitro*) testing. More research is also needed to link bioassay responses to actual impacts on the aquatic system. The application of passive samplers for bioassay assessment of water also warrants special attention.

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

Water bioassays recommended for use in different monitoring strategies are well described in OECD, ASTM, ISO, SETAC and ICES test protocols (see also USEPA, 1995; Tonkes *et al.*, 2005). Bioassays are widely recognized within Europe to be an efficient way to assess water quality. Bioassays are also applied on national level by several countries (ICES, 2004). The uptake of water bioassays, such as the oyster embryo assay (Thain *et al.*, 1991), in monitoring programmes across the OSPAR maritime area is however still poor (so far, only UK; see ICES, 2004). *In vivo* bioassays (and *in vitro* tests with micro-organisms) are now also frequently used as tools in estimating the potential risk of contaminants of estuarine and marine waters (e.g. Thomas *et al.*, 2002; Murk *et al.*, 2002; Klamer *et al.*, 2003; Akerman *et al.*, 2004).

The standard for bioassays described and proposed is based on a 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 mechanizm.

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, such as DR-Luc/DR-CALUX are laboratory tests using prepared cells or sub cellular fractions isolated from organisms or modified bacteria. These tests are mechanizm-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 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.

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

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

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

Preconditions and criteria for *in vivo and in vitro* 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 must include inter and intra laboratory Quality Assurance procedures. These are provided using agreed international procedures and through BEQUALM and intercalibration exercises. Several further requirements are listed and discussed below. The basic principle is that these tools should allow the ecosystem to be protected as much as possible. The ideal set of bioassays would be representative of all organisms and trophic levels in the ecosystem in question and that the most sensitive species are used. The idea being that the ecosystem as a whole will be protected if a number of 'trigger species' from several taxonomic groups are protected. Furthermore, in such an ideal situation, the response from the set of bioassays should enable all possible substances to be covered, at both the acute and the chronic level. The set should therefore also have the following qualities:

Ecologically and/or toxicologically relevant

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

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

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

Covering all effects of all possible substances and action mechanizms, 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 mechanizms of toxicity and indeed no in *vitro* bioassay that is capable of detecting all substances or possible action mechanizms. The best way to address this issue is to use a set of *in vivo* and *in vitro* bioassays that cover as many different action mechanizms as possible (see also de Zwart and Sterkenburg, 2002). However, some action mechanizms 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 mechanizms 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 are well tried and intercalibrated and as such the inherent variability of the endpoints of each assay is well documented. Therefore, it is possible to design sampling and test strategies with adequate replication to provide good discriminatory power between test samples.

Reliable and reproducible

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

Availability of test species

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

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

Financial

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

Laboratory availability

The introduction of bioassays into the CEMP will place major demands on the available laboratory capacity. This capacity should therefore ideally be expanded. There should preferably be more contract laboratories that can routinely perform bioassays. The bioassays recommended in the JAMP CEMP have well documented protocols and the procedures are easy to learn and in most cases do not require expensive or sophisticated equipment or capital expenditure. Current methods tend to be 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 of course be consistent with the proposed metrics.

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

Towards a normative framework for bioassays

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

It is usually necessary, when conducting rapid, acute *in vivo* tests and *in vitro* 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 confounding or interfering 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 mechanizm, such as substances with an estrogenic effect.

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

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

There are also drawbacks, however. Usually only a proportion of the substances are extracted and the efficiency of the extraction process will depend on the medium and solvent used. Metals, in particular, tend to get left behind in the current procedures. This restricts our view of the total toxicity of the surface water, forcing us to overlook the combined effects of several substance groups with different action mechanizms, 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.

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

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

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

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

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

The OSPAR JAMP CEMP lists water bioassays as Category-II-rated. The corresponding Technical Annexes to 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, and *Skeletonema* 72 hour growth.

Quality assurance through BEQUALM is in place or currently running (JAMP, 1998; ASMO, 2003; ICES, 2005). So far, uptake of water bioassays in BEQUALM has been slow but is increasing. Protocols exist for water extracts, but they have not been agreed, standardized and "transcribed" into OSPAR guidelines. A standardized protocol for bioassay extractions is required to ensure consistency of application between laboratories and member states and comparability of reported data for assessment purposes. Also these protocols are used as standard procedures for BEQUALM intercalibrations. The Protocol for Extraction Methods for Bioassays will be published in the *ICES Techniques in Marine Environmental Sciences* series on Biological Effects of Contaminants (expected publication date: Autumn 2011).

Synergism between CEMP, MSFD and WFD

Though bioassays are not included as ecological quality elements in the monitoring for the Water Framework Directive (WFD) (CIS, 2003), it is generally accepted that they will be able to contribute to investigative monitoring and 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. Further chemical analysis can be combined with water bioassays at smaller interval time points for the purposes of trend monitoring. In this way, bioassays can be used as a partial replacement for chemical analysis of priority and/or other relevant substances and prioritizing locations for further chemical analysis. This "bioanalysis approach" can lead to more cost-efficient and cost-effective monitoring and would put the precautionary principle called for in the WFD into practice. Pilot studies carried out in the Netherlands to explore these possibilities have had promising results (van de Heuvel *et al*., 2005; Maas *et al*., 2005). It can be concluded that clear opportunities exist for synergism between the CEMP or the MSFD and WFD for bioassay applications in coastal and estuarine areas, but that further work and agreement are needed.

Thresholds and assessment tools

General

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

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

Methods for water *in vivo* bioassays currently in JAMP

Water *in vivo* bioassays

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

- Copepod (*Tisbe battaglii* and *Acartia* sp); 48 hour exposure using mortality as the endpoint.
- Bivalves (*Crassostrea gigas, Mytilus* spp) embryos: 24 hour exposure using Percent Net Response as the endpoint.
- Sea urchin (*Paracentrotus lividus*): 24 hour 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 are 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 responses 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 substance (usually a Zn salt). A *reference water* may also be included for site-specific programmes and may be considered as the control water for the sampling area or region under investigation and ideally should give the same response as the control water.

The methodology for the extraction or concentration generally requires sample manipulation and/or concentration techniques, and clean-up using extraction procedures analogous to those used in chemical analysis. These procedures and QA are being developed and documents will be published in the ICES TIMES series.

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

Assessment of background response level of available data for water bioassays

A derivation of background response levels was attempted for the water bioassays using *Tisbe bataglii*, bivalve embryo and echinoderm embryo. Data from controls were collected for several tests from different sources. When individual datasets were obtained, these were averaged per sample and listed in a database with standard deviation. From resulting samples, the averaged per lab/country was calculated together with the 0.1, 0.5 (median) and 0.9 quantile. Where more datasets were available, the same was done with lab/countries datasets. The current assessment thresholds are given in Table 3.

Table 2. Template of data used for calculations of background responses for water bioassays (Median, Min and max are optional).

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

Table 3. Assessment criteria for water *in vivo* bioassays.

Biological Effect	Qualifying comments	Background Response Range	Elevated Response Range	High and Cause for Concern Response
Bioassays; % mortality	Water, copepod	0–10	> 10–< 50	> 50
Bioassays; % abnormality	Water, bivalve embryo	0–20	> 20–< 50	> 50
	Water, sea urchin embryo	0–10	> 10–< 50	> 50
Bioassay; % growth	Water, sea urchin embryo	0–30	> 30–< 50	> 50

Ecotoxicological assessment criteria for *in vivo* and *in vitro* bioassays

This method is available but needs further before it can be implemented.

Assessment framework: metric and criteria

Experience in the Netherlands

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

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

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

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

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

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

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

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

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

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

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

- in sediment, unlike in 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 and Sterkenburg, 2002). Using this technique on water samples makes it easier to scale up the results of *in vivo* bioassays using aquatic organisms to the 'full' metric NE-MPE-SE (so including SE).

Standard for *in vivo* bioassays for surface water

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

Table 4. Details of the metrics for surface water.

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

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

The method works as follows:

- At least four 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.

The MPE on the metric for surface water thus corresponds to the level at which no effect is measured in three *chronic* tests with different taxonomic groups on the 'as is' sample (Cf = 1). On the basis of three *acute* tests the MPE corresponds to the level at which no effect (in practice<EC_{50}) is measured when the sample is concentrated by a factor 10 (Cf = 10) relative to the 'as is' sample. This factor 10 is based on the ACR of 10 (see above).

The 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. Table 5 shows the results of a preliminary effect assessment using the test results of the three bioassays.

Table 5. Indication of toxicity in surface water of the Western Scheldt estuary on basis of three different bioassay responses allowing a preliminary effect assessment as proposed in Maas *et al.*, 2003.

		Cf (ECf50)*			Cf (MTE)
location	date	Daphnia	Algae	Microtox	(from PAF5)
SvOD-1	12-2-2000	42	20	19	
SvOD-2	9-4-2000	28	16	24	
SvOD-3	11-6-2000	54	2.4	23	
SvOD-4	2-8-2000	56	3.5	35	
SvOD-5	17-10-2000	96	4.5	62	
SvOD-6	15-12-2000	87	9	31	
SvOD-1	13-01-2005	95	20	27	
SvOD-2	9-03-2005	87	30	29	
SvOD-3	2-05-2005	127	17	43	
SvOD-4	27-6-2005	197	14	44	
SvOD-5	23-8-2005	251	10	38	
SvOD-6	19-10-2005	94	12	70	
W.Scheldt Vlissingen	4-6-2003	416	52	15	2.0
W Scheldt Honte	4-6-2003	180	56	38	3.2
W Scheldt Terneuzen	4-6-2003	403	28	57	4.0
W Scheldt Hansweert	2-6-2003	243	16	84	17.2
W Scheldt Boei s.v WO3	2-6-2003	271	15	97	3.2
Scheldt Bath	3-6-2003	271	9	52	1.8
Schaar vo Doel (SvoD)	3-6-2003	92	9	50	1.6
Scheldt Antwerpen	18-6-2003	144	2	23	0.4

corrected for recovery

Expected chronic effect in surface water:

green = negligible effect (NE)

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

red = serious effect (SE)

Experience in the UK

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

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

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

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

Conclusions

- Water *in vivo* bioassays are available for immediate deployment within the OSPAR JAMP CEMP. These bioassays have been recommended by ICES and are of sufficient standing in terms of methodological development, ease of use and application for uptake across the whole OSPAR area. Quality assurance procedures are in place for most of the bioassays and are provided for by BEQUALM. Therefore bioassay data can be submitted to the ICES database for subsequent assessment as appropriate by ICES/OSPAR.
- Bioassays should be deployed as a "battery of tests" and should include a minimum basic set, possibly of three or more. However, the composition of what the set needs to comprise of requires further work. The range of bioassays needs to be expanded to include all trophic levels and phyla such as echinoderms.
- The sampling strategy and design of water quality monitoring for spatial and temporal monitoring purposes needs to be clearly defined and in particular the role of water concentrates. In this respect there is an important need to validate appropriate protocols for extraction methods and subsequent *in vivo* and *in vitro* testing.
- Background response levels and assessment criteria for water bioassays currently in JAMP are available.

References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

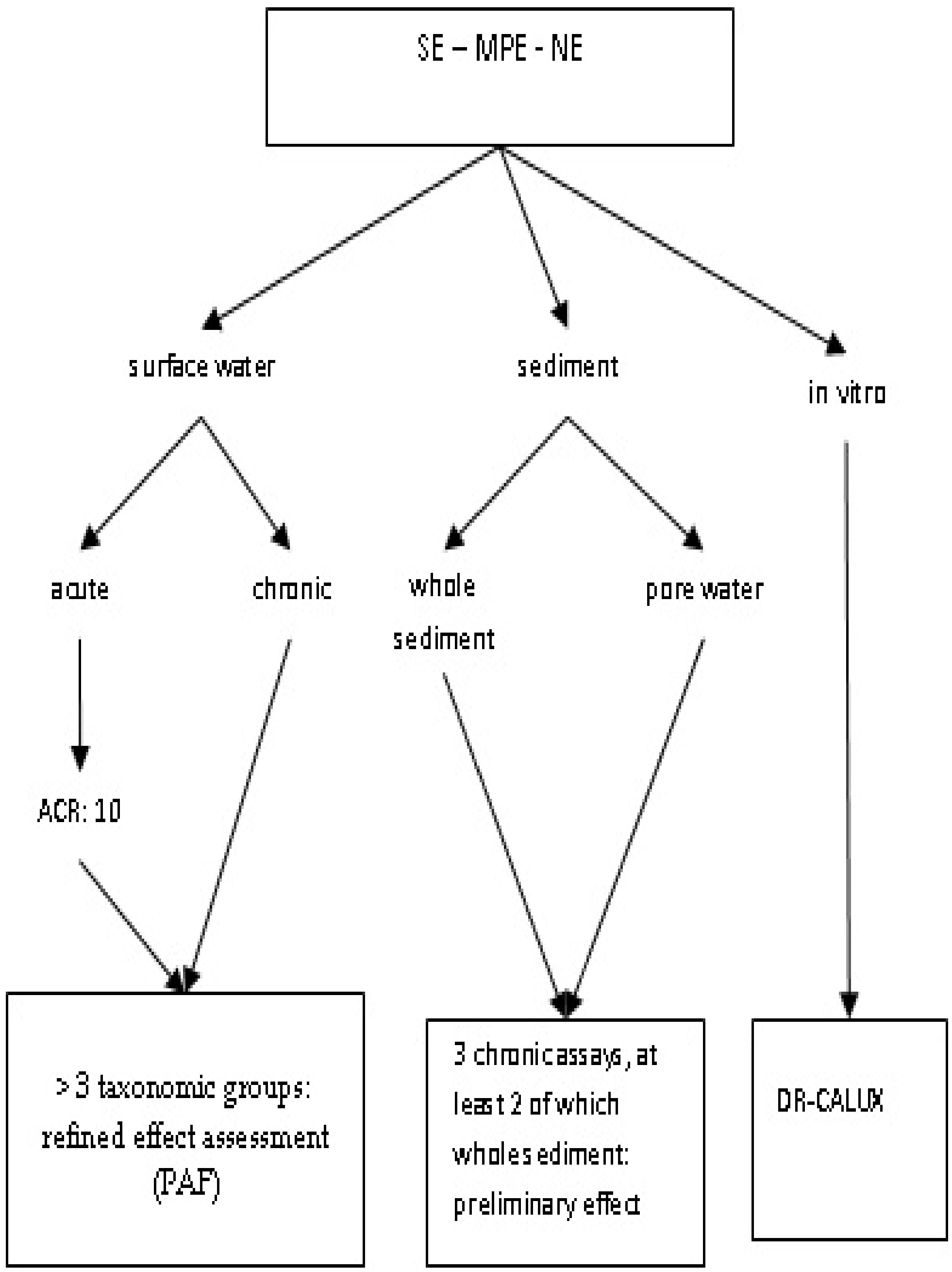


Figure 1. Summary of the metrics based on *in vivo* bioassays for surface water (and sediment, and on *in vitro* bioassays) (ACR: acute-chronic ratio; PAF: potentially affected fraction).

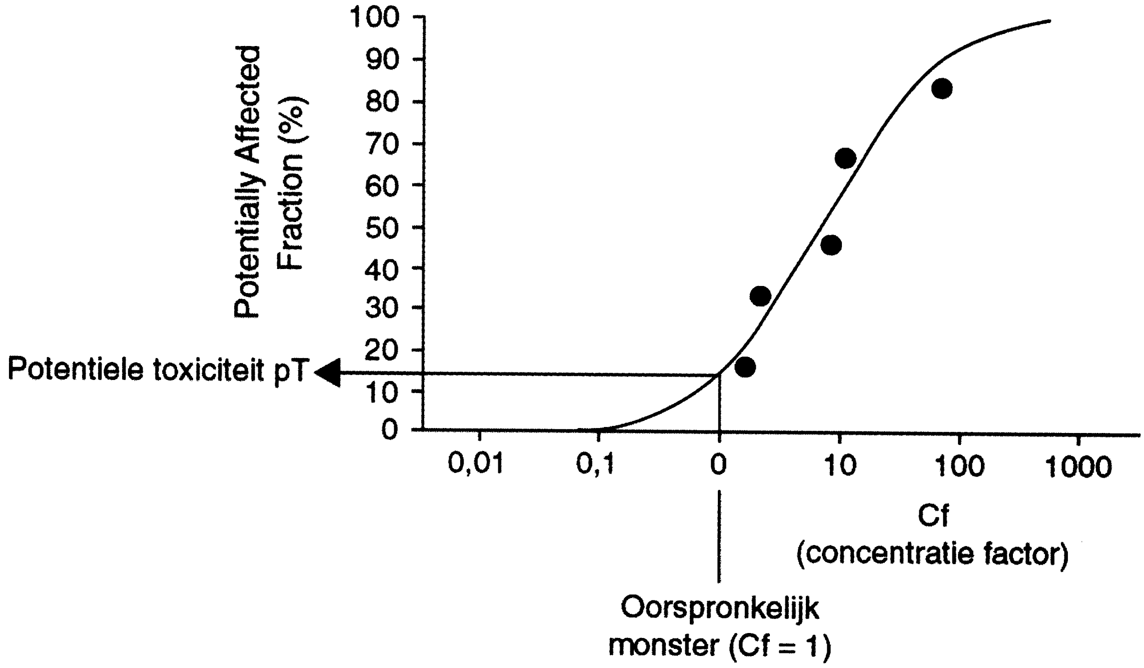


Figure 2. Use of a response curve to estimate the potentially affected fraction (PAF%). (Oorspronkelijk monster = original sample)

Annex 18: Background document (revised): Externally visible fish diseases, macroscopic liver neoplasms and liver histopathology

1.1 Summary

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

Status of quality assurance. Quality assurance procedures for externally visible fish diseases, macroscopic liver neoplasms and liver histopathology are in place and operational through ICES activities and under BEQUALM (www.bequalm.org). 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 ringtests and intercalibration workshops were held. ICES TIMES series publications have been published (nos. 19 and 38).

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

Assessment of thresholds. For externally visible diseases Background Assessment Criteria (BAC) and Environmental Assessment Criteria (EAC) have been defined. For non-specific liver histopathology, BACs have been defined (*EAC awaiting actual ICES data*). Additionally, significant changes in disease prevalence levels and trends serve as 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)(*modification possible, depends on actual ICES data*) (ICES 2009).

Proposals for assessment tools. The WGPDMO developed a Fish Disease Index (FDI) to be used for the analysis and assessment of fish disease data. BAC and EAC have been agreed upon during the 2011 meeting. At the 2009 ICES/OSPAR Workshop on Assessment Criteria for Biological Effects Measurements (WKIMC), assessment criteria

for macroscopic liver neoplasms and for contaminant-specific liver histopathology were proposed.

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

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

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

On an international level, fish disease data have been used for environmental assessments in the framework of the North Sea Task Force and its Quality Status Report (North Sea Task Force, 1993), the OSPAR Quality Status Report 2000 (OSPAR Commission, 2000) and in the 3rd and 4th HELCOM assessments (HELCOM, 1996, 2002). Studies on externally visible diseases, macroscopic liver neoplasms (= liver nodules) and liver histopathology are on the list of techniques for general and contaminant-specific biological effects monitoring as part of the OSPAR pre-CEMP (see Table 2 and Annex 1).

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

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

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

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

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

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

1.3 Status of quality assurance techniques for fish diseases and liver pathology

Since the early 1980s, ICES has played a leading role in the initiation and coordination of fish disease surveys and has contributed considerably to the development of standardized methodologies. Through the work of the ICES Working Group on Pathology and Diseases of Marine Organisms (WGPDMO), its offspring, the Subgroup/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 DataCentre and to data assessment.

A number of practical ICES sea-going workshops on board research vessels were organized by WGPDMO in 1984 (southern North Sea), 1988 (Kattegat), 1994 (Baltic Sea, co-sponsored by the Baltic Marine Biologists, BMB) and 2005 (Baltic Sea) in order to intercalibrate and standardize methodologies for fish disease surveys (Dethlefsen *et al.*, 1986; ICES, 1989, 2006a; Lang and Mellergaard, 1999) and to prepare guidelines. While 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 standardization of methodologies used for studies on liver pathology of flatfish were a major issue of the 1996 ICES Special Meeting on the Use of Liver Pathology of Flatfish for Monitoring Biological Effects of Contaminants (ICES, 1997). This formed the basis from which the BEQUALM programme developed for the application of liver pathology in biological effects monitoring (Feist *et al.*, 2004) (Table 1).

Table 1. BEQUALM categories of histopathological liver lesions in fish that should be used for the CEMP General and PAH-specific Biological Effects Monitoring.

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

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

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

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

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

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

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

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

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

The presence of macroscopic liver neoplasms and of certain types of histopathological liver lesions is a more direct indicator of contaminant effect and has been used for many years in environmental monitoring programmes around the world. Liver neoplasms (either detected macroscopically or by histopathological analysis) are likely to be associated to exposure to carcinogenic contaminants, including PAHs, and are therefore considered appropriate indicators for General and for PAH-specific biological effects monitoring. Therefore, monitoring of macroscopic liver neoplasms in the CEMP should not only be part of the CEMP general biological effects monitoring but also of the CEMP PAH-specific biological effects monitoring. The study of liver histopathology (comprises the detection of more lesion categories (non-specific, neoplastic and non-neoplastic toxicopathic lesions), reflecting responses to a wider range of contaminants (including PAHs) but also to other environmental stressors and is, therefore, considered an appropriate indicator for both General and PAH-specific biological effects monitoring.

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

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

As indicated above, ICES has developed requirements for the international reporting of fish diseases over many years in order to minimize variation between laboratories regarding the accuracy and reproducibility of data generated. These have been reviewed by BEQUALM and produced in CD-ROM format. Each grossly visible disease (lymphocystis, acute and healing skin ulcerations, epidermal hyperplasia/papilloma and liver nodules, etc.) has a minimum number of examined individuals requirement for reporting. 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.

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

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

1.6 Proposals for assessment tools

The development of assessment tools for externally visible diseases, macroscopic neoplasms and liver histopathology has been carried out by the ICES Working Group on Pathology and Diseases of Marine Organisms (WGPDMO) (ICES, 2006b, 2007, 2008, 2009, 2011), Further additions were proposed at the 2009 ICES/OSPAR Workshop on Assessment Criteria for Biological Effects Measurements (WKIMC).

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

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

- Data on diseases of the common dab (*Limanda limanda*) (can be adapted to other fish species, provided that sufficient appropriate data are available);
- Information on the presence or absence of a range of diseases monitored on a regular basis, categorized as externally visible diseases (EVD: nine key diseases, incl. three parasites), macroscopic liver neoplasms (MLN: two key diseases) and liver histopathology (LH: five key diseases) (see Table 2);

- For most diseases, data on three severity grades (reflecting a light, medium or severe disease status) are included;
- Disease-specific weighting factors, reflecting the impact of the diseases on the host (assigned based on expert judgements);
- Adjustment factors for effects of size and sex of the fish as well as for season effects.

Table 2. Disease categories and key diseases to be used for calculating the Fish Disease Index for dab (*Limanda limanda*) (ICES 2009).

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

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. Usually a sample in the present sense consists of the data collected in an ICES statistical rectangle during one cruise. All assessment is based on mean FDI values calculated from these samples. Depending on the data available, FDIs can be calculated either for single disease categories or for combinations thereof.

The assessment of the mean FDI data considers (a) long-term FDI level changes, (b) FDI trends in the recent five years time window and (c) comparing each FDI to its BAC and EAC where these are defined. While assessments (a) and (b) are done on a region-wise basis, global BAC and EAC are used by assessment (c). The assessment approaches (a) and (b) do not apply any global background or reference values or assessment criteria as is often done for chemical contaminants or for biochemical biomarkers. Instead, these assessment approaches use the development of the mean FDI within the geographical units (usually ICES rectangles) 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 implausible 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 five years) for every geographical area to be assessed. The assessment approach (c) ignores the known regional differences and involves globally defined assessment criteria with the consequence that within-region variation might be dominated by general differences in regional levels. However, the FDI can also be used for exploratory monitoring in areas not studied before or for newly installed fish disease monitoring programmes after some modification.

The final products of the assessment procedure are:

- graphs showing the temporal changes in mean FDI values in a geographical unit over the entire observation period; and
- maps in which the geographical units assessed are marked with green, yellow or red smiley faces, indicating long-term changes (e.g. comparing the past five years to the preceding 5-years period) in health status of the fish population (green: improvement of the health status; yellow: indifferent variation; red: worsening of the health status, reason for concern and motivation for further research on causes);
- maps in which the geographical units assessed are marked with green, yellow or red smiley faces, indicating trends in health status of the fish population during the past five years (green: improvement of the health status; yellow: indifferent variation; red: worsening of the health status, reason for concern and motivation for further research on causes);
- maps in which the geographical units assessed are marked with green, yellow or red smiley faces, indicating the level of the FDI for external diseases observed at a defined point in time (green: below the BAC; yellow: between BAC and EAC; red: above the EAC, reason for concern and motivation for further research on causes);
- maps in which the geographical units assessed are marked with green or red smiley faces, indicating the level of the FDI for macroscopic neoplasms observed at a defined point in time (green: below the BAC; red: above the EAC, reason for concern and motivation for further research on causes).

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 DataCentre twice in 2008 and, using an extended dataset, in 2009 (ICES, 2008, 2009). The results will be included in the OSPAR QSR 2010 as a case study.

At the 2009 ICES/OSPAR Workshop on Assessment Criteria for Biological Effects Measurements (WKIMC), additional assessment criteria for macroscopic liver neoplasms and for the contaminant-specific components of liver histopathology were proposed. These are provided in Table 3, which also contains the BAC and EAC for the FDI-EVD, which had been agreed upon at the 2011 meeting of the WGPDMO.

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

Disease category	Background	Elevated response/ above background	Significant response/ unacceptable effects
Externally visible diseases (to be used as additional information for the assessment)	see Table 4	Statistically significant increase in mean FDI level in the assessment period compared to a prior observation period or Statistically significant upward trend in mean FDI level in the assessment period or BAC ≤ FDI level < EAC	Statistically significant increase in mean FDI level in the assessment period compared to a prior observation period or Statistically significant upward trend in mean FDI level in the assessment period or EAC ≤ FDI level
Liver histopathology: non-specific (to be used as additional information for the assessment)	Not applicable	Statistically significant increase in mean FDI level in the assessment period compared to a prior observation period or Statistically significant upward trend in mean FDI level in the assessment period	Statistically significant increase in mean FDI level in the assessment period compared to a prior observation period or Statistically significant upward trend in mean FDI level in the assessment period
Liver histopathology: contaminant-specific	Mean FDI <2	Mean FDI ≥ 2 A value of FDI = 2 is, e. g., reached if the prevalence of liver tumours is 2 % (e. g., one specimen out of a sample of 50 specimens is affected by a liver tumour). Levels of FDI ≥ 2 can be reached if more fish are affected or if combinations of other toxicopathic lesions occur.	Mean FDI ≥ 2 A value of FDI = 2 is, e. g., reached if the prevalence of liver tumours is 2 % (e. g., one specimen out of a sample of 50 specimens is affected by a liver tumour). Levels of FDI ≥ 2 can be reached if more fish are affected or if combinations of other toxicopathic lesions occur.
Macroscopic liver neoplasms	Mean FDI <2	Mean FDI ≥ 2 A value of FDI = 2 is reached if the prevalence of liver tumours (benign or malignant) is 2 % (e. g., one specimen out of a sample of 50 specimens is affected by a liver tumour). If more fish are affected, the value is FDI > 2.	Mean FDI ≥ 2 A value of FDI = 2 is reached if the prevalence of liver tumours (benign or malignant) is 2 % (e. g., one specimen out of a sample of 50 specimens is affected by a liver tumour). If more fish are affected, the value is FDI > 2.

Table 4. Assessment criteria for the assessment of the FDI for externally visible diseases in common dab (*Limanda limanda*). Abbreviations used: Ac, *Acanthochondria cornuta*; Ep, Epidermal hyperplasia/papilloma; Fi, Acute/ healing fin rot/erosion; Hp, Hyperpigmentation; Le, *Lepeophtheirus* sp.; Ly, Lymphocystis; St, *Stephanostomum baccatum*; Ul, Acute / healing skin ulcerations; Xc, X-cell gill disease.

Sex	diseases/ parasites involved in FDI (see legend for abbreviations)	Background Assessment Criteria		Environmental Assessment Criteria	
		ungraded diseases	graded diseases	ungraded diseases	graded diseases
F	Ep, Ly, Ul	1.32	0.216	---	54.0
M	Ep, Ly, Ul	0.96	0.232	---	47.7
F	Ac, Ep, Fi, Hp, Le, Ly, St, Ul, Xc	1.03	0.349	50.6	19.2
M	Ac, Ep, Fi, Hp, Le, Ly, St, Ul, Xc	1.17	0.342	38.8	16.1
F	Ac, Ep, Hp, Le, Ly, St, Ul, Xc	1.09	0.414	48.3	21.9
M	Ac, Ep, Hp, Le, Ly, St, Ul, Xc	1.18	0.398	35.2	16.5

1.7 Final remarks

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

- In the JAMP Guidelines for PAH-specific biological effects monitoring, Chapter 4.1 and 5, the term 'Liver pathology' should be changed to 'Liver histopathology' and the term 'external diseases' should be changed to 'externally visible diseases' because these terms more correctly describe the technique to be applied.
- In the table of contents of the JAMP Guidelines for PAH-specific biological effects monitoring, the terms 'histopathology' and 'liver pathology' should be replaced by 'liver histopathology' because this term more correctly describes the technique to be applied.

1.8 Literature cited

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

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

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

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

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

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

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

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

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

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

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

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

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

ICES. 2009. Report of the ICES/OSPAR Workshop on Assessment Criteria for Biological Effects Measurements (WKIMC), ICES CM 2009/ACOM:50, 133 pp.

ICES. 2011. Report of the Working Group on Pathology and Diseases of Marine Organisms (WGPDMO), ICES CM 2011/SSGHIE:04.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Appendix 1: Fish disease monitoring in the OSPAR Coordinated Environmental Monitoring Programme (CEMP) reflecting ICES advice (ICES 2005)

Table 2a. PAH-specific biological effects monitoring.

	Species	Diseases	Numbers	Guidelines
Macroscopic liver neoplasms	Dab (1st priority) (*Limanda limanda*)	Macroscopic liver nodules > 2 mm in diameter, subsequent quantification of histologically identified liver neoplasms	Size group ≥ 25 cm: 50 (if not available in sufficient numbers, include size group 20–24 cm)	JAMP Guidelines based on: Bucke *et al*., 1996. Common diseases and parasites of fish in the North Atlantic: Training guide for identification. ICES TIMES No. 19. Relevant in addition: ICES 1989. Methodology of fish disease surveys. ICES Coop. Res. Rep. 166. Feist *et al*., 2004. Biological effects of contaminants: Use of liver pathology of the European flatfish dab (*Limanda limanda* L.) and flounder (*Platichthys flesus* L.) for monitoring. ICES TIMES 38, 42 pp. BEQUALM
	Flounder (*Platichthys flesus*)		Size group ≥ 30 cm: 50 (if not available in sufficient numbers, include size group 25–29 cm)	
Liver histopathology	Dab (1st priority) (*Limanda limanda*)	Non-specific lesions Early toxicopathic non-neoplastic lesions Foci of cellular alteration Benign neoplasms Malignant neoplasms	Size group 20–24 cm : 50	JAMP Guidelines based on: ICES 1997. Report of the Special Meeting on the Use of Liver Pathology of Flatfish for Monitoring Biological Effects of Contaminants. ICES CM 1997/F:2. Relevant in addition: Feist *et al*., 2004. Biological effects of contaminants: Use of liver pathology of the European flatfish dab (*Limanda limanda* L.) and flounder (*Platichthys flesus* L.) for monitoring. ICES TIMES 38, 42 pp. BEQUALM
	Flounder (*Platichthys flesus*)		Size group 25–29 cm : 50	
	Dragonet (*Callionymus* spp.)		Size group 10–15 cm : 50	No JAMP guidelines so far. Relevant: Feist *et al*., 2004. Biological effects of contaminants: Use of liver pathology of the European flatfish dab (*Limanda limanda* L.) and flounder (*Platichthys flesus* L.) for monitoring. ICES TIMES 38, 42 pp.

Table 2b. General biological effects monitoring.

	Species	Diseases	Numbers	Guidelines
Externally visible fish diseases	Dab (1st priority) (*Limanda limanda*)	Lymphocystis Epidermal hyperplasia/papilloma Acute/healing skin ulcers X-cell gill disease Hyperpigmentation	Size group 15–19 cm: 100 Size group 20–24 cm: 100 Size group ≥ 25 cm : 50	JAMP Guidelines based on: Bucke *et al*., 1996. Common diseases and parasites of fish in the North Atlantic: Training guide for identification. ICES TIMES No. 19. Relevant in addition: ICES 1989. Methodology of fish disease surveys. ICES Coop. Res. Rep. 166. BEQUALM
	Flounder (*Platichthys flesus*)	Lymphocystis Acute/healing skin ulcers	Size group 20–24 cm: 100 Size group 25–29 cm: 100 Size group ≥ 30 cm: 50	
	Cod (*Gadus morhua*)	Acute/healing skin ulcers Skeletal deformities Pseudobranchial swelling *Cryptocotyle* sp.	Size group < 29 cm: 100 Size group 30–44 cm: 100 Size group ≥ 45 cm: 50	
	Whiting (*Merlangius merlangus*)	Epidermal hyperplasia/papilloma *Lernaeocera branchialis* *Diclidophora merlangi* *Clavella adunca*	Size group 15–19: 100 Size group 20–29: 100 Size group ≥ 30: 50	No JAMP guidelines so far Relevant: Bucke *et al*., 1996. Common diseases and parasites of fish in the North Atlantic: Training guide for identification. ICES TIMES No. 19.
Macroscopic liver neoplasms	Dab (1st priority) (*Limanda limanda*)	Macroscopic liver nodules > 2 mm in diameter, subsequent quantification of histologically identified liver neoplasms	Size group ≥ 25 cm: 50 (if not available in sufficient numbers, include size group 20–24 cm)	JAMP Guidelines based on: Bucke *et al*., 1996. Common diseases and parasites of fish in the North Atlantic: Training guide for identification. ICES TIMES No. 19. Relevant in addition: ICES 1989. Methodology of fish disease surveys. ICES Coop. Res. Rep. 166. Feist *et al*., 2004. Biological effects of contaminants: Use of liver pathology of the European flatfish dab (*Limanda limanda* L.) and flounder (*Platichthys flesus* L.) for monitoring. ICES TIMES 38, 42 pp. BEQUALM

	Species	Diseases	Numbers	Guidelines
	Flounder (*Platichthys flesus*)		Size group ≥ 30 cm: 50 (if not available in sufficient numbers, include size group 25–29 cm)	
Liver histopathology	Dab (1st priority) (*Limanda limanda*)	Non-specific lesions Early toxicopathic non-neoplastic lesions Foci of cellular alteration Benign neoplasms Malignant neoplasms	Size group 20–24 cm: 50	JAMP Guidelines based on: Bucke *et al*., 1996. Common diseases and parasites of fish in the North Atlantic: Training guide for identification. ICES TIMES No. 19. Relevant in addition: ICES 1989. Methodology of fish disease surveys. ICES Coop. Res. Rep. 166. Feist *et al*., 2004. Biological effects of contaminants: Use of liver pathology of the European flatfish dab (*Limanda limanda* L.) and flounder (*Platichthys flesus* L.) for monitoring. ICES TIMES 38, 42 pp. BEQUALM
	Flounder (*Platichthys flesus*)		Size group 25–29 cm: 50	
	Dragonet (*Callionymus* spp.)		Size group 10–15 cm : 50	No JAMP guidelines so far for Dragonet Relevant: Feist *et al*., 2004. Biological effects of contaminants: Use of liver pathology of the European flatfish dab (*Limanda limanda* L.) and flounder (*Platichthys flesus* L.) for monitoring. ICES TIMES 38, 42 pp.

Annex 19: Collation publication resolution

A report on the work of the Study Group on Integrated Monitoring of Contaminants and biological Effects (SGIMC), collating the extensive advice prepared by SGIMC for OSPAR (request 2008/8) on an integrated approach to marine environmental monitoring, to be edited by Ian Davies (UK), and Dick Vethaak (NL) was approved by the Chair of the Advisory Committee, and will be published in the ICES Cooperative Research Report series. The estimated number of pages is 170.

The Study Group on Integrated Monitoring of Contaminants and Biological Effects (SGIMC), agrees to submit the final draft of the proposed publication by 31 October 2011.[3]

Supporting information

Priority:	This has a high priority because of the topicality of the work in relation to Descriptor 8 of MSFD GES, and the extensive package of advice documents that has been prepared. They bring together the work of SGIMC, and of its predecessor WKIMON, both of which had strong links with WGBEC.
Scientific justification:	The forthcoming ICES Cooperative Research Report represents a synthesis of the most recent scientific work on integration of chemical and biological effects monitoring of the sea. SGIMC has had access to authoritative expertise in this are.
Resource requirements:	The material in the report is fairly straightforward, and therefore no specific additional costs are necessary. Some colour images will be required.
Participants:	Approximately two weeks' work is required by the SGIMC editors to finalize the draft.
Secretariat facilities:	About [one month] of the services of Secretariat Professional and General Staff will be required.
Financial:	Cost of production and publication of a 170-page CRR.
Linkages to advisory committees:	This product has been endorsed by ACOM.
Linkages to other committees or groups:	None.
Linkages to other organizations:	EU member states planning their work on MSFD Descriptor 8 will welcome the publication of these document sin a coherent and collated form.

[3] Extension of this deadline can be requested up to one month before the deadline's expiration. If an extension of the deadline is not agreed upon or if the final draft is not forthcoming, the ICES Secretariat will have the option of cancelling the resolution.

Annex 20: Technical annex: Updating Advice from 2010

Annex 9. Technical Annex on sampling and analysis for integrated chemical and biological effects monitoring in fish and shellfish

Introduction

ICES/OSPAR WKIMON and associated groups have progressively developed an integrated approach to the use of biological effects and chemical measurements in environmental monitoring and assessment to meet the objectives of the OSPAR Strategy for Hazardous Substances. In relation to hazardous substances, the OSPAR Joint Assessment and Monitoring Programme seeks to address the following questions:

- What are the concentrations in the marine environment, and the effects, of the substances on the OSPAR List of Chemicals for Priority Action ("priority chemicals")? Are they at, or approaching, background levels for naturally occurring substances and close to zero for manmade substances?
- Are there any problems emerging related to the presence of hazardous substances in the marine environment? In particular, are any unintended/unacceptable biological responses, or unintended/unacceptable levels of such responses, being caused by exposure to hazardous substances?

Integration of chemical and biological effects measurements in OSPAR CEMP

The primary means of addressing these questions on an OSPAR wide basis is the Coordinated Environmental Monitoring Programme (CEMP; OSPAR Agreement 2005, 5). Advice on updated Guidelines for the Integrated Monitoring and Assessment of Contaminants and their Effects were presented by ICES to OSPAR in 2011 in response OSPAR request 2008/8.

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 sediment, fish or shellfish (mussels, gastropods), and overviews of these are included in the Guidelines for the Integrated Monitoring and Assessment of Contaminants and their Effects . 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. More detailed schemes for integrated monitoring are included in the Guidelines for the Integrated Monitoring and Assessment of Contaminants and their Effects, and are reproduced below as Figures 1 and 2.

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.

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, and is the topic of an OSPAR EcoQO. While it is theoretically possible for other substances to disrupt the hormonal systems of snails in a similar way, it is generally accepted that TBT is the primary marine contaminant responsible for the effects.

There is clearly great attraction in the recognition of a highly specific response to a particular narrow class of contaminants, particularly if chemical analysis at concentrations known to be associated with the effects is difficult. However, generally such close relationships are rare. For example, a range of effects measurements have been applied to the effects of planar organic contaminants in the sea, i.e.

- the concentration of PAH-metabolites in fish bile;
- CYP1A/EROD induction;
- Indices of genotoxicity (e.g. DNA adducts of PAH, COMET assay, micronucleus assay, etc);
- liver (microscopic) neoplasms;
- liver histopathology.

However, these effects show varying degrees of specificity for PAH as opposed to other planar organic contaminants such as planar CBs, or dioxins. The concentration of PAH-metabolites in fish bile is clearly specific to the PAH compounds detected, but CYP1A/EROD induction is a property of a range of groups of compounds.

In general, it is found that while subcellular responses can commonly be linked to 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 integrated monitoring frameworks and in Figures 1 and 2 under the headings of subcellular responses, tissues responses and whole organism responses. Subcellular responses such as EROD, bile metabolite concentrations and metallothionein are recognized as biomarkers of exposure to contaminants, while whole organism and tissue level responses are more clearly markers of effect.

Sampling and analysis strategies for integrated fish and bivalve monitoring

The integration of contaminant and biological effects monitoring requires a strategy for sampling and analysis that includes the:

1) sampling and analyses of same tissues and individuals;
2) sampling of individuals for effects and chemical analyses from the same population as that used for disease and/or population structure determination at a common time;
3) sampling of water, the water column and sediments at the same time and location as collecting biota; and

4) more or less simultaneous sampling for and determination of primary and support parameters (e.g. hydrographic parameters) at any given location.

Examples of sampling strategies for the integrated fish and shellfish schemes are shown in Figures 1 and 2. The numbers of individual organisms required are driven primarily by the assessment of external diseases and macroscopic liver nodules (fish) and histopathology (bivalves), because these require the largest number of individuals. A subsample of individuals within the primary sample is further sampled for liver histopathology (fish) and biomarkers (fish and bivalves) to meet Requirements 1 and 2 above.

In the specified target species, further subsampling of the same individuals for chemical analysis is often restricted by insufficient remaining tissue, e.g. liver in fish. In order to meet Requirement 2, subsamples for chemical analysis are taken from the same combined hauls/population as those for disease/biomarkers.

In order to integrate sediment, water chemistry and associated bioassay components, with the fish and bivalve schemes, sediment and water samples should be collected at the same time as fish/bivalve samples and from a site or sites that are representative of the defined station/sampling area.

Additional integrated sampling opportunities may arise from trawl/grab contents, for example, gastropods for imposex or benthos, and these should be exploited where possible/practicable.

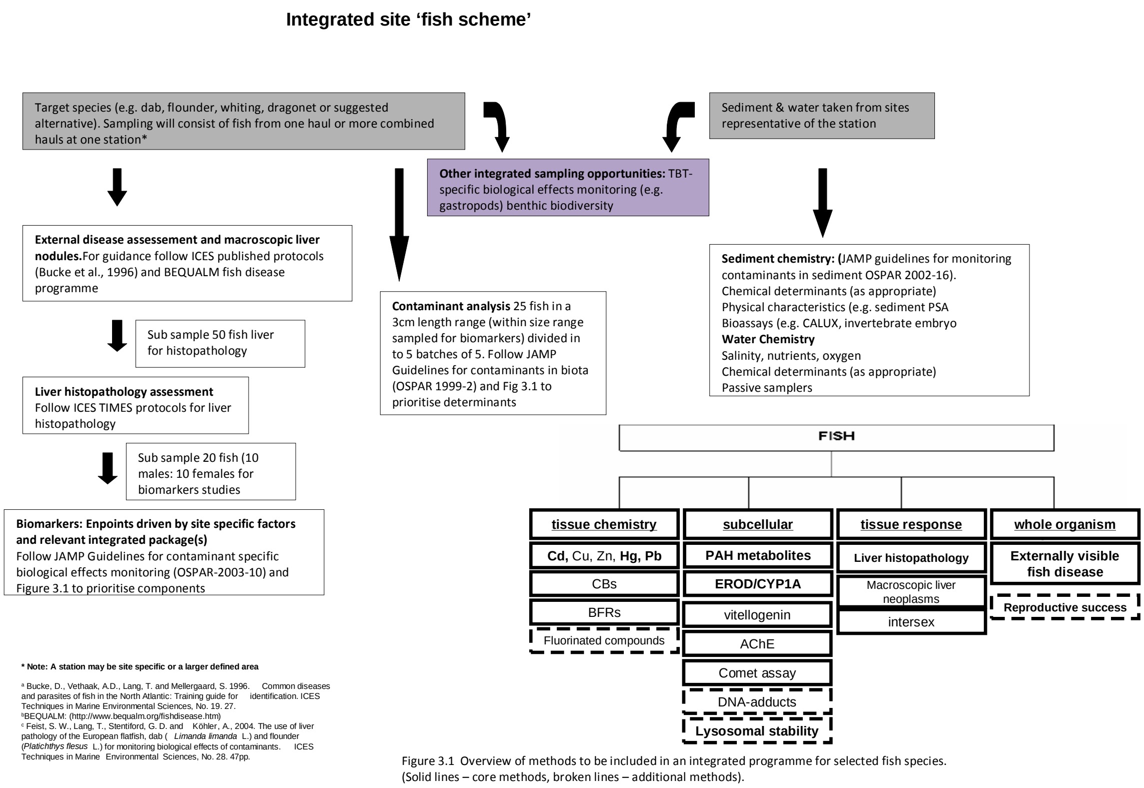


Figure 1. Sampling strategy for integrated fish monitoring.

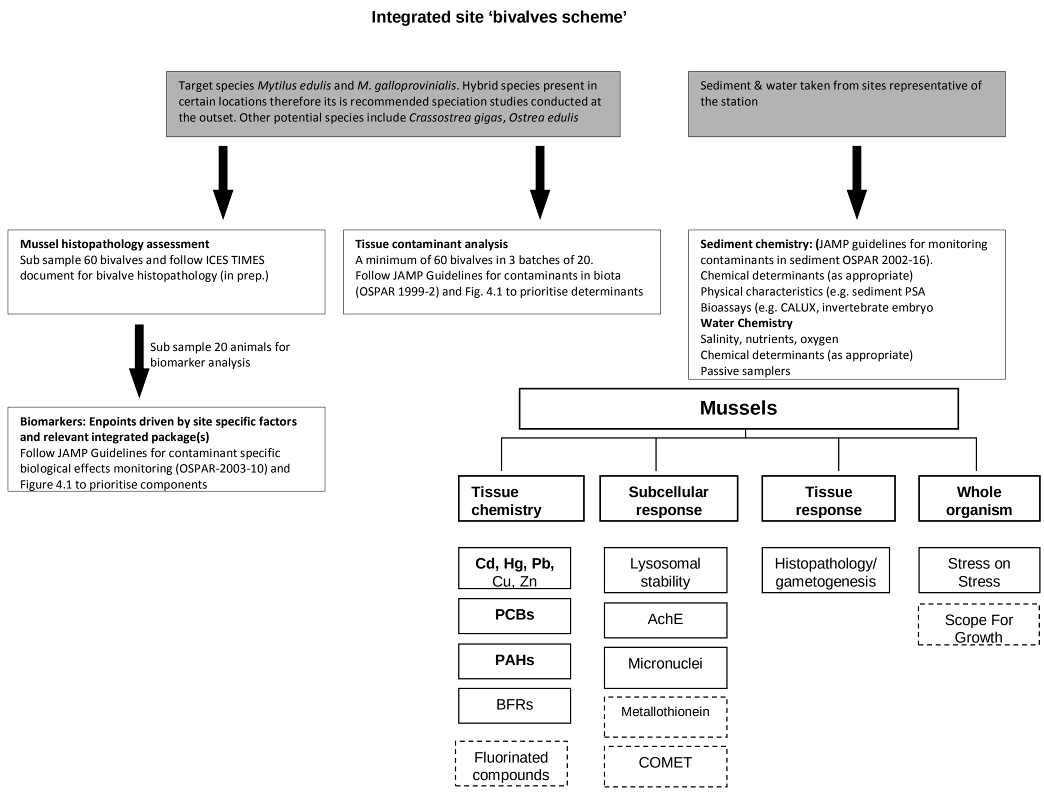


Figure 4.1 Overview of methods to be included in an integrated programme for selected bivalve species. (Solid lines – core methods, broken lines – additional methods).

Figure 2. Sampling strategy for integrated bivalve monitoring.

Guidance on sampling and analysis for integrated monitoring of biological effects and chemical measurements

Some aspects of the details of fish and shellfish sampling and analysis are covered in the OSPAR JAMP Guidelines. Integration of chemical and biological effects data in coordinated monitoring programmes was not a primary consideration when the components of these Guidelines were developed. Some revisions have therefore been made to ensure that the information correctly covers the requirements for integrating chemical and biological effects sampling.

The following tables address aspects of technical guidance on sampling design and supporting parameters.

Tables 1–3 cover methods to be used for integrated fish, bivalve and gastropod monitoring, Tables 4 and 5 cover methods for monitoring of water and sediments.

Table 1. Overview of selected methods for integrated fish monitoring (2007 WKIMON Report, revised).

SUBJECT	PARAMETER	COMMENT
Species	Primary species: dab, flounder, Whiting, eelpout Alternative species: plaice, cod, herring, eelpout, hake, dragonet or other	Alternative species may be used if primary species are not available.
Sex	females and/or males	For certain biomarkers or chemical measurements, only females or only males are used (see relevant JAMP guidelines)
Health condition	Specimens free of external visible diseases should be used for chemical and biomarker analysis.	Certain biomarkers are affected by disease conditions.
Size ranges	Dab: ≥ 15 cm (according to suggested new JAMP guidelines for externally visible diseases). Flounder: ≥ 20 cm (according to suggested new JAMP guidelines for externally visible diseases). Whiting: ≥ 15 cm (according to suggested new JAMP guidelines for externally visible diseases). Dragonet: ≥ 10 cm (according to suggested new JAMP guidelines for liver histopathology). Eelpout: Pregnant females 15–30 cm , 50 fish per station.	For integrated monitoring encompassing chemistry, histopathology and biomarkers, the mid size groups are preferable which are: 20–24 cm (dab) 20–29 cm (flounder)20–24 cm (whiting)10–15 cm (dragonet).
Sample size	Depending on the parameter measured, according to JAMP Guidelines.	Sample sizes have to fulfill statistical requirements for spatial and/or temporal trend monitoring. Preferably, all measurements should be done in individual fish and pooling should be avoided (with the possible exception of contaminant measurements).
Sampling time and frequency	Sampling for all parameters should be carried out at the same time, outside the spawning season, and at least once a year in the same time window	Justification is provided in the OSPAR JAMP Guidelines

SUBJECT	PARAMETER	COMMENT
Sampling location	Sampling for all parameters should be carried out at the same site	The location, size and number of sampling sites depend on the purpose of the monitoring. For offshore sampling targeted at fish, it is recommended to use ICES statistical rectangles as sampling sites. A number of repeated samplings (= hauls) (replicates) should be carried out in each of these rectangles. For coastal and estuarine waters, sites should be selected based on existing WFD and other chemical/biological monitoring sites, taking account of potential hot-spot areas or areas at risk. The number of sampling sites should be sufficient to reflect the environmental conditions ij the survey area, and meet the purposes of the monitoring programme.
Chemical determinands	Metals: Hg, Cd, Pb, Cu, Zn CBs: ICES 7 CBs + CB77, CB81, CB126, CB169 + CB105, CB114, CB123, CB156, CB157, CB167, CB189. Brominated flame retardants: congeners of the penta-mix, octa-mix and deca-mix PBDE formulations; hexabromocyclododecane, tetrabromobisphenol-A. Lindane. TBT	In addition, in situ PAH measurements (e.g.., using UV-fluorescence spectrometry) may be employed under specific circumstances (e.g. after oil spill or PAH-related point source discharges). Besides the contaminants already covered by the OSPAR CEMP, there are a number of other compounds from the OSPAR List of Chemicals for priority action that should be monitored because of their toxicity and environmental relevance. The list provided is, therefore, not complete.
Biological effects measurements	Biological effect techniques as specified in the OSPAR Guidelines for the Integrated Monitoring and Assessment of Contaminants and their Effects, as in Figures 1 and 2 above	Additional opportunities for the inclusion of new methods is likely to emerge through the implementation of MSFD and as science develops. Potential examples include indicators of immunocompetence, and embryo-malformation.
Supporting parameters	Length, weight, gender, age, somatic indices, stage of gonadal maturation, grossly visible anomalies, lesions, parasites, hydrography (temperature, salinity, oxygen content)	In the list, parameters are provided that are known to affect both the biological effects responses and the concentration of contaminants. The data can be of assistance in data interpretation.
Haul duration	Haul durations should be harmonized between monitoring authorities. An appropriate value would be 30 minutes, but may be less than this if conditions require.	The purpose is to standardize the stress experienced by fish during capture
Duration and conditions of storage of live fish prior to dissection	Fish should be maintained alive in flowing seawater on the sampling vessel for periods not exceeding 8 hours.	Storage for longer periods or under poor conditions can stress the fish and alter some biomarker responses.

Table 2. Overview of selected methods for integrated shellfish monitoring (2007 WKIMON Report, revised).

SUBJECT	PARAMETER	COMMENT
Species	Primary species: *Mytilus edulis* Alternative species: *Mytilus galloprovincialis*, *Crassostrea gigas*, *Ostrea edulis*	The first choice shellfish species is not available in all parts of the OSPAR area. In such cases, other species should be selected, such as oysters. For Mytilus sp., speciation studies are recommended in order to confirm species identity.
Sex	Females and/or males	For certain biomarkers or chemical measurements, only females or only males are used (see relevant JAMP guidelines)
Size range	Mussel: >=40 mm, ideally in the range between 40–55mm. Pacific oyster: 9–14 cm	Based on JAMP Guidelines for chemical monitoring
Sample size	Depending on the parameter measured, according to JAMP Guidelines.	Sample sizes have to fulfil statistical requirements for spatial and/or temporal trend monitoring. For some parameters, sample size has still to be defined. Preferably, all measurements should be done in individual mussels and pooling should be avoided (except where recommended, for example for the measurement of contaminant concentrations).
Sampling time and frequency	Sampling for all parameters should be carried out at the same time, outside the spawning season, and at least once a year in the same time window	Justification is provided in the OSPAR JAMP Guidelines
Sampling Location	Sampling for all parameters should be carried out at the same site.	The location, size and number of sampling sites depend on the purpose of the monitoring. For coastal and estuarine waters, sites should be selected based on existing sites used for WFD or other purposes, taking account of hot-spot areas and areas at potential risk. The number of sampling sites should be sufficient to reflect the environmental conditions ij the survey area, and meet the purposes of the monitoring programme. For coastal and offshore studies, caging of mussels should be considered.

SUBJECT	PARAMETER	COMMENT
Chemical determinands	Metals: Hg, Cd, Pb, Cu PAHs: EPA 16 + NPD CBs: ICES 7 + CB 77,81,126,169 + CB 105,114,123,156,157, 167,189 Brominated flame retardants: congeners of the penta-mix, octa-mix and deca-mix PBDE formulations; hexabromocyclododecane, tetrabromobisphenol-A. Lindane Organotin compounds	In addition, total hydrocarbon measurements (e.g.., using UV-fluorescence spectrometry) may be employed under specific circumstances (e.g. after oil spill or PAH-related point source discharges). Besides the contaminants already covered by the OSPAR CEMP, there are a number of other compounds from the OSPAR List of Chemicals for priority action that should be monitored because of their toxicity and environmental relevance. The list provided is not complete.
Biological effects measurements	Biological effect techniques as specified in the OSPAR Guidelines for the Integrated Monitoring and Assessment of Contaminants and their Effects, as in Figures 1 and 2 above	Additional opportunities for the inclusion of new methods is likely to emerge through the implementation of MSFD and as science develops. Potential examples include indicators of immuno-competence, and embryo-malformation.
Supporting parameters	Shell length, shell and soft body weight, gender, stage of gonadal maturation, grossly visible anomalies, lesions, parasites, sampling depth, hydrography (temperature, salinity, oxygen content, turbidity), nutrients/eutrophication	In the list, parameters are provided that are known to affect both the biological effects responses and the concentration of contaminants. The data can be of use for normalization.
Sampling depth	Subtidal or intertidal mussels can be used. Deployed mussels offshore can be positioned at depths 0–8m	Intertidal specimens may be subject to greater biomarker variability. Subtidal specimens are less robust post-sampling and effects measurements may be more susceptible to post-sampling stress.
Storage and transport of bivalves	Transport of bivalves should be completed within than 24 hours. They should be transported in an insulated container at 4oC in a damp atmosphere maintained by absorbent materials (such as seaweed and/or paper towel) wetted with seawater.	

Table 3. Overview of methods and species for integrated gastropod/organotin monitoring (2007 WKIMON Report, revised).

SUBJECT	PARAMETER	COMMENT
Species	Intertidal species: *Nucella lapillus* *Nassarius reticulata* *Littorina littorea* Offshore species: *Buccinum undatum* *Neptunea antiqua*	
Sex	Females and/or males	
Size range	Size ranges are to be selected in accordance with the JAMP Guidelines	
Sample size	Depending on the parameter measured, according to JAMP Guidelines.	All measurements should be done in individual gastropods and pooling should be avoided.
Sampling time and frequency	Sampling for all parameters should be carried out at the same time. Sampling frequency according to JAMP Guidelines.	
Sampling Location	Sampling for all parameters should be carried out at the same site.	For coastal and estuarine waters, sites should be selected based on existing WFD sites (where they are established) and TBT hot-spot areas like harbours and major shipping routes (see relevant JAMP guidelines).
Chemical Determinands	Organotin compounds in tissue	Guidelines for chemical measurements in biota will be published shortly in ICES TIMES series, and in a Technical Annex to the JAMP Guidelines.
Biological effects measurements	Imposex or intersex (species-dependent endpoints, as in the JAMP Guideline) ICES TIMES document on intersex in *Littorina* provides methodological advice.	
Supporting parameters	Shell length, organotin compounds in sediment.	

Table 4. Environmental parameters for inclusion in monitoring programmes (water) (2007 WKI-MON Report, revised).

SUBJECT	PARAMETER	COMMENT
Chemistry	Salinity, nutrients, oxygen	
Chemical determinands	Metals: Hg, Cd, Pb, Cu, Zn PAHs: EPA 16 + Naphthalene, phenanthrene, dibenzothiophene and their alkylated derivatives CBs: ICES 7 CBs Brominated flame retardants: congeners of the penta-mix, octa-mix and deca-mix PBDE formulations; hexabromocyclododecane, tetrabromobisphenol-A. Lindane Organotin compounds	Consideration should be given to bioavailability. To answer the JAMP question relating to concentrations approaching background or zero, there may be a requirement to measure a broader range of chemicals.
Physical	Temperature, content of suspended matter	
Biology	Phyto- and zooplankton	Information might be useful in the case of specific events, such as blooms affecting fish health

Table 5. Environmental parameters for inclusion in monitoring programmes (sediment) (2007 WKIMON Report, revised).

SUBJECT	PARAMETER	COMMENT
Chemistry	TOC, water content, Al, Li	Al and Li (or other elements as appropriate to the sediment type) are used for normalization of contaminant concentrations.
Chemical determinands	Metals: Hg, Cd, Pb, Cu, Zn PAHs: EPA 16 + Naphthalene, phenanthrene, dibenzothiophene and their alkylated derivatives CBs: ICES 7 CBs+ CB77, CB81, CB126, CB169 + CB105, CB114, CB123, CB156, CB157, CB167, CB189. Brominated flame retardants: congeners of the penta-mix, octa-mix and deca-mix PBDE formulations; hexabromocyclododecane, tetrabromobisphenol-A. Lindane Organotin compounds	Consideration should be given to bioavailability. To answer the JAMP question relating to concentrations approaching background or zero, there may be a requirement to measure a broader range of chemicals.
Physical	Sediment type, particle size, colour, index, information on anthropogenic disturbances, sedimentation rates, current flow rates	Anthropogenic disturbance such as trawling or sand and gravel extraction may affect the sediment structure.

Annex 10. Technical Annex for Integrated chemical and biological monitoring of Mussel (*Mytilus* sp.)

Background

The basis for the technical annex is the mussel integrated monitoring strategy incorporating biological effect techniques at the subcellular, tissue and whole organism responses and tissue chemistry. This is outlined below (Figure 1):

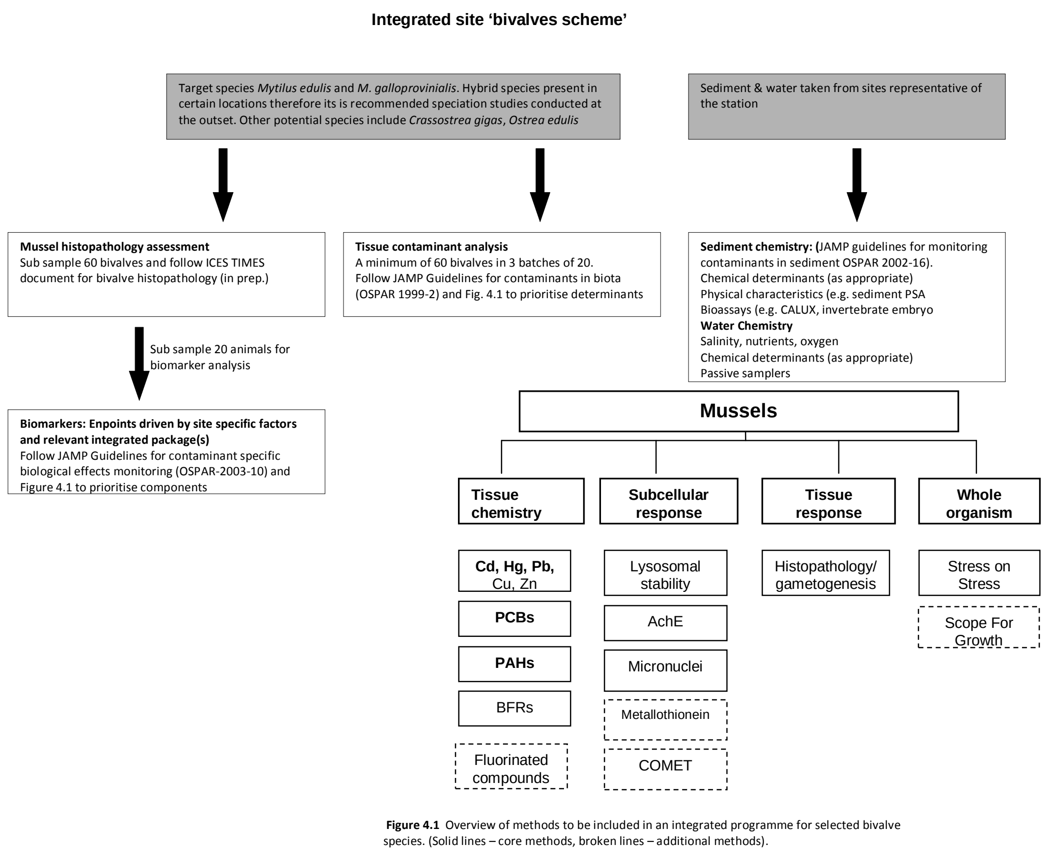


Figure 1. Sampling strategy for integrated bivalve monitoring.

In any mussel integrated monitoring programme the core components as indicated should be included as a bare minimum.

Purpose of work

The integrated approach described above can be used for:

- *Status and trend monitoring*; contaminant and biological effect responses are measured over geographic areas and repeated over time. The purpose here may be to compare biological effect responses between sites, to compare changes in response with time and to observe if the "health status" is improving, at a steady state or declining.
- *Investigative monitoring*; most frequently used as a screening step to assess if biological effects are occurring in relation to a suspected contaminant

gradient, pollution event or if biological effects are suspected for any reason (e.g. tissue chemical residues have been observed to be high).

- *Hot spot – site-specific monitoring*; usually in relation to risk assessment at pollution sites e.g. oil platform investigations.

Offshore and coastal

Mussels (*Mytilus* species) are infrequently found in the sub littoral zone. But populations do exist in shallow waters and are found on the seabed, usually close to the coastline, in general within the 12 mile limit. They may also be found offshore attached to navigation buoys, chains, and oil and gas platforms. For monitoring purposes these mussels can be used but care needs to be exercised in sampling the organisms, to ensure that they are not damaged during sampling and that the correct size range can be obtained. For offshore monitoring purposes it is usually more applicable to use *in situ* caging methods (see below). Advantages of using caged organisms are; choice of site deployment (including reference sites), selection of depth of deployment (e.g. may be critical for oil platform studies, but generally within 8 m of the sea surface); standardization of origin (same source/supply), size and species. Disadvantages are: cost of deployment in respect of mooring systems and ship time for deployment and retrieval; in addition some techniques require immediate sampling and analysis which may not be feasible on a research vessel offshore.

If caging is used then hydrographical conditions must be considered with special attention given to water currents and stratification.

Shoreline

Mussels may be regarded as ubiquitous on rocky shore coastlines and therefore, ideal for monitoring purposes. Sampling sites can be selected easily, organisms collected with little cost and reference sites located without difficulty. In addition, if mussels are not present at a site of interest then organisms can be caged on the seashore or in estuaries on piers or similar structures.

Sampling information

Details required

- Date, time and location on the shoreline (if applicable e.g. low water) and exposure (e.g. highly exposed Atlantic rocky shore or enclosed sheltered bay).
- Position in Lat. Long.
- Type of site; reference, pollution gradient, status or trend.
- At caging sites information on water temperature, depth of deployment, time of immersion, water column depth and information on currents and stratification if available, water temperature and salinity.
- Source of mussels for caging studies; for any caging study it is important that the mussels are sourced from a clean site, and that day 0 values are determined for tissue contaminant chemistry and biological effect responses.
- For shoreline monitoring, ideally the mussels must be sampled in a uniform manner between sites i.e. tidal height and similar salinity profile.

Confounding factors

For *in situ* transplants/caging the mussels must be deployed for at least three weeks in order to allow sufficient time for contaminants to accumulate in the tissues and reach a state of equilibrium. Failure to do this may produce spurious data. Also of note is that in many countries there are regulations controlling the movement and deposit of shellfish and these must be observed (i.e. prevention of transfer of disease).

Reproductive state and gametogenic cycle; Mussels generally spawn in early spring, with spawning occurring later in more northern populations. At spawning there is a major loss in body lipid and a subsequent fall in condition; therefore sampling in or shortly after this period should be avoided for all aspects of tissue chemistry analysis and biological effect determinations.

Salinity; be aware that low salinities affect the biomarker response, of particular importance for caging work in estuaries.

Temperature: mussels on the shoreline can be subject to extremes of temperature, cold in winter and extreme heat in summer. Avoid sampling when extremes are likely to occur as this may compromise the biological effects response.

Parasites; mussels with severe parasite infections should not be used.

Algal blooms: in spring and late summer and autumn intense algal blooms may occur and sampling of mussels at such times should be avoided.

Species; on some coastlines mussels are solely of one species whereas at other locations they are mixed or hybrids. It is unclear whether species difference will affect interpretation of data but wherever possible attempts should be made to determine the species under observation.

In caging studies (shoreline or offshore) care should be taken in sourcing mussels from a "clean site". If rope grown mussels are chosen then particular attention must be given to transporting the mussels as they tend to have weak adductor muscles and easily gape and become stressed during transportation which may give rise to initial mortalities or erroneous biological effect responses. Therefore, the source of mussels should be taken account of in the experimental design.

Supporting measurements

- Condition index; dry meat relative to whole live weight or internal shell volume.
- Gonad state; index of reproductive state.
- Lipid content; usually a determined and measured along with tissue chemistry and useful for interpretation of biomarker responses.
- Real growth; if available measured using growth of marked intervals over time, usually months.
- Water quality measurements; salinity, temperature are recommended, and where possible suspended solids or turbidity, DO, and chlorophyll.
- Chemical analysis of tissues; this is essential to interpretation of biological effects data and for the implementation of the integrated chemical biological effect strategy as outlined above. Prioritized contaminants are Cd, Cu, Hg, Zn, Cd, PAHs and PCBs. As a minimum 50 mussels (>40 mm in length) should be collected, taken to the laboratory and held in running seawater for 24 hours to eliminate gut contents (e.g. sediment, etc). The

tissues should then be extracted from the mussel and placed in acid washed hexane rinsed glass/plastic / metal containers (as appropriate for the particular analysis), stored at -20°C for subsequent chemical analysis using ICES or appropriate protocols.

Sampling for biological effects

For some methods the samples require immediate processing at the time of sampling whereas for other techniques processing is undertaken in the laboratory. An overview of this is shown in the table below (Table 1), and also includes the number of animals typically sampled for each method. Ideally the size of individual mussels for all methods is >40 mm.

Table 1. Overview of sampling procedures for mussels.

Method and minimum numbers of animals usually sampled per site in brackets.	When analytical sampling is undertaken	Acclimation	Comments and aspects that are crucial
SFG (10)	24 hr	Ca 10 hr	Crucial
AChE (10)	Immediate in field	Not applicable	Stored immediately in liquid nitrogen
Mt (10)	Any time within 24 hr on live mussel	Not applicable	Take tissue sample – freeze in liquid nitrogen
COMET	Within 24 hr	Store for no more than 24 hr in cool damp conditions. Must be consistant in startegy	Do as quickly as possible
Micronuclei (20)	Within 3 days	None	Mussels can be kept out of water but cool
NRR (10)	Within 24 hr	Store for no more than 24 hr in cool damp conditions. Must be consistant in startegy	Do as quickly as possible
Lysosomal histochemical method (10)	Freeze immediately	Not applicable	In liquid nitrogen
Stress on stress (40)	Not aplicable	Transport at low temperatures for no more than 24 hr	Analysis done at 18°C
Histopathology and gametogenesis (30–50)	Sample immediately if possible	Anything more than 6 hr delay in sampling place in water for 48 hr acclimation	Dessication must be avoided, correct dissection to incude all organs
Tissue chemistry (50)	Place in 24 hr clean running seawater	Not applicable	Depuration of sediment is crucial

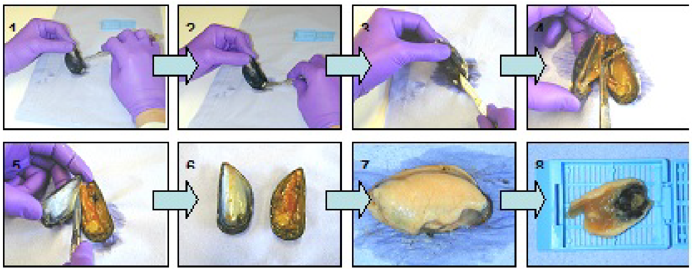
Mussels are attached to each other or to a substratum by a byssal thread. When mussels are sampled care should be used not to pull the mussels and byssal threads too vigorously as this can damage and stress the mussels. If mussels have to be transported this should be kept to a minimum and they should be kept damp and cool and if possible the temperature logged during the transport.

For some techniques such as SFG the mussels will need to be carefully cleaned. It should be noted that there are limitations of analysis for some methods e.g. for SFG and NRR where time-wise it may be difficult to process more than two samples in a single day.

For histological sampling it is essential that the dissection is conducted in a precise manner and this is described below:

The technical procedure essential to correct mussel sampling for histology (taken from draft TIMES doc. under preparation, provided by J Bignell UK, Cefas):

- Insert scalpel into ventral byssal cavity and move knife down so it cuts the posterior adductor muscle.
- Open shell and remove byssal thread.
- Remove mussel from one shell half. Repeat for remaining half.
- Analyse tissue for presence of parasites, pearls or other abnormalities.
- Obtain a standardized section as shown in photographs 1–8 in order to include all organs of interest in one section and place into histo-cassette.



- Samples should be preserved for a minimum of 24 hours in Bakers Formal Calcium, and subsequently transferred to 70% alcohol until processed.
- The correct ratio of mussels to fixative is 30 samples per 800 ml (approx) of fixative. This is the recommended volume of fixative to ensure adequate fixation.
- Samples should be agitated periodically to ensure thorough fixation. A rocker plate facilitates this perfectly.

Methods to be used

These are listed in the mussel integrated strategy above. An overview of the methods is given in the table below (Table 2) with references to the analytical procedures.

Table 2. Overview of methods and reference to analytical procedure.

Method	Issue addressed	Biological significance	References
AChE inhibition	Organophosphates and carbamates or similar molecules Possibly algal toxins	Measures exposure to a wide range of compounds and a marker of stress.	1–2
Metallothionein induction	Measures induction of metallothionein protein by certain metals (e.g. Zn, Cu, Cd, Hg)	Measures exposure and disturbance of copper and zinc metabolism.	3–4
Lysosomal stability (including NRR)	Not contaminant-specific, but responds to a wide variety of xenobiotic contaminants and metals	Measures cellular damage and is a good predictor of pathology. Provides a link between exposure and pathological endpoints. Possibly, a tool for immunosuppression studies in white blood cells.	5–19
Scope for growth	Responds to a wide variety of contaminants	Integrative response, a sensitive sublethal measure of energy available for growth.	20–21
Stress on stress	Responds to a wide variety of contaminants and other environmental conditions	Integrative response , a measure of stress, condition, health and well-being.	26
Micronuclei	Exposure to aneugenic and clastogenic	Exposure to aneugenic and clastogenic	22–23
Histopathology and gametogenesis	Not contaminant-specific	General responses	24–25 ++
COMET	Genotoxic compounds	DNA strand breaks	See OSPAR Background Document

Quality assurance

Wherever possible all analytical methods must be supported with quality assurance procedures. These should be through international intercalibration exercises where they exist and through internal quality controls.

The current position with quality assurance is:

- NRR – currently being developed across OPSAR, exists in MEDPOL, for internal QA a dual assessment with a colleague on the same samples is recommended.
- Ache – not yet developed but include internal standard.
- Mt – MEDPOL have intercalibration exercises, elsewhere there have been *ad hoc* intercalibrations and additionally an internal standard should be included.
- SFG – none at present.
- Stress on Stress – none at present but will be addressed by MEDPOL/ICES workshop in 2010.

- Histology and gametogenesis – TIMES doc and circulation of reference material.
- Lysosomal histochemical procedures – none currently available but include an internal standard. In addition will be addressed by MEDPOL / ICES workshop in 2010.
- Micronuclei formation – currently being addressed through MEDPOL and may be extended to include a wider participation.
- COMET – none at present but being addressed through ICES WGBEC.

Reporting requirements

Biological effect responses; these should be reported in-line with requirements detailed in each analytical method. When different biological effect measurements are made on the same individual mussel then the data should be identified in the reporting and data assessment.

Contaminants: reported in line with standard analytical procedures.

Supporting parameters

Essential; date and time of sampling, Lat. Long. position, organism length, whole weight, site characterization (e.g. position on shore, or caging, DO, salinity, etc); for caged studies the source of organisms and duration of exposure.

Desirable; identification of species particularly if in a hybrid zone.

References

Bocquené, G., Bellanger, C, Cadiou, Y., and Galgani, F. 1995. Joint action of combinations of pollutants on the acetylcholinesterase activity of several marine species. Ecotoxicology, 4: 266–279

Rickwood, C.J., and Galloway, T.S. 2004. Acetylcholinesterase inhibition as a biomarker of adverse effect. A study of *Mytilus edulis* exposed to the priority pollutant chlorfenvinphos. Aquatic Toxicology, 67: 45–56.

Roesijadi, G., Unger, M.E., and Morris, J.E. 1988 Immunochemical quantification of metallothioneins of a marine mollusc. Canadian Journal of Fisheries and Aquatic Sciences, 45: 1257–1263.

Viarengo, A., Ponzano, E., Dondero, F., and Fabbri, R. 1997. A simple spectrophotometric method for metallothionein evaluation in marine organisms: an application to Mediterranean and Antarctic molluscs. Marine Environmental Research, 44: 69–84.

Ringwood, A.H., Conners, D.E., and Keppler, C.J. 1999. Cellular responses of oysters, *Crassostrea virginica,* to metal-contaminated sediments. Marine Environmental Research, 48: 427–437.

Ringwood, A.H., Conners, D.E., and Hoguet, J. 1998. Effects of natural and anthropogenic stressors on lysosomal destabilization in oysters *Crassostrea virginica.* Marine Ecology Progress Series, 166: 163–171.

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

Cheung, V.V., Wedderburn, R.J., and Depledge, M.H. 1998. Molluscan lysosomal responses as a diagnostic tool for detection of a pollution gradient in Tolo Harbour, Hong Kong. Marine Environmental Research, 46: 237–241.

Etxeberria, M., Cajaraville, M.P., and Marigomez, I. 1995. Changes in digestive cell lysosomal structure in mussels as biomarkers of environmental stress in the Urdaibai Estuary (Biscay Coast, Iberian Peninsula). Marine Pollution Bulletin, 30: 599–603.

Fernley, P.W., Moore, M.N., Lowe, D.M., Donkin, P., and Evans, S. 2000. Impact of the *Sea Empress* oil spill on lysosomal stability in mussel blood cells. Marine Environmental Research, 50: 451–455.

Lin, S.Y., and Steichen, D.J. 1994. A method for determining the stability of lysosomal membranes in the digestive cells *of Mytilus edulis.* Marine Ecology Progress Series, 115: 237–241.

Lowe, D.M., and Fossato, V.U. 2000. The influence of environmental contaminants on lysosomal activity in the digestive cells of mussels *(Mytilus galloprovincialis)* from the Venice Lagoon. Aquatic Toxicology, 48: 75–85.

Lowe, D.M., Soverchia, C, and Moore, M.N. 1995. Lysosomal membrane responses in the blood and digestive cells of mussels experimentally exposed to fluoranthene. Aquatic Toxicology, 33: 105–112.

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

Regoli, F. 1992. Lysosomal responses as a sensitive stress index in biomonitoring heavy-metal pollution. Marine Ecology Progress Series, 84: 63–69.

Ringwood, A.H., Conners, D.E., and Di Novo, A. 1998. The effects of copper exposures on cellular responses in oysters. Marine Environmental Research, 46: 591–595.

Ringwood, A.H., Connors, D.E., and Hoguet, J. 1998. Effects of natural and anthropogenic stressors on lysosomal destabilisation in oysters, *Crassostrea virginica.* Marine Ecology Progress Series, 166: 163–171.

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

Tremblay, R., Myrand, B., and Guderley, H. 1998. Temporal variation of lysosomal capacities in relation to susceptibility of mussels, *Mytilus edulis,* to summer mortality. Marine Biology, 132: 641–649.

Widdows, J., and Salkeld, P. 1992. Practical procedures for the measurement of scope for growth. MAP Technical Reports Series, 71: 147–172.

Widdows, J., and Johnson, D. 1988. Physiological energetics of *Mytilus edulis:* Scope for growth. Marine Ecology Progress Series, 46(1-3): 113–121.

Al-Sabti K and Metcalfe C. 1995. Fish micronuclei for assessing genotoxicity in water. Mutat Res 343:121–135.

Grisolia CK and Cordeiro CMT. 2000. Variability in micronucleus induction with different mutagens applied to several species of fish. Genetics and Molecular Biology 23(1):235–239.

Cajaraville, M.P., Díez, G., Marigómez, J.A., and Angulo, E. 1990. Responses of the basophilic cells of the digestive gland of mussels to petroleum hydrocarbon exposure. Diseases of Aquatic Organisms, 9: 221–228.

Cajaraville, M.P., Marigómez, J.A., Díez, G., and Angulo, E. 1992. Comparative effects of the WAF of three oils on mussels. 2. Quantitative alterations in the structure of the digestive tubules. Comparative Biochemistry and Physiology, 102C(1): 113–123.

Viarengo, A., Lowe, D., Bolognesi, C., Fabbri, E., and Koehler, A. 2007. The use of biomarkers in biomonitoring: a 2-tiered approach assessing the level of pollutant-induced stress syndrome in sentinel organisms. Comp Biochem Phys. 146(3), 281-300.

Annex 21: Guidelines for the Integrated Monitoring and Assessment of Contaminants and their Effects

General introduction

Our seas and oceans are dynamic and variable. They represent a fundamental component of global ecosystems and as such we need to be able to assess the health status of the marine environment. Furthermore, we need to be able to detect anthropogenically induced changes in seas and oceans and to be able to identify the reasons for these changes. It is only through such understanding that we can advise on necessary and appropriate remedial responses, such as regulatory action, as well as report on any improvements resulting from OSPAR measures. There is a need to express clearly what is meant by the 'health' of the marine environment and for that purpose we require indicators of components of ecosystem health.

The marine environment receives inputs of hazardous substances through riverine inputs and direct discharges, as well as by atmospheric deposition. The marine environment is the ultimate repository for complex mixtures of persistent chemicals. This means that organisms are exposed to a range of substances, many of which have the potential to cause metabolic disorders, an increase in disease prevalence and, potentially, effects on populations through changes in e.g. growth, reproduction and survival. There is general agreement that the best way to assess the environmental quality of the marine environment, with respect to hazardous substances, is by using a suite of chemical and biological measurements in an integrated fashion. In the past, monitoring to assess the 'impact' of hazardous substances has been based primarily on measurements of concentration. This was because the questions being asked concerned concentrations of such substances in water, sediment and biota and such measurements were possible. However, in order to more fully assess the health of our maritime area, questions about the bioavailability of hazardous substances and their impact on marine organisms or processes are now being posed. Biological effects techniques have become increasingly important in recent years. The specific focus from OSPAR is on determining whether there are any unintended/unacceptable biological responses, or unintended/unacceptable levels of such responses, as a result of exposure to hazardous substances. Sometimes a biological response can be observed when the causative substance is below current chemical analytical detection limits; the development of imposex in gastropod molluscs due to tributyltin (TBT) is a point in case.

This guidance document is intended to replace the original JAMP Guidelines for monitoring contaminants in biota and sediment or biological effects which do not provide guidance for the optimum approach to monitoring to support the integrated assessment of concentrations and effects of contaminants across the OSPAR Maritime Area, although some of them contain references to supporting measurements (chemical data, physical data, biological data) which aid the interpretation of monitoring data. Consequently, chemical analytical and biological effects data have usually been collected, reported and assessed separately. Also, in some cases, the original Guidelines do not provide guidance on the specific substances which should be determined in order to be able to explicitly link concentrations and effects. An integrated approach to monitoring is based on the simultaneous measurement of contaminant concentrations (in biota, sediments and, in some cases, water or passive samplers), biological effects parameters and a range of physical and other chemical measurements so as to permit normalization and appropriate assessment.

Integrated monitoring of contaminants and their effects requires coordination of field sampling and sample handling techniques, utilizing the same species/population/individual for both types of measurement, from the same area and sampled within the same time frame. Furthermore, a set of supporting parameters should be measured at the same time and such data have to be available for use in the final assessment, because biological effects may be influenced by e.g. temperature, stage of maturation or size. Integration of effort in this way will yield additional information in a cost-effective manner, while also reducing the interannual variance of the data.

OSPAR has obligations to measure and monitor the quality of the marine environment and its compartments (water, sediments, and biota), the activities and inputs that can affect that quality and the effects of those activities and inputs, and to assess what is happening in the marine environment as a basis for identifying priorities for action. OSPAR, together with HELCOM, have agreed on an ecosystem approach to managing the marine environment under which OSPAR has committed to monitoring the ecosystems of the marine environment, in order to understand and assess the interactions between, and impact of, human activities on marine organisms. Integrated monitoring and assessment of contaminants in the marine environment and their effects will contribute effectively to the integrated assessment of the full range of human impacts on the quality status of the marine environment, as part of the ecosystem approach.

The OSPAR Hazardous Substances Strategy

The objective of the OSPAR Hazardous Substances Strategy (OSPAR Agreement 2003–2021) is to prevent pollution of the maritime area by continuously reducing discharges, emissions and losses of hazardous substances, with the ultimate aim of achieving concentrations in the marine environment near background values for naturally occurring substances and close to zero for man-made synthetic substances. The Hazardous Substances Strategy further declares that the Commission will implement this Strategy progressively by making every endeavour to move towards the target of the cessation of discharges, emissions and losses of hazardous substances by the year 2020. In association with this, and the other five OSPAR strategies, OSPAR has developed a Joint Assessment and Monitoring Programme (JAMP). This provides the basis for the monitoring activities undertaken by Contracting Parties to assess progress towards achieving OSPAR objectives. In relation to hazardous substances, the JAMP seeks to addresses the following questions:

- What are the concentrations in the marine environment, and the effects, of the substances on the OSPAR List of Chemicals for Priority Action ("priority chemicals")? Are they at, or approaching, background levels for naturally occurring substances and close to zero for manmade substances?
- Are there any problems emerging related to the presence of hazardous substances in the marine environment? In particular, are any unintended/ unacceptable biological responses, or unintended/ unacceptable levels of such responses, being caused by exposure to hazardous substances?

There is a need to adopt an integrated approach to the monitoring of contaminants in the marine environment and the biological responses to the presence of hazardous substances. Such an approach would provide greater interpretative power in assessments of the state of the OSPAR Maritime Area with respect to hazardous substances and an improved assessment of progress towards achieving the objectives of the OSPAR Hazardous Substances Strategy.

EU Water Framework Directive and Marine Strategy Framework Directive

The marine environment is a precious heritage that must be protected, restored and treated as such with the ultimate aim of providing biologically diverse and dynamic oceans and seas that are safe, clean, healthy and productive. It is in this context that the European Union has over the last decade developed its water policies such that significant European Legislation incorporating marine waters and the lakes and rivers which ultimately flow into our coastal ecosystems. The Water Framework Directive (Directive 2000/06/EC) establishes a framework for Community action in the field of water policy, central to which is good ecological status for water bodies. This is described on the basis of biological quality elements, hydromorphological quality elements and physico-chemical quality elements. More recently, the European Union has implemented the Marine Strategy Framework Directive (Directive 2008/56/EC). At its heart is the concept of "Good Environmental Status" for all European waters and the provision of a framework for the protection and preservation of the marine environment, the prevention of its deterioration and where practicable the restoration of that environment in areas where it has been adversely affected. Good Environmental Status (GES) will be assessed on a regional basis and as such the programmes of the various Regional Sea Conventions, including OSPAR, will provide a valuable source of data for the assessments that will be required.

The Directive specifies that GES will be assessed against eleven qualitative Descriptors. Descriptor 8 (Concentrations of contaminants are at levels not giving rise to pollution effects) has been interpreted as requiring assessments of contaminant concentrations and their biological effects.

A Task Group set up by ISPRA interpreted this as meaning that the concentrations of contaminants should not exceed established quality standards (e.g. EQSs, EACs) and that the intensity of biological effects attributable to contaminants should not indicate harm at organism or higher levels of organization. Commission Decision (2010/477/EU) noted that progress towards good environmental status will depend on whether pollution is progressively being phased out, i.e. the presence of contaminants in the marine environment and their biological effects are kept within acceptable limits, so as to ensure that there are no significant impacts on or risk to the marine environment.

It is clear that assessment for Descriptor 8 will require both chemical and biological effects measurements. It is likely that a robust and holistic approach will seek to integrate the assessment chemical and biological effects data into a single process.

Purpose of this Guideline

The purpose of this document is to provide guidance on integrated chemical and biological effects monitoring within the OSPAR area, in the context of the Coordinated Environmental Monitoring Programme (CEMP) issues and the list of OSPAR priority chemicals. In addition, it provides the context for the associated Technical Annexes describing biological effects techniques include a list of the supporting parameters which are required in an integrated programme, as well as the chemical determinands relevant to the effects being studied.

The Guideline is supported by associated Background Documents which provide information on the scientific background to the contaminants and biological effects measurements included in the Programme, and on the derivations and values of assessment criteria (Background Concentrations, Background Assessment Concentrations, and Environmental Assessment Criteria for chemical contaminants, and analogous assessment criteria for biological effects measurements).

Quantitative objectives; Temporal Trend and Spatial Programmes

The ultimate objectives of OSPAR monitoring activities relating to hazardous substances are:

- to assess status (existing level of marine contamination and its effect) and trends of hazardous substances across the OSPAR maritime area;
- to assess the effectiveness of measures taken for the reduction of marine contamination;
- to assess harm (unintended/unacceptable biological responses) to living resources and marine life;
- to identify areas of serious concern/hot spots and their underlying causes;
- to identify unforeseen impacts and new areas of concern;
- to create the background to develop prediction of expected effects and the verification thereof (hindcasting); and
- to direct future monitoring programmes.

By being clear about the objective of the monitoring, the parameters for inclusion in the programme of work, the sampling strategy, methods of statistical analysis and assessment methods can all be developed and specified. In the context of integrated monitoring, the planning aspect is crucial as it will ensure that operating procedures can be put in place that clearly detail all the chemical, physical and biological samples and data to be collected.

There is a need to perform monitoring which will identify differences over time and across geographical space. This will divide monitoring into two generic types:

- Spatial monitoring: monitoring to identify geographical variation within the OSPAR maritime area;
- Temporal Monitoring: monitoring aimed at identifying changes over time.

Although these two types of monitoring have been described separately, there is no reason why the two activities cannot be carried out simultaneously, as long as this is incorporated into the design of the programme. The processes of integration for both these types of monitoring are closely related and hence should be developed simultaneously.

The integrated approach

The contribution made by the 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. By bringing together monitoring disciplines which have tended to be conducted separately, an integrated assessment can further lead to an improved ability to explain the causes for hot spots detected during monitoring programmes. An integrated approach also has the advantage of combining and coordinating the various disciplines to achieve a greater understanding among those performing marine assessments of the contributions from the different components of a monitoring programme. This has the clear technical advantage that sampling of all relevant parameters at any particular sampling location will be assured. The economic benefit of an integrated approach comes from the fact that the samples and data are gathered during a single cruise and that the data can be directly compared/used with holistic assessment tools to provide truly integrated assessments.

The integration of sampling has four distinct connotations:

- sampling and analyses of same tissues and individuals;
- sampling of individuals for effects and chemical analyses from the same population as that used for disease and/or population structure determination at the same time;
- sampling of water, the water column (if included) and sediments at the same time and location as collecting biota; and
- simultaneous measurement of support parameters (e.g. hydrographic parameters) at any given sampling location.

Fundamental aspects of the design of an integrated programme include key environmental matrices (water, sediment and biota), the selection of appropriate combinations of biological effects and chemical measurements and the design of sampling programmes to enable the chemical concentrations, the biological effects data and other supporting parameters to be combined for assessment. The basic structure of an integrated programme is illustrated in Figure 1.

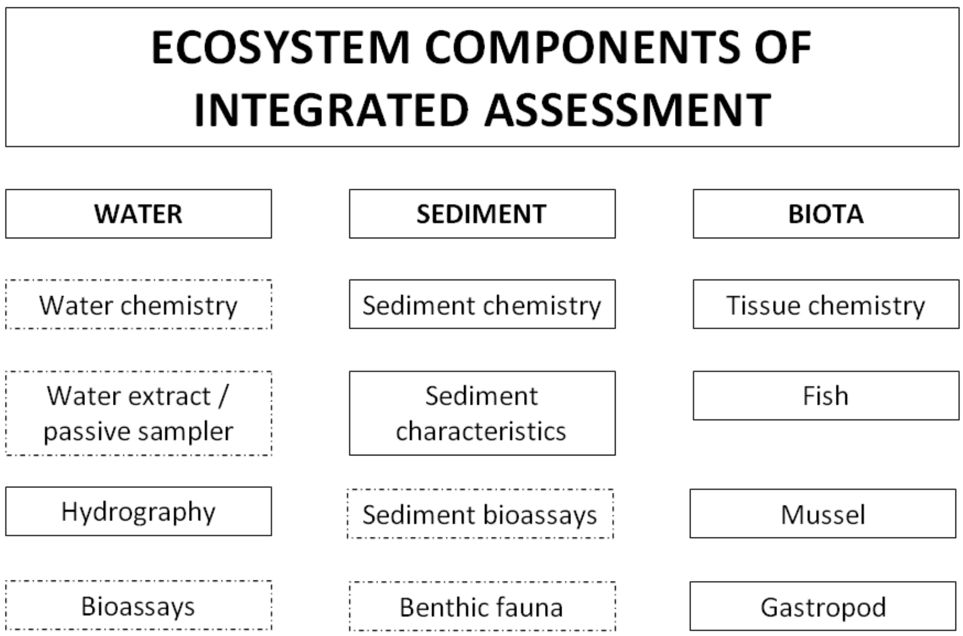


Figure 1. Overview of components in a framework for the integrated monitoring programme chemical contaminants and their biological effects. Solid lines – core methods, broken lines – additional methods.

Chemical analyses to be included in an integrated programme for OSPAR purposes should cover the OSPAR priority hazardous substances. Analytical methods should be sufficiently sensitive to detect variation in environmental quality, and supported by appropriate quality control and assurance. Biological effects methods to be included in an integrated programme have been identified by the ICES Working Group on the Biological Effects of Contaminants (WGBEC). They require the following characteristics:

1) the ability to separate contaminant-related effects from influence by other factors (e.g. natural variability, food availability, etc);
2) sensitivity to contaminants, i.e. provide "early warning";
3) the suite of methods used should cover a range of mechanizms of toxic action, e.g. estrogenicity/androgenicity, carcinogenicity, genotoxicity and mutagenicity;
4) the range of methods applied in an integrated programme should include at least one that measures the "general health" of the organism.

Biological effects and chemical methods were selected for the biota matrix (separated as fish and mussel) using these criteria. In addition, some physiological characteristics of individual fish are required including gonad somatic index (GSI), liver somatic index (LSI) and condition factor, as described in supporting Technical Annexes. Similarly, spawning status is relevant to mussel effect assessment. General designs for integrated monitoring of fish are presented in Figure 2 and of mussel in Figure 3. Designs for water, sediment and gastropod monitoring are included as Figures 4, 5 and 6 respectively.

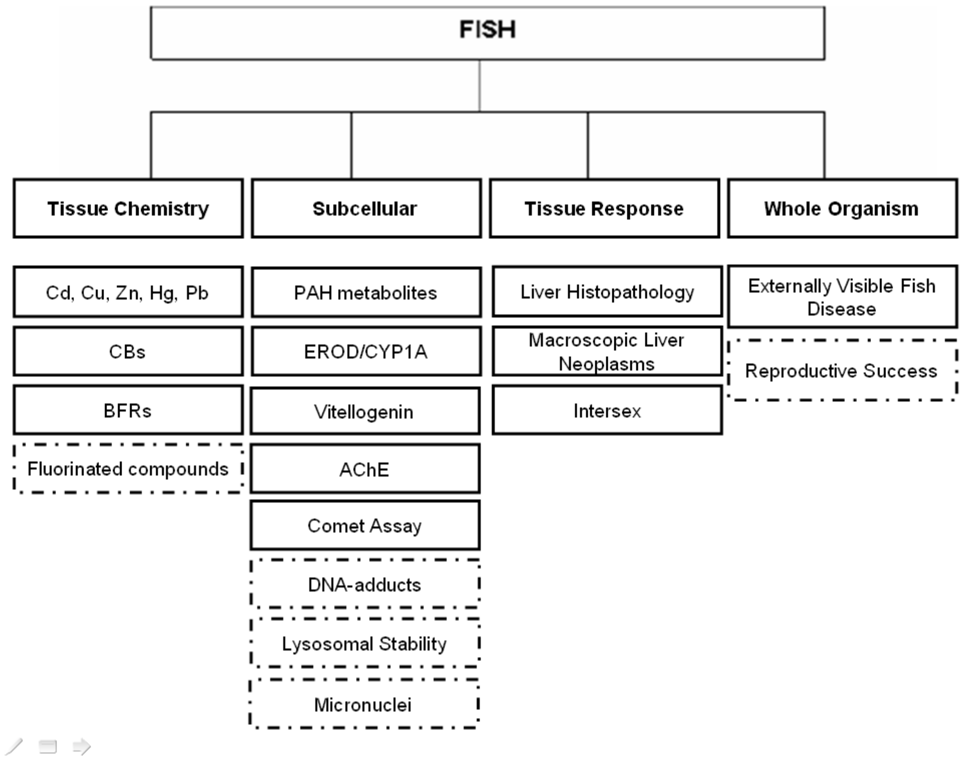


Figure 2. Methods included in the fish component of the integrated monitoring framework; solid lines – core methods, broken lines – additional methods.

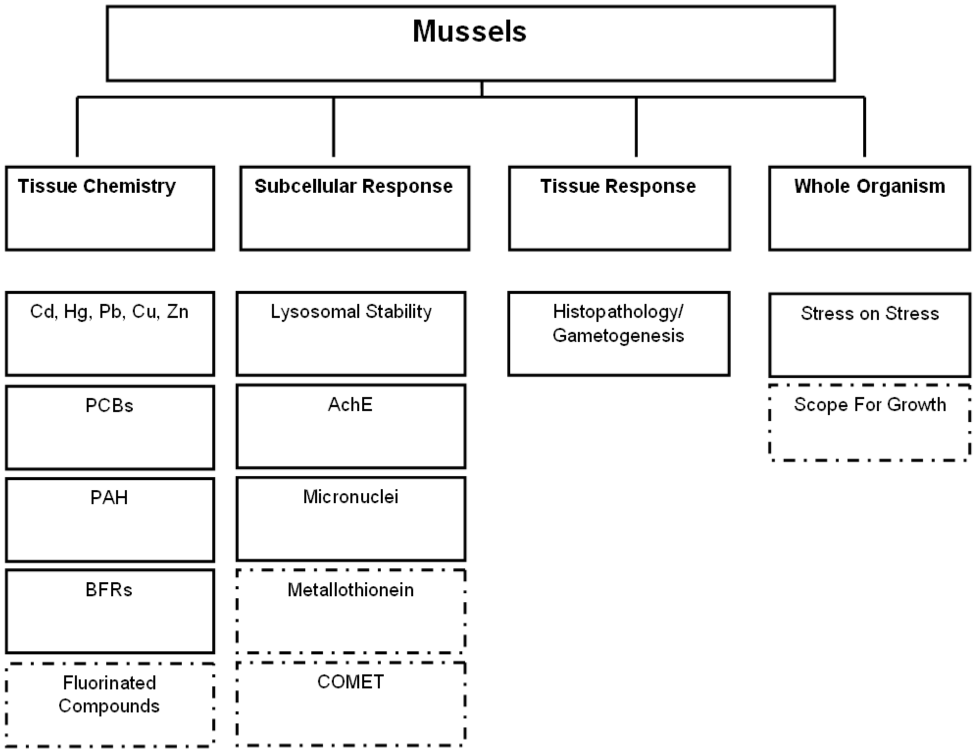


Figure 3. Methods included in the mussel component of the integrated monitoring framework; solid lines – core methods, broken lines – additional methods.

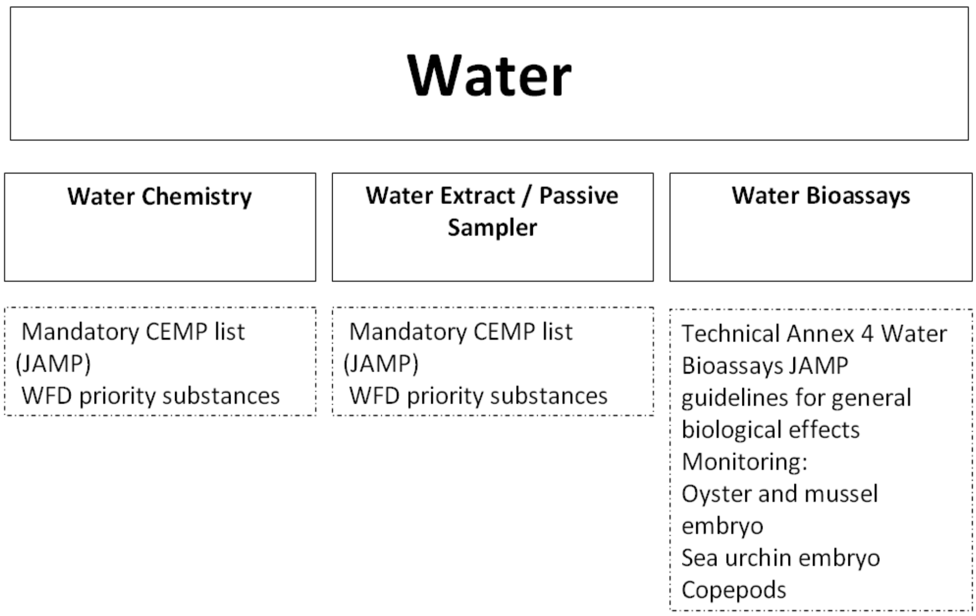


Figure 4. Methods included in the water component of the integrated monitoring framework; solid lines – core methods, broken lines – additional methods.

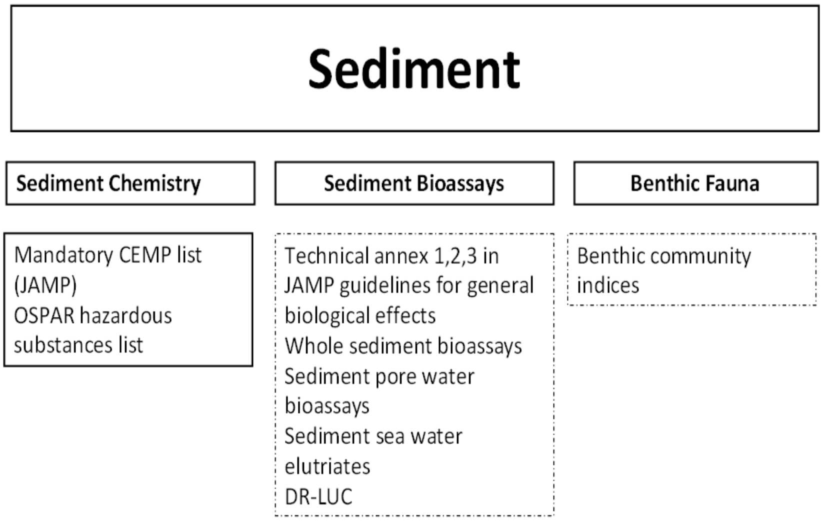


Figure 5. Methods included in the sediment component of the integrated monitoring framework; solid lines – core methods, broken lines – additional methods.

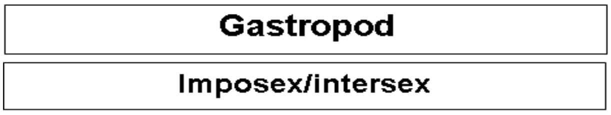


Figure 6. Methods included in the gastropod component of the integrated monitoring framework; solid lines – core methods, broken lines – additional methods.

Sampling and analysis strategies for integrated fish and bivalve monitoring

The integration of contaminant and biological effects monitoring requires a strategy for sampling and analysis that includes the:

- sampling and analyses of same tissues and individuals;
- sampling of individuals for effects and chemical analyses from the same population as that used for disease and/or population structure determination at a common time;
- sampling of water, the water column and sediments at the same time and location as collecting biota; and
- more or less simultaneous sampling for and determination of primary and support parameters (e.g. hydrographic parameters) at any given location.

Examples of sampling strategies for the integrated fish and shellfish schemes are shown in Figures 7 and 8. In order to integrate sediment, water chemistry and associated bioassay components, with the fish and bivalve schemes, sediment and water samples should be collected at the same time as fish/bivalve samples and from a site or sites that are representative of the defined station/sampling area.

Additional integrated sampling opportunities may arise from trawl/grab contents, for example, gastropods for imposex or benthos, and these should be exploited where possible/practicable.

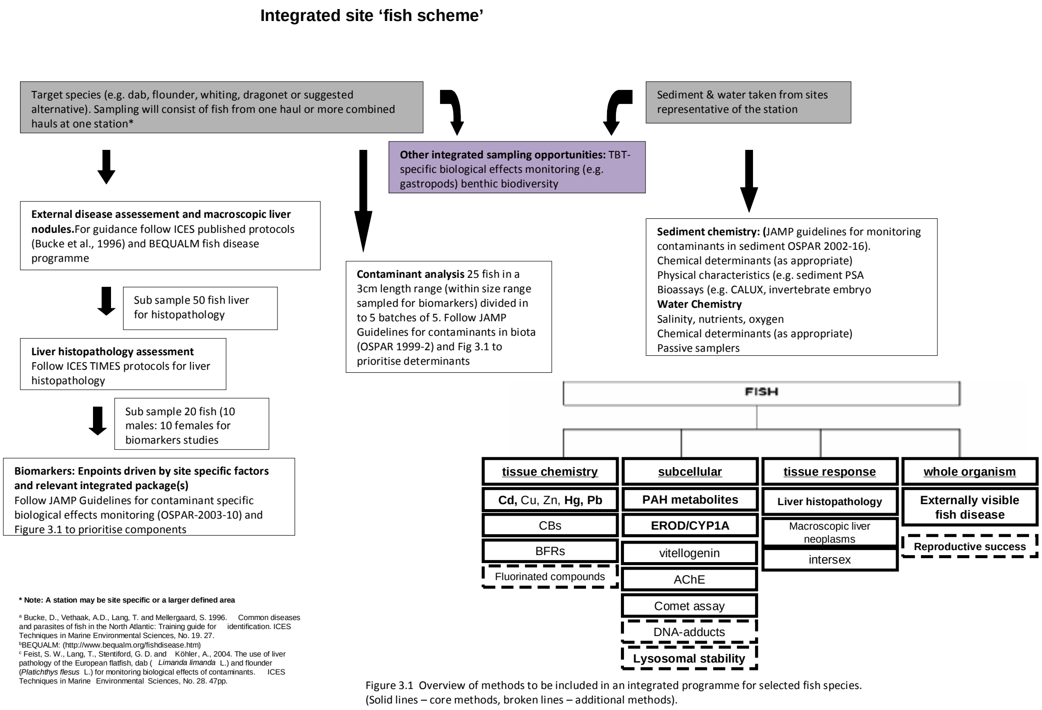


Figure 7. Sampling strategy for integrated fish monitoring.

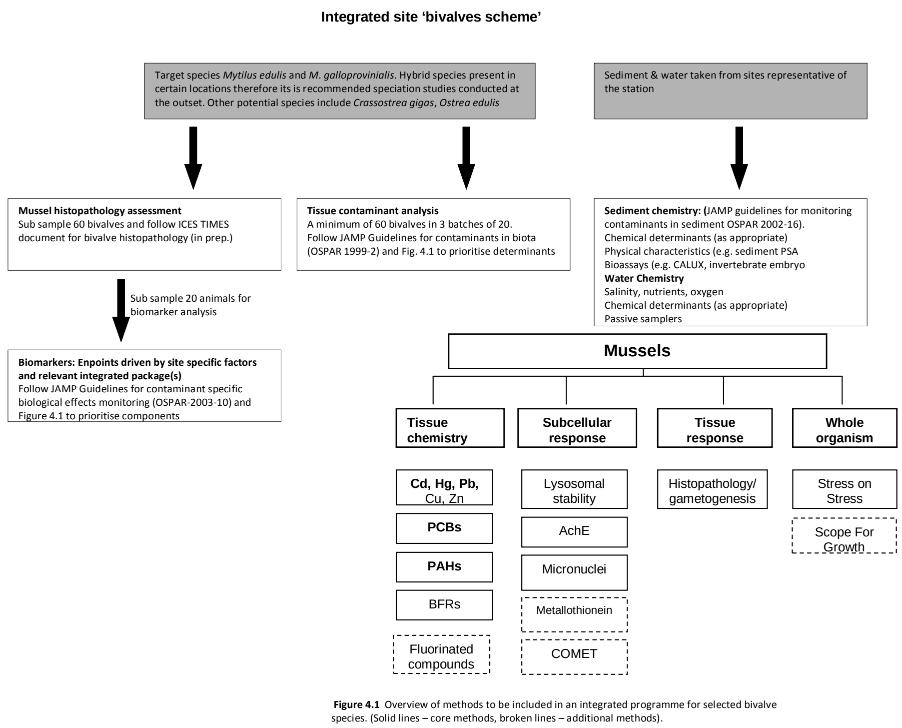


Figure 8. Sampling strategy for integrated bivalve monitoring.

The integrated assessment

It is not sufficient simply to coordinate sampling; integration must also involve a combined assessment of the monitored parameters, which must themselves be selected with the assessment aim in mind. Such a combined assessment may involve using environmental parameters as covariates in statistical analyses or they may be used to standardize effect-variables, e.g. temperature or seasonal effects on biomarker responses. Similarly, normalization procedures for the expression of contaminant concentrations in biota and sediment have been established, for example the use of defined bases (e.g. dry weight or lipid weight) for biota analyses, and normalization of sediment analyses to organic carbon or aluminium to minimize the influence of differences in bulk sediment properties. These are described in detail in the CEMP Monitoring Manual.

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). In order to assess progress towards the objectives of the OSPAR Hazardous Substances Strategy, OSPAR has developed assessment criteria for contaminant concentration data. These are Background Concentrations (BCs), Background Assessment Concentrations or Criteria (BACs) and Environmental Assessment Criteria (EACs). The use of these in data assessment, on both local and large (OSPAR Convention area) scales, is described in the CEMP Manual. The Manual also describes the statistical approaches to be used in comparing field data with assessment criteria to ensure rigorous and consistent assessments.

In the same way, OSPAR, with assistance from ICES, has more recently developed coherent sets of analogous assessment criteria for biological effects measurements. The concept of a background level of response has been found to be applicable to all effects measurements. Assessment criteria analogous to EACs, i.e. representing levels of response below which unacceptable responses at higher (e.g. organism or population) levels would not be expected, have been found to be applicable for some many biological effects measurements, and these have been termed biomarkers of effect. In other cases, the link to higher level effects is less clear and these measurements have been termed biomarkers of exposure, in that they indicate that exposure to hazardous substances has occurred. Importantly, the processes used to derive BACs and their biological analogues, and EACs and their analogues have been applied consistently to all chemical and effects measurements. The consequence is that the OSPAR objective of achieving background or near background concentrations/effects represents targets based upon the same criteria across all parameters, and that EACs and analogues represent similar levels of environmental risk. A table of the current assessment criteria for biological effects is presented as Annex 23 to the ICES SGIMC 2011 report.

This coherence across the broad range of assessment criteria forms the basis for integrated assessment schemes. The presentation of progress towards the objectives of the Hazardous Substances Strategy in the QSR 2010 document, in that the status of all OSPAR priority contaminants could be presented in directly comparable "traffic light" formats (Figure 7).

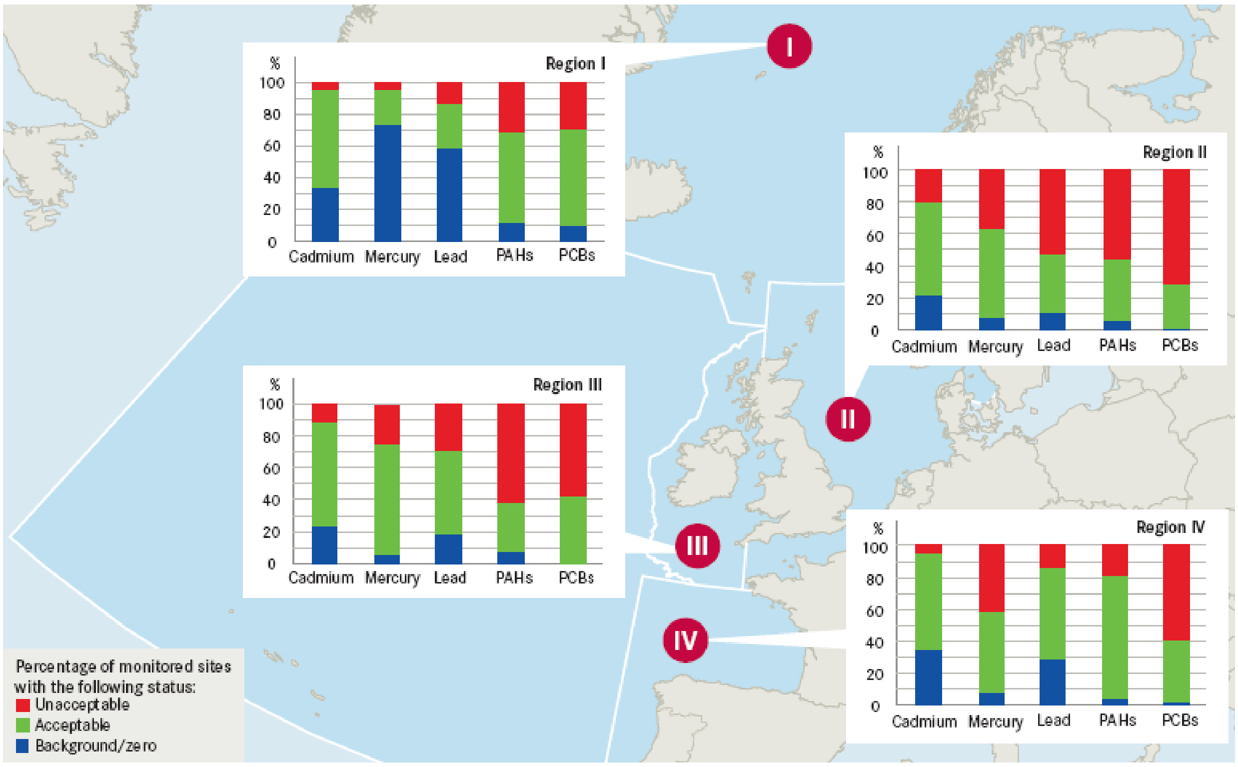


Figure 7. OSPAR regional level integration of the concentrations of priority contaminants in fish, shellfish and sediment, from the OSPAR QSR 2010, Hazardous Substances chapter.

A comparable approach can be used in the assessment of biological effects data, for which EACs and/or BACs have been developed. Furthermore, the coherence of assessment criteria across both chemistry and biological effects measurements allows these two types of data to be brought together into a single integrated assessment scheme. The "traffic light" presentation is equally applicable to biological effects data and can be used to present data integrated on a range of geographical scales from the single sampling site to the Regional scale, as required under MSFD. The current state of development of this approach is described in Section 8 ii) and Annex 25 of ICES/OSPAR SGIMC 2011 report.

Annex 22: Supporting documentation summary

Table 1. Biological effect techniques relevant to the ecosystem components for integrated monitoring and assessment of chemical and biological effects data. Status regarding availability of Background Documents, assessment criteria, and quality assurance.

Biological effect technique	Background document	Assessment Criteria	Quality Assurance
Oyster and mussel embryo test	X	X	A
Sea urchin embryo test	X	X	B
Copepod test (*Tisbe*)	X	X	A
Whole sediment bioassays	X	X	A
Sediment pore-water bioassays	X	X	A
Sediment seawater elutriates	X	X	A
DR-LUC	X	X	B (in future)
PAH metabolites	X	X	C, D
Cytochrome P4501A activity (EROD)	X	X	A, B, D, F
Vitellogenin	X	X	E
Acetylcholinesterase	X	X	B, E
Comet assay	X	X	E
Micronucleus formation	X	X	B, F
DNA adducts	X	X	
Metallothionein	X	X	A (fish), F (mussels)
Lysosomal stability (Cytochemical and neutral Red)	X	X	B(fish), B, F(mussels)
Liver pathology including neoplasia/hyperplasia	X	X	A
Macroscopic liver neoplasms	X	X	A
Intersex in fish	X	X	B (in future)
Mussel histopathology (gametogenesis)	X	X	B (in future)
Imposex/ Intersex in gastropods	X	X	C
Stress on Stress (SoS)	X	X	not required
Scope for growth	X	X	B
Externally visible fish diseases	X	X	A
Reproductive success in eelpout	X	X	A

A: BEQUALM

B: Between particular independent laboratories

C: QUASIMEME

D: BEAST

E: WGBEC

F: MEDPOL

Annex 23: Technical annex: Assessment criteria for biological effects measurements

Table 1: Assessment criteria for biological effects measurements. Values are given for both background assessment levels (BAC) and environmental assessment criteria (EAC), as available.

BIOLOGICAL EFFECT	APPLICABLE TO:	BAC	EAC
VTG in plasma; µg/ml	Cod	0.23	
	Flounder	0.13	
Reproduction in eelpout; mean frequency (%)	Malformed larvae	1	
	Late dead larvae	2	
	Growth retarded larvae	4	
	Frequency of broods with malformed larvae	5	
	Frequency of broods with late dead larvae	5	
EROD; pmol/mg protein pmol/min/ mg protein S9 * pmol/min/ mg microsomal protein	Dab (F)	178	
	Dab (M)	147	
	Dab (M/F)	680*	
	Flounder (M)	24	
	Plaice (M)	9.5	
	Cod (M/F)	145*	
	Plaice (M/F)	255*	
	Four spotted megrim (M/F)	13*	
	Dragonet (M/F)	202*	
	Red mullet (M)	208	
PAHs Bile metabolites; [1] ng/ml; HPLC-F [2] pyrene-type µg/ml; synchronous scan fluorescence 341/383 nm [3] ng/g GC/MS * 1-OH pyrene ** 1-OH phenanthrene	Dab	16 [1]* 3.7 [1]** 0.15 [2]	 22[2]
	Cod	21 [1]* 2.7 [1]** 1.1 [2]	483 [3]* 528 [3]** 35 [2]
	Flounder	16 [1]* 3.7 [1]** 1.3 [2]	 29[2]
	Haddock	13 [1]* 0.8 [1]** 1.9 [2]	 35[2]
DR-Luc; ng TEQ/kg dry wt, silica clean up	Sediment (extracts)	10	40

DNA adducts; nm adducts mol DNA	Dab	1	6
	Flounder	1	6
	Cod	1.6	6
	Haddock	3.0	6
Bioassays; % mortality	Sediment, Corophium	30	60
	Sediment, Arenicola	10	50
	Water, copepod	10	50
Bioassays; % abnormality	Water, oyster and mussel embryo	20	50
	Water, sea urchin embryo	10	50
Bioassay; % growth	Water, sea urchin embryo	30	50
Lysosomal stability; minutes	Cytochemical; all species	20	10
	Neutral Red Retention: all species	120	50
Micronuclei; ‰ (frequency of micronucleated cells) [1] Gill cells [2] Haemocytes [3] Erythrocytes	*Mytilus edulis*	2.5 [1] 2.5 [2]	
	Mytilus galloprovincialis	3.9 [2]	
	Mytilus trossulus	4.5 [2]	
	Flounder	0.0-0.3 [3]	
	Dab	0.5 [3]	
	Zoarces viviparus	0.3-0.4 [3]	
	Cod	0.4 [3]	
	Red mullet	0.3 [3]	
Comet Assay; % DNA Tail	*Mytilus edulis*	10	
	Dab	5	
	Cod	5	
Stress on Stress; days	*Mytilus* sp.	10	5
AChE activity; $nmol.min^{-1}$ mg $prot^{-1}$ [1] gills [2] muscle tissue [3] brain tissue * French Atlantic waters ** Portuguese Atlantic waters [+] French Mediterranean Waters [++] Spanish Mediterranean Waters	*Mytilus edulis*	30 [1*]	21 [1*]
		26 [1**]	19 [1**]
	Mytilus galloprovincialis	29[1+]	20[1+]
		15 [1++]	10 [1++]
	Flounder	235 [2*]	165 [2*]
	Dab	150 [2*]	105 [2*]
	Red mullet	155 [2+] 75 [3++]	109 [2+] 52 [3++]
Externally visible diseases*** Ep,Ly,Ul Ep,Ly,Ul Ac,Ep,Fi,Hp,Le,Ly,St,Ul,Xc	Dab	Fish Disease Index (FDI): F: *1.32*, 0.216 M: *0.96*, 0.232 F: *1.03*, 0.349	Fish Disease Index (FDI): F: *NA*, 54.0 M: *NA*, 47.7 F: *50.6*, 19.2

Ac,Ep,Fi,Hp,Le,Ly,St,Ul,Xc Ac,Ep,Hp,Le,Ly,St,Ul,Xc Ac,Ep,Hp,Le,Ly,St,Ul,Xc Italics: ungraded, bold: graded NA: Not applied		M: *1.17*, 0.342 F: *1.09*, 0.414 M: *1.18*, 0.398 M: males F: females	M: *38.8*, 16.1 F: *48.3*, 21.9 M: *35.2*, 16.5
Liver histopathology-non specific	Dab	NA	Statistically significant increase in mean FDI level in the assessment period compared to a prior observation period *or* Statistically significant upward trend in mean FDI level in the assessment period
Liver histopathology- contaminant-specific	Dab	Mean FDI <2	Mean FDI ≥ 2 A value of FDI = 2 is, e. g., reached if the prevalence of liver tumours is 2 % (e. g., one specimen out of a sample of 50 specimens is affected by a liver tumour). Levels of FDI ≥ 2 can be reached if more fish are affected or if combinations of other toxicopathic lesions occur.
Macroscopic liver neoplasms	Dab	Mean FDI <2	Mean FDI ≥ 2 A value of FDI = 2 is reached if the prevalence of liver tumours (benign or malignant) is 2 % (e. g., one specimen out of a sample of 50 specimens is affected by a liver tumour). If more fish are affected, the value is FDI > 2.
Intersex in fish; % prevalence	Dab Flounder Cod Red mullet *Zoarces viviaprus*	5	
Scope for growth Joules/hr/g dry wt.	Mussel (*Mytilus* sp.) (provisional, further validation required)	5	-2
Hepatic metallothionein µg/g (w.w.)	*Mussel edulis*	0.6 [1*] 2.0 [2*]	

[1] Whole animal [2]Digestive gland [3]Gills * Differential pulse polarography		0.6 [3*]	
	Mytilus galloprovincialis	2.0 [1*] 3.9[2*] 0.6 [3*]	
Histopathology in mussels	VVbas: Cell type composition of digestive gland epithelium; $\mu m^3/\mu m^3$ *(quantitative)*	0.12	0.18
	MLR/MET: Digestive tubule epithelial atrophy and thinning; µm/µm *(quantitative)*	0.7	1.6
	VVLYS & Lysosomal enlargement; $\mu m^3/\mu m^3$ *(quantitative)*	VvLYS 0.0002	V>0.0004
	S/VLYS: $\mu m^2/\mu m^3$	4	
	Digestive tubule epithelial atrophy and thinning *(semi-quantitative)*	STAGE ≤1	STAGE 4
	Inflammation *(semi-quantitative)*	STAGE ≤1	STAGE 3
Imposex/intersex in snails	Gastropod molluscs	See OSPAR adopted criteria	See OSPAR adopted criteria

*****: Assessment criteria for the assessment of the Fish Disease Index (FDI) for externally visible diseases in common dab (*Limanda limanda*). Abbreviations used: Ac, *Acanthochondria cornuta;* Ep, Epidermal hyperplasia/papilloma; Fi, Acute/healing fin rot/erosion; Hp, Hyperpigmentation; Le, *Lepeophtheirus sp.*; Ly, Lymphocystis; St, *Stephanostomum baccatum*; Ul, Acute/healing skin ulcerations; Xc, X-cell gill disease.**

Full details of the assessment criteria and how they were derived can be found in the SGIMC 2010 and SGIMC 2011 and WKIMON 2009 reports on the ICES website and in the OSPAR Background Documents for individual biological effects methods.

Data for biomarkers in some northern fish species have been obtained through the IRIS BioSea JIP programme (funded by Total E&P Norge & EniNorge) and the Biomarker Bridges programme (funded by Research Council of Norway) and have been used to develop EAC and BAC values for Arctic fish.

Annex 24: Development of MIME 2

ANNEX 7

(Ref. §8.5a)

OSPAR Convention for the Protection of the Marine Environment of the Northeast Atlantic

Working Group on Monitoring and on Trends and Effects of Substances in the Marine Environment (MIME)

Copenhagen: 7–10 December 2010

Further development by ICES/OSPAR Study Group on Integrated Monitoring of Contaminants and Biological Effects (SGIMC), Copenhagen: 14–18 March 2011

Biological effects relevant to good environmental status in the OSPAR area

This provides advice to HASEC 2011 for the development of a consolidated list of biological effects techniques that can act as targets and indicators for good environmental status in the OSPAR area.

Background

Descriptor 8 for good environmental status under the Marine Strategy Framework Directive requires: "Concentrations of contaminants are at levels not giving rise to pollution effects".

Within Commission Decision 2010/477/EU on criteria and methodological standards for GES in marine waters, the suggested criteria include:

8.1. Concentration of contaminants

— Concentration of the contaminants mentioned above, measured in the relevant matrix (such as biota, sediment and water) in a way that ensures comparability with the assessments under Directive 2000/60/EC (8.1.1)

8.2. Effects of contaminants

— Levels of pollution effects on the ecosystem components concerned, having regard to the selected biological processes and taxonomic groups where a cause/effect relationship has been established and needs to be monitored (8.2.1)

— Occurrence, origin (where possible), extent of significant acute pollution events (e.g. slicks from oil and oil products) and their impact on biota physically affected by this pollution (8.2.2).

Considerations for developing a list of substances relevant to GES in the OSPAR area

The OSPAR Commission, in cooperation with ICES, has developed a number of biological effects monitoring techniques and associated assessment criteria to measure response within marine organisms. This includes contaminant-specific techniques and general techniques which reflect responses to multiple contaminants. The most robust technique with a clear cause/effect relationship is the measurement of TBT-specific effects (imposex) in gastropods. Other techniques reflect mechanizms of toxicity and therefore respond to groups of substances that have a similar mode of action (for example, EROD responds to planar organic contaminants).

In many cases, it has not been possible so far to link chemical monitoring with observations of effects in species in such a way that conclusions can be drawn about the impact of the on the functioning of ecosystems at a regional level. Any list of biological effects relevant to good environmental status in the OSPAR area would need to allow evolution as new knowledge becomes available.

Recommended starting point for a list of relevant biological effects techniques which could act as targets and indicators for good environmental status is the recommendations from ICES/OSPAR SGIMC on ecosystem components for integrated assessment of chemical and biological effects data. These have been developed for seawater, sediment, fish, mussel and gastropods, and for the group of substances covered by the CEMP for concentration monitoring: metals; PCBs and polychlorinated dibenzodioxins and furans; PAHs and alkylated PAHs.

The suites of biological effects techniques included in the integrated assessment approach have been identified on the basis of the following characteristics of the maturity of the methods, in terms of their scientific and practical basis, and the ability to assess the data in an integrated manner:

1) the ability to separate contaminant-related effects from influence by other factors (e.g. natural variability, food availability, etc);
2) sensitivity to contaminants, i.e. provide "early warning";
3) the suite of methods used should cover a range of mechanizms of toxic action, e.g. estrogenicity/androgenicity, carcinogenicity, genotoxicity and mutagenicity;
4) the range of methods applied in an integrated programme should include at least one that measures the "general health" of the organism;
5) the availability of Background Documents;
6) the availability of assessment criteria; and
7) the stage of development quality assurance.

Many are already included in the pre-CEMP. Contracting Parties have made progress in standardizing reference methods for monitoring biological indicators, but have not yet implemented a fully coordinated biological effects monitoring programme. This will be needed to support the regional assessment of hazardous substances.

Table 1 lists the biological effects techniques identified in the ecosystem components of integrated monitoring and assessment of chemical and biological effects monitoring for contaminants (Appendix 1) and highlights the status of development of supporting documentation for coordinated monitoring, assessment tools and quality assurance.

Assessment criteria for both chemical concentrations and biological effects are keys to the development integrated assessments. The QSR 2010 Chapter 5 (Hazardous substances) contained integrated assessments of the concentrations of contaminants in sediment, fish and shellfish, based upon a coherent set of assessment criteria (Background Assessment Criteria (BACs) and Environmental Assessment Criteria (EACs)). The former were used to assess progress towards the OSPAR Hazardous Substances Strategy objective of concentrations at or close to background. The EACs are concentrations below which unacceptable biological responses are unlikely to occur, and represent equivalent levels of environmental risk for each contaminant. This consistency of approach has been extended to the biological effects, and has led to the development of assessment criteria for effects that are analogous to BACs and EACs.

The assessment criteria for biological effects developed by ICES WGBEC and ICES/OSPAR SGIMC and included background documents published by OSPAR are summarized in Appendix 2, and cover all the effects measurements in the ecosystem components of integrated monitoring and assessment of chemical and biological effects monitoring for contaminants (Appendix 1). For some measurements (biomarkers of effect), both BAC and EAC analogues have been developed. In other cases, only BACs have been developed, and these methods are considered to be biomarkers of exposure to contaminants. Except for TBT-specific effects (imposex), assessment criteria have not yet been formally agreed by OSPAR.

The development of a coherent suite of assessment criteria for both chemical concentrations and biological effects measurements lays the foundation for the extension of the integrated assessment of chemical data (e.g. Appendix 3) to the integration of chemical and biological effects data in a single presentation. Data from single effects measurements could be presented, as for the contaminants in Appendix 3, or could be grouped in various ways. For example, data for biomarkers of effect could be separately grouped from biomarkers of exposure, or biomarkers of relevance to particular groups of contaminants could be grouped. Spatial integration can be undertaken at a wide range of scales, from the individual sampling site to the MSFD regional scale. Details of the current state of development of this approach are described in Section 8 ii) and Annex 25 of the report from SGIMC 2011.

SGIMC 2011 gave some preliminary consideration to how this form of data integration might be used to assess GES for Descriptor 8. SGIMC noted that measurements of various parameters in various environmental matrices at various stations can be progressively summarized into simple visual representations of status at different degrees of data aggregation. At the highest level, data for both contaminant concentrations and their effects can be represented at MSFD Regional level by a single three colour “traffic light”. SGIMC consider that the critical boundary for GES assessment should be the green – red boundary, representing comparisons with EACs. SGIMC considered that GES could be expressed as some high percentage compliance with this boundary.

SGIMC consider that 100% compliance is impractical, as it amounts to a “one out all out” approach, and is therefore highly susceptible to perturbations by a small number of errors in sampling, analysis or data handling, or occasional short-term variations in environmental quality. SGIMC therefore suggest that 95% compliance at the highest level of data aggregation would be an appropriate threshold for GES compliance.

The OSPAR assessment criteria for contaminant concentrations in environmental matrices are largely confined to the OSPAR priority contaminants. It is possible that the suite of chemicals to be monitored for MSFD may include additional substances, such as those included in aspects of monitoring under the Water Framework Directive. The inclusion of these additional substances in the integrated assessment will require there to be assessment criteria for these substances in monitoring matrices or ecosystem components. As a starting point, EQSs developed according to EU guidance could be considered as satisfactory analogues to EACs. The principles developed by OSPAR for the definition of BACs should also be applicable additional substances.

Table 1. Biological effects techniques relevant to the ecosystem components for integrated monitoring and assessment of chemical and biological effects data. Status regarding availability of Background Documents, assessment criteria, and quality assurance.

Biological effect technique	Background document	Assessment Criteria	Quality Assurance
Oyster and mussel embryo test	X	X	A
Sea urchin embryo test	X	X	B
Copepod test (*Tisbe*)	X	X	A
Whole sediment bioassays	X	X	A
Sediment pore-water bioassays	X	X	A
Sediment seawater elutriates	X	X	A
DR-LUC	X	X	B (in future)
PAH metabolites	X	X	C, D
Cytochrome P4501A activity (EROD)	X	X	A, B, F
Vitellogenin	X	X	E
Acetylcholinesterase	X	X	B, E
Comet assay	X	X	E
Micronucleus formation	X	X	B, F
DNA adducts	X	X	
Metallothionein	X	X	A (fish), F (mussels)
Lysosomal stability (Cytochemical and neutral Red)	X	X	B(fish), B, F(mussels)
Liver histopathology	X	X	A
Macroscopic liver neoplasms	X	X	A
Intersex in fish	X	X	B (in future)
Mussel histopathology (gametogenesis)	X	X	B (in future)
Imposex/ Intersex in gastropods	X	X	C
Stress on Stress (SoS)	X	X	not required
Scope for growth	X	X	B
Externally visible fish diseases	X	X	A
Reproductive success in eelpout	X	X	A

A: BEQUALM

B: Between particular independent laboratories

C: QUASIMEME

D: BEAST

E: WGBEC

F: MEDPOL

Appendix 1. Ecosystem components of integrated monitoring and assessment of chemical and biological effects monitoring for contaminants

(Source: Report of ICES/OSPAR SGIMC 2011, and developed from 1.5.5.1 ICES Advice 2010, Book 1)

Fundamental aspects of the design of an integrated programme include key environmental matrices (water, sediment and biota), the selection of appropriate combinations of biological effects and chemical measurements and the design of sampling programmes to enable the chemical concentrations, the biological effects data and other supporting parameters to be combined for assessment. The basic structure of an integrated programme is illustrated in Figure 1.

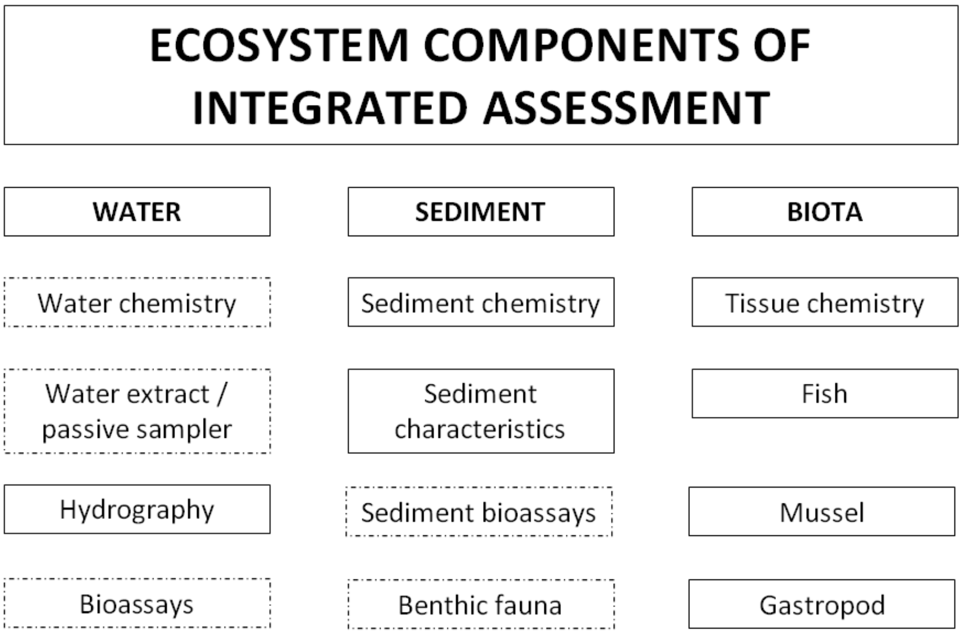


Figure 1. Overview of components in a framework for the integrated monitoring programme chemical contaminants and their biological effects. Solid lines – core methods, broken lines – additional methods.

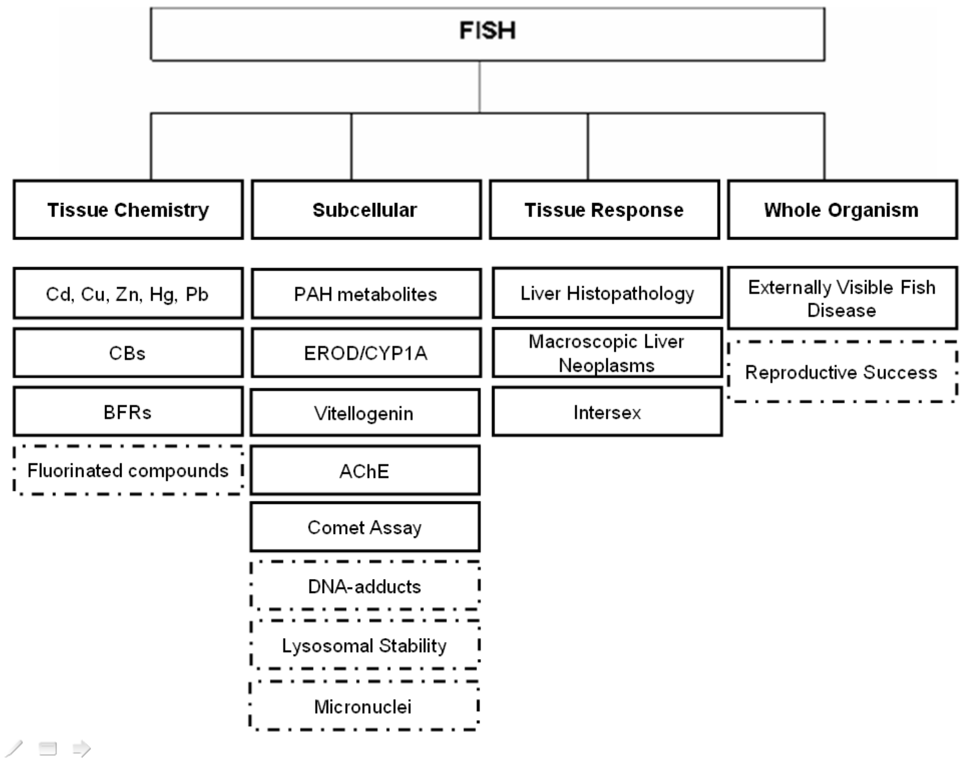


Figure 2. Methods included in the fish component of the integrated monitoring framework; solid lines – core methods, broken lines – additional methods.

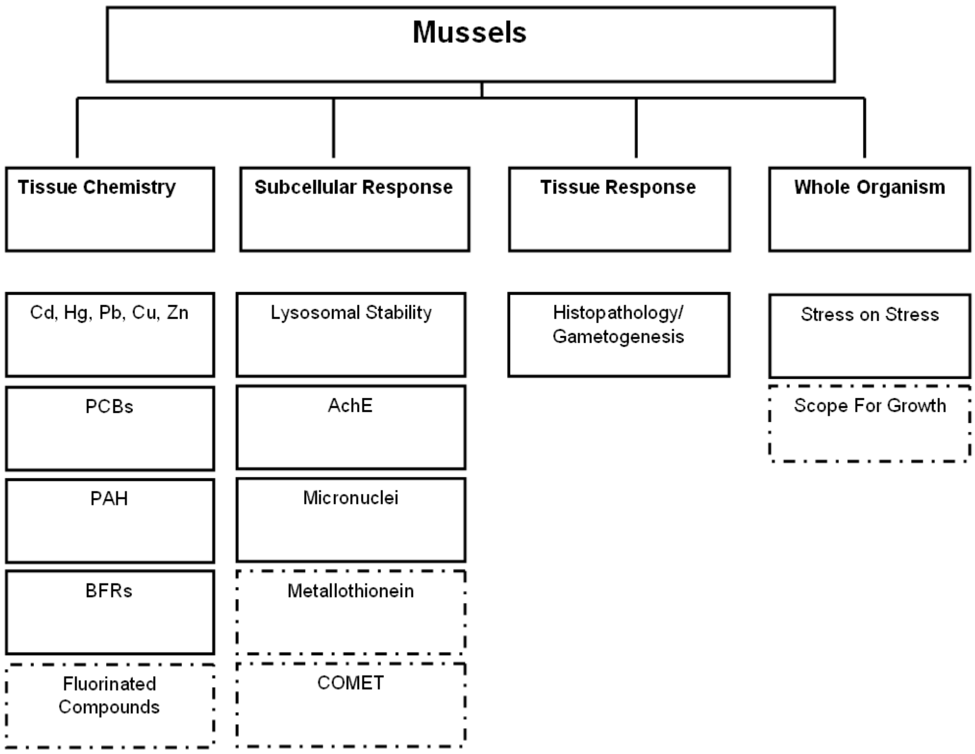


Figure 3. Methods included in the mussel component of the integrated monitoring framework; solid lines – core methods, broken lines – additional methods.

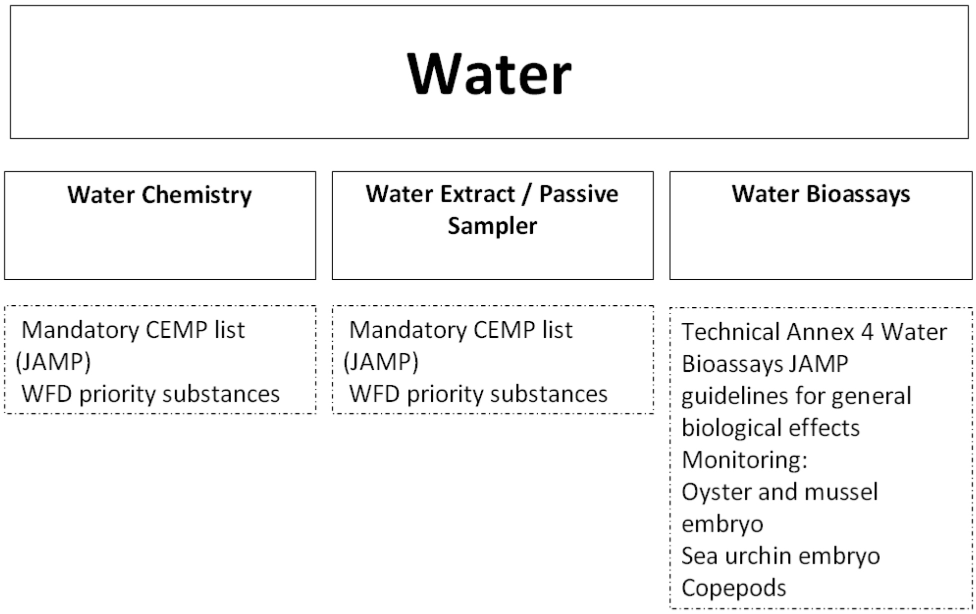


Figure 4. Methods included in the water component of the integrated monitoring framework; solid lines – core methods, broken lines – additional methods.

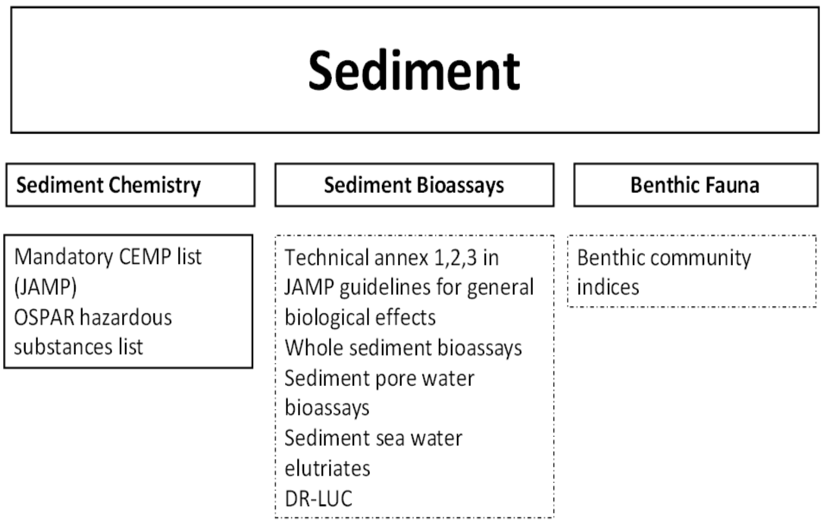


Figure 5. Methods included in the sediment component of the integrated monitoring framework; solid lines – core methods, broken lines – additional methods.

Figure 6. Methods included in the gastropod component of the integrated monitoring framework; solid lines – core methods, broken lines – additional methods.

Appendix 2. Assessment criteria for biological effects measurements

Values are given for both background assessment levels (BAC) and environmental assessment criteria (EAC), as available.

(Source: Annex 23 of the ICES/OSPAR SGIMC 2011 report).

Biological Effect	Applicable to:	BAC	EAC
VTG in plasma; µg/ml	Cod	0.23	
	Flounder	0.13	
Reproduction in eelpout; mean frequency (%)	Malformed larvae	1	
	Late dead larvae	2	
	Growth retarded larvae	4	
	Frequency of broods with malformed larvae	5	
	Frequency of broods with late dead larvae	5	
EROD; pmol/mg protein pmol/min/ mg protein S9 * pmol/min/ mg microsomal protein	Dab (F)	178	
	Dab (M)	147	
	Dab (M/F)	680*	
	Flounder (M)	24	
	Plaice (M)	9.5	
	Cod (M/F)	145*	
	Plaice (M/F)	255*	
	Four spotted megrim (M/F)	13*	
	Dragonet (M/F)	202*	
	Red mullet (M)	208	
PAHs Bile metabolites; (1) ng/ml; HPLC-F (2) pyrene-type µg/ml; synchronous scan fluorescence 341/383 nm (3) ng/g GC/MS * 1-OH pyrene ** 1-OH phenanthrene	Dab	16 (1)* 3.7 (1)** 0.15 (2)	 22(2)
	Cod	21 (1)* 2.7 (1)** 1.1 (2)	483 (3)* 528 (3)** 35 (2)
	Flounder	16 (1)* 3.7 (1)** 1.3 (2)	 29(2)
	Haddock	13 (1)* 0.8 (1)** 1.9 (2)	 35(2)
DR-Luc; ng TEQ/kg dry wt, silica clean up	Sediment (extracts)	10	40
DNA adducts; nm adducts mol	Dab	1	6

DNA	Flounder	1	6
	Cod	1.6	6
	Haddock	3.0	6
Bioassays; % mortality	Sediment, Corophium	30	60
	Sediment, Arenicola	10	50
	Water, copepod	10	50
Bioassays; % abnormality	Water, oyster and mussel embryo	20	50
	Water, sea urchin embryo	10	50
Bioassay; % growth	Water, sea urchin embryo	30	50
Lysosomal stability; minutes	Cytochemical; all species	20	10
	Neutral Red Retention: all species	120	50
Micronuclei; ⁰/₀₀ (frequency of micronucleated cells) 1 Gill cells 2 Haemocytes 3 Erythrocytes	*Mytilus edulis*	2.5 [1] 2.5 [2]	
	Mytilus galloprovincialis	3.9 [2]	
	Mytilus trossulus	4.5 [2]	
	Flounder	0.0-0.3 [3]	
	Dab	0.5 [3]	
	Zoarces viviparus	0.3-0.4 [3]	
	Cod	0.4 [3]	
	Red mullet	0.3 [3]	
Comet Assay; % DNA Tail	*Mytilus edulis*	10	
	Dab	5	
	Cod	5	
Stress on Stress; days	*Mytilus* sp.	10	5
AChE activity; nmol.min^{-1} mg prot^{-1} 1 gills 2 muscle tissue 3 brain tissue * French Atlantic waters ** Portuguese Atlantic waters + French Mediterranean Waters ++ Spanish Mediterranean Waters	*Mytilus edulis*	30 [1*]	21 [1*]
		26 [1**]	19 [1**]
	Mytilus galloprovincialis	29[1+]	20[1+]
		15 [1++]	10 [1++]
	Flounder	235 [2*]	165 [2*]
	Dab	150 [2*]	105 [2*]
	Red mullet	155 [2+] 75 [3++]	109 [2+] 52 [3++]
Externally visible diseases*** Ep, Ly, Ul Ep, Ly, Ul Ac, Ep, Fi, Hp, Le, Ly, St, Ul, Xc Ac, Ep, Fi, Hp, Le, Ly, St, Ul, Xc Ac, Ep, Hp, Le, Ly, St, Ul, Xc Ac, Ep, Hp, Le, Ly, St, Ul, Xc	Dab	 *F, 1.32*, 0.216 *M, 0.96*, 0.232 *F, 1.03*, 0.349 *M, 1.17*, 0.342 *F, 1.09*, 0.414	 F, *NA*, 54.0 M, *NA*, 47.7 *F, 50.6*, 19.2 *M, 38.8*, 16.1 *F, 48.3*, 21.9

Italics: ungraded, bold: graded NA: Not applied		*M, 1.18*, 0.398	*M, 35.2*, 16.5
Liver histopathology-non specific	Dab	NA	Statistically significant increase in mean FDI level in the assessment period compared to a prior observation period *or* Statistically significant upward trend in mean FDI level in the assessment period
Liver histopathology: contaminant-specific	Dab	Mean FDI <2	Mean FDI ≥ 2 A value of FDI = 2 is, e. g., reached if the prevalence of liver tumours is 2 % (e. g., one specimen out of a sample of 50 specimens is affected by a liver tumour). Levels of FDI ≥ 2 can be reached if more fish are affected or if combinations of other toxicopathic lesions occur.
Macroscopic liver neoplasms	Dab	Mean FDI <2	Mean FDI ≥ 2 A value of FDI = 2 is reached if the prevalence of liver tumours (benign or malignant) is 2 % (e. g., one specimen out of a sample of 50 specimens is affected by a liver tumour). If more fish are affected, the value is FDI > 2.
Intersex in fish; % prevalence	Dab Flounder Cod Red mullet *Zoarces viviaprus*	5	
Scope for growth Joules/hr/g dry wt.	Mussel (*Mytilus* sp.) (provisional, further validation required)	5	-2
Hepatic metallothionein μg/g (w.w.) [1] Whole animal [2]Digestive gland [3]Gills * Differential pulse polarography	*Mussel edulis*	0.6 [1*] 2.0 [2*] 0.6 [3*]	
	Mytilus galloprovincialis	2.0 [1*] 3.9[2*] 0.6 [3*]	
Fish Disease Index	Dab, flounder, cod, whiting, haddock	2.5% quantile	97.5% quantile

Histopathology	VVbas: Cell type composition of digestive gland epithelium; $\mu m^3/\mu m^3$ *(quantitative)*	0.12	0.18
	MLR/MET: Digestive tubule epithelial atrophy and thinning; µm/µm *(quantitative)*	0.7	1.6
	VVLYS & Lysosomal enlargement; $\mu m^3/\mu m^3$ *(quantitative)*	VvLYS 0.0002	V>0.0004
	S/VLYS: $\mu m^2/\mu m^3$	4	
	Digestive tubule epithelial atrophy and thinning *(semi-quantitative)*	STAGE ≤1	STAGE 4
	Inflammation *(semi-quantitative)*	STAGE ≤1	STAGE 3
Imposex/intersex in snails	Gastropod molluscs	See OSPAR adopted criteria	See OSPAR adopted criteria

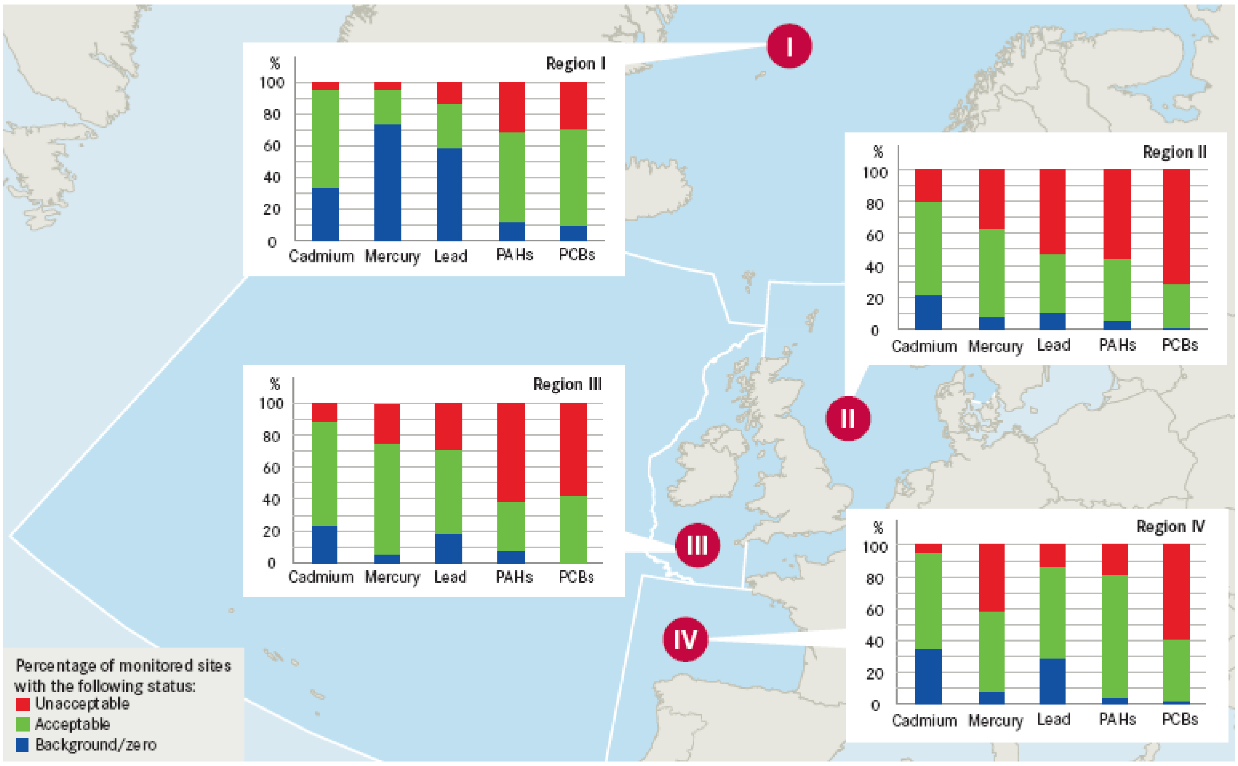
*****: Assessment criteria for the assessment of the FDI for externally visible diseases in common dab (*Limanda limanda*). Abbreviations used: Ac, *Acanthochondria cornuta;* Ep, epidermal hyperplasia/papilloma; Fi, acute/ healing fin rot/erosion; Hp, hyperpigmentation; Le, *Lepeophtheirus sp.*; Ly, lymphocystis; St, *Stephanostomum baccatum*; Ul, acute / healing ulcerations; Xc, X-cell gill disease.**

Full details of the assessment criteria and how they were derived can be found in the SGIMC 2010 and SGIMC 2011 and WKIMON 2009 reports on the ICES website and in the OSPAR Background Documents for individual biological effects methods.

Data for biomarkers in some northern fish species have been obtained through the IRIS BioSea JIP programme (funded by Total E&P Norge & EniNorge) and the Biomarker Bridges programme (funded by Research Council of Norway) and have been used to develop EAC and BAC values for arctic fish.

Appendix 3. OSPAR regional level integration of the concentrations of priority contaminants in fish, shellfish and sediment

(Source: OSPAR QSR 2010, Hazardous Substances chapter.)



Annex 25: Technical annex: Integrated assessment framework for contaminants and biological effects

The development of a framework with which to assess contaminant and biological effects data together is essential to the delivery of integrated monitoring and assessment. A multi-step process is proposed which follows on from experience of the assessment of contaminants data for sediment, fish and shellfish in OSPAR contexts. The process is informed initially by the individual assessment of determinands (contaminants or effects) in specific matrices at individual sites against the defined assessment criteria (BAC and EAC). Such assessment criteria for biological effects have been developed over recent years and are included in OSPAR Background Documents, and for contaminants have been used by OSPAR groups, for example in the QSR 2010. Initial comparisons determine whether the determinand and site combinations are <BAC (blue), between the BAC and EAC (green) or >EAC (red). This summarized indicator of status for each determinand can then be integrated over a number of levels: matrix (sediment, water, fish, mussel, gastropod), site and region and expressed with varying levels of aggregation to graphically represent the proportion of different types of determinands (or for each determinand, sites within a region) exceeding either level of assessment criteria.

Such an approach has several advantages. The integration of data can be simply performed on multiple levels depending on the type of assessment required and the monitoring data available. The representation of the assessment maintains all the supporting information and it is easy to identify the causative determinands that may be responsible for exceeding EAC levels. In addition, any stage of the assessment can be readily unpacked to a previous stage to identify either contaminant or effects measurements of potential concern or sites contributing to poor regional assessments.

This approach builds on the OSPAR MON regional assessment tool developed for contaminants. The development of BAC and EAC equivalent assessment criteria for biological effects, which represent the same degree of environmental risk as indicated by BACs and EACs for contaminants, allows the representation of these monitoring data alongside contaminant data using the same graphical representation approach. The inclusion of biological effects data to the system adds considerable value to the interpretation of assessments. Where sufficient effects monitoring data are available, confidence can be gained that contaminants are not having significant effects even where contaminant monitoring data are lacking. In instances where contaminant concentrations in water/sediment are >EAC, a lack of EAC threshold breach in appropriate effects data can provide some confidence that contaminant concentrations are not giving rise to pollution effects (due for example to lack of availability to marine biota). Similarly, the inclusion of effects data in the assessment framework can indicate instances where contaminants are having significant effects on biota, but have not been detected or covered in contaminant-specific chemical monitoring work.

Application to determination of GES for Descriptor 8 of MSFD

The assessment framework described below provides an appropriate tool for assessment of environmental monitoring data to determine whether Good Environmental Status is being achieved for Descriptor 8 of MSFD (concentrations of contaminants are at levels not giving rise to pollution effects). Determinands with EAC or EAC equivalent assessment criteria provide appropriate indicators with quantitative targets. The assessment of contaminant and effects monitoring data against these EAC-level assessment criteria provides information both on concentrations of contaminants likely to give rise to effects and the presence/absence of significant effects in marine biota.

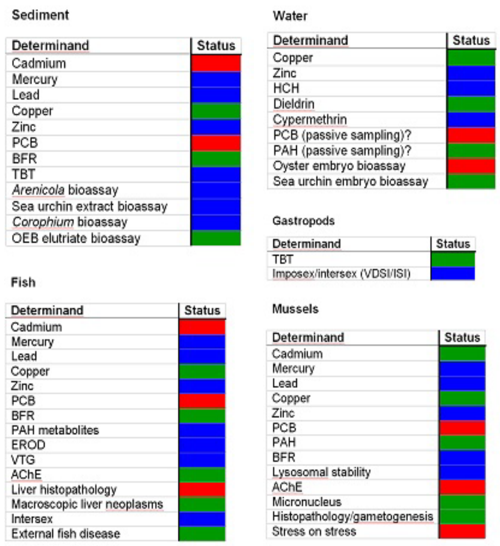
Due to the relatively large number of determinands monitored under the integrated approach, it is inappropriate to adopt an approach whereby EAC level failure of a single determinand results in failure of GES for a site or region. A more appropriate approach would involve the setting of a threshold (%) of proportion of determinands that should be <EAC to achieve GES. Such an approach would avoid the failure of sites or regions due to occasional outlying, erroneous results for particular determinands. The setting of an appropriate threshold for overall regional assessment for MSFD will require consideration and revision in the light of testing the framework described here with real monitoring data, however an initial threshold of 95% <EAC (to ensure that the vast majority but not all contaminants/effects measurements should be <EAC) is proposed here for the purposes of testing the system.

Example application of the integrated assessment framework

In order to best demonstrate how monitoring data (assessed against BAC and EAC) can be integrated for matrices, sites and regions and ultimately provide an assessment that could be useful for determination of GES for Descriptor 8, a worked example is provided below following a five step process.

Step 1 Assessment of monitoring data by matrix against BAC and EAC

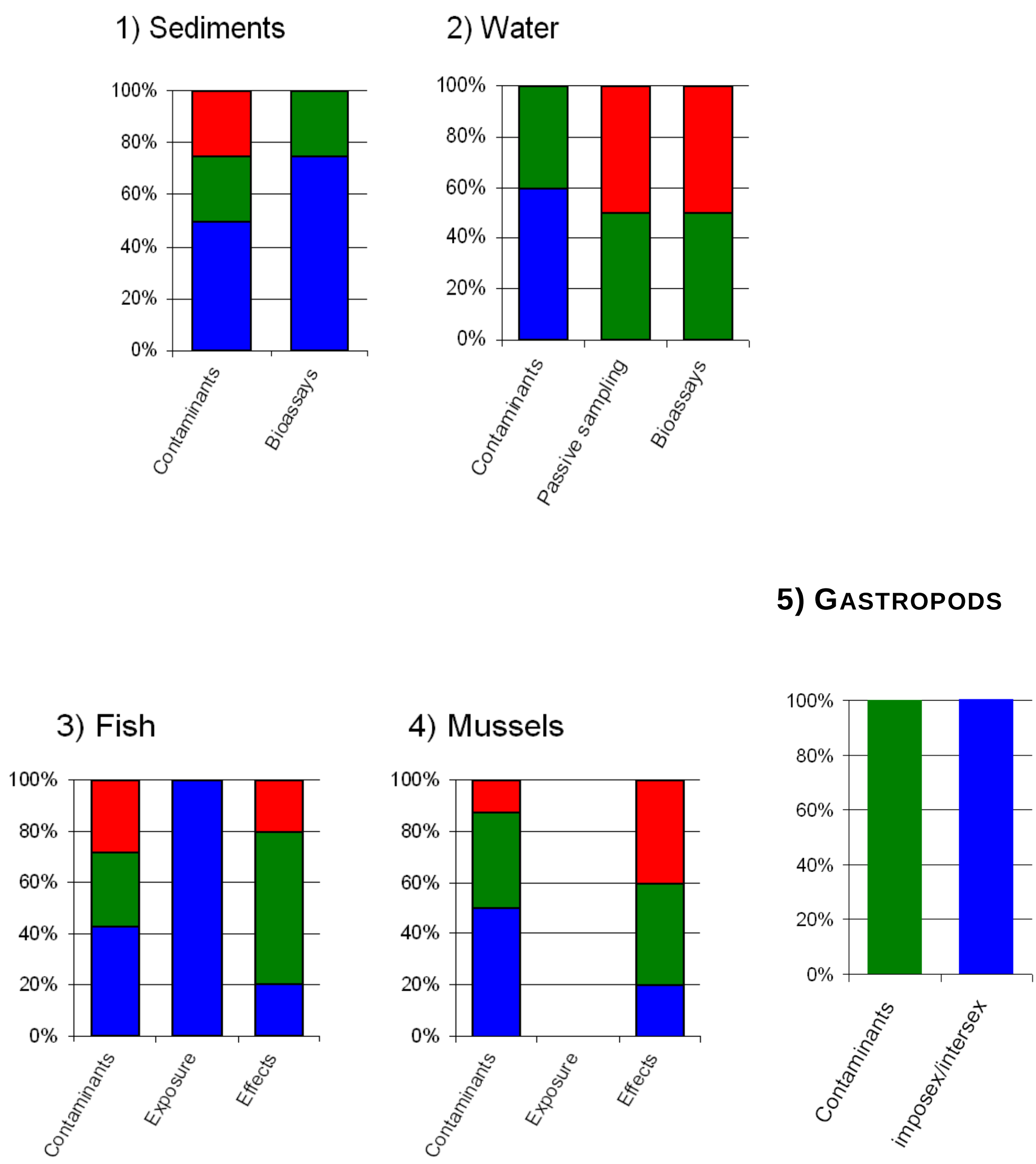
All determinands available for a specific site assessment are compiled with results presented by monitoring matrix and expressed as a colour depending on whether the value exceeds BAC or EAC. In the example provided below, determinands and their status are provided for illustrative purposes only, to show how subsequent integration can be performed. A red classification indicates that the EAC is exceeded, blue indicates compliance with the BAC, while green indicates concentrations or levels of effects are between the BAC and EAC.



Step 2 Integration of determinands by matrix for a given site

For each of the five matrices, the results of the individual determinand assessments are aggregated into categories: contaminants, exposure indicators, effects indicators and for sediment/water matrices also passive sampling and bioassay categories. It is necessary to separate the biological effects measurements into different categories depending on whether an EAC-equivalent assessment criterion (AC) has been set or not. Otherwise aggregated information on the proportion of determinands exceeding the separate AC will be incorrect. For simplicity, these categories have been termed 'exposure indicators' (where an EAC has not been set) and 'effects indicators' where an EAC (equivalent to significant pollution effect) has been set for the measurement. On subsequent aggregation / integration of these indicators across matrices for a specific site, bioassays are considered 'effects indicators' as EAC are available. It should be possible to include data from passive sampling in both the water and sediment schemes when assessment criteria have become available. They are nominally included in the example here to show how they could be included.

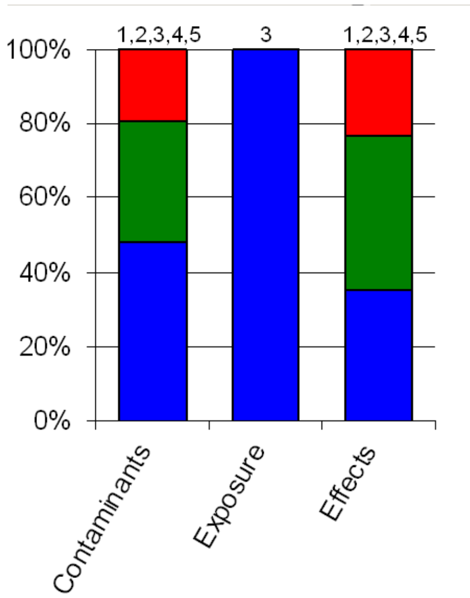
The integration by matrix and category of determinand can be expressed by tri-coloured bars showing the proportions of determinands that exceed the BAC and EAC as shown below. Note that for mussels in this instance, no exposure indicators are used, because all the biological effects measurements have EAC available.



Step 3 Integration of matrices for a site assessment

In order to express the results of assessment for a particular site simply, information can be aggregated across matrices and expressed by determinand category as shown below. In order to achieve this, results from passive sampling from sediment and water categories could be integrated into the contaminant indicator graphic and bioassays and gastropod intersex/intersex integrated into 'effects indicators'. Thus the outcome of assessment of all determinands from all matrices can be expressed for a whole site. For some assessments, this will be the highest level of aggregation required. However, for assessments covering larger geographical areas (subregional, regional, national, regional seas for MSFD, etc) where assessments need to be undertaken across multiple sites, a further level of integration is required (steps 4 and 5).

For transparency, each determinand grouping is labelled with the matrices from which it is comprised. Thus it can quickly be determined whether the site assessment is comprised of all or just a subset of the monitoring matrices. In the example below, all five matrices have been used to determine the overall site assessment, however only for fish (matrix 3) were there any effects measurements that did not have available EAC for assessment. Therefore the exposure indicators graphic is labelled to show that only matrix 3 contributed to the site assessment.

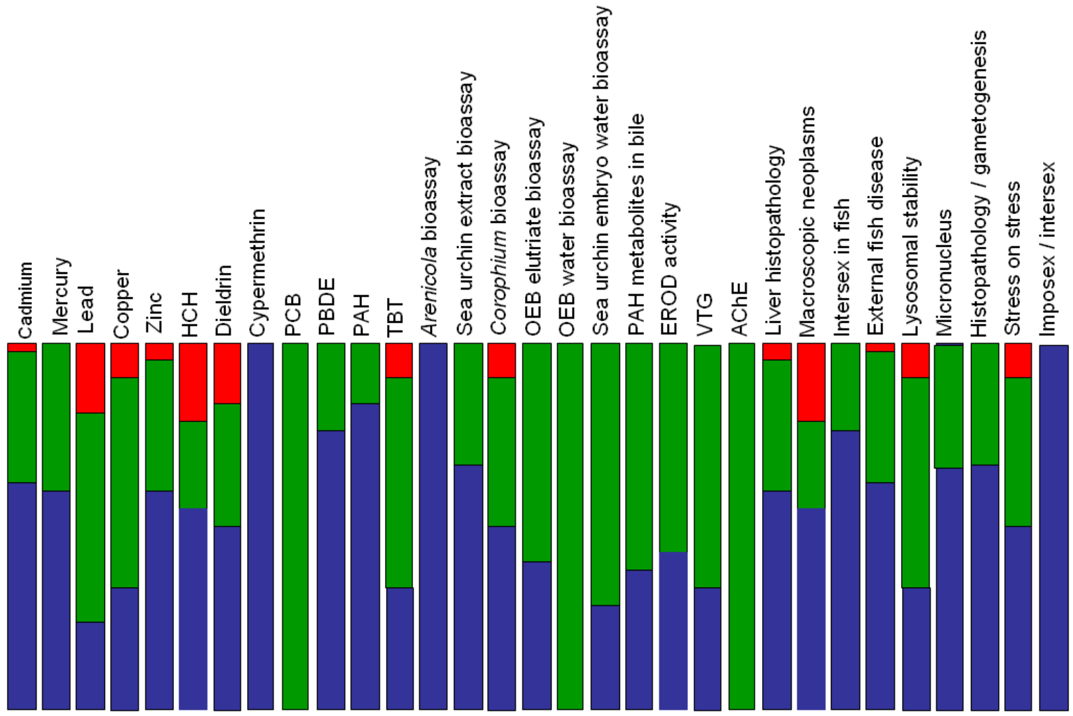


Step 4 Regional assessment across multiple sites

This can be done at multiple levels (aggregation of data at the subregional, regional and national levels) in different ways to express both the overall assessment of proportion of determinands (across all matrices) exceeding both assessment thresholds (BAC/EAC) (approach A) and by determinand for the region showing the proportion of sites assessed in the region that exceed the thresholds (approach B). Both approaches show the overall proportion of determinand/site incidences of threshold exceedence. However approach A shows most clearly which determinands are responsible for any EAC exceedence, while approach B shows a more aggregated, summarized representation of the same information by determinand category. Both can be constructed directly from the output of Step 1.

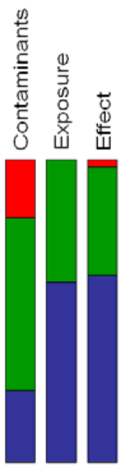
4A Regional assessment of sites by determinand

This shows a graphical representation of the proportion of sites falling into each status class for each determinand across all relevant matrices (many determinands are only relevant to one or some of the matrices).



4B Regional assessment of sites by determinand category

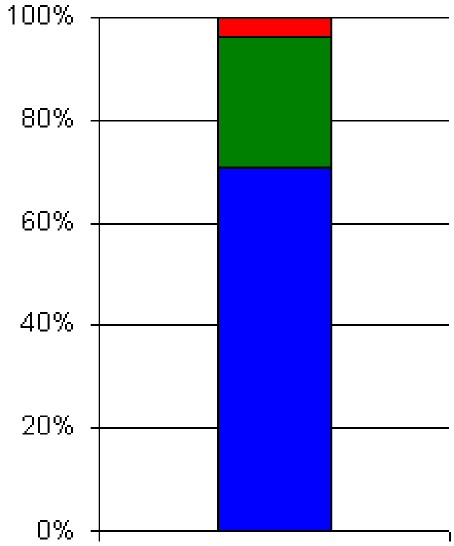
The above regional assessment can be summarized by determinand category as was demonstrated in step 3 for the site assessment and shown below.



Step 5 Overall assessment

The assessment by region can be aggregated further into a single schematic showing the proportion all determinands across all sites that exceed BAC and EAC. This can be used for the purposes of an overall assessment and it is proposed that a simple threshold figure (e.g. 95%) <EAC is used to determine whether Good Environmental Status for Descriptor 8 is met in this assessment. The overall assessment can be easily unpacked through the steps above to determine which sites and determinands (effects types or contaminants) are contributing to, for example, the proportion of red

(greater than EAC) data, and thereby potentially leading to failure to achieve GES for a region.



Conclusion

A potential assessment framework has been presented which integrates across contaminant and biological effects monitoring data and allows assessments to be made across matrices, sites and regions. It is simple and transparent and allows for multiple levels of aggregation for different assessment requirements. Such an approach has been used with success for a wide range of contaminants data in the OSPAR QSR 2010. It is proposed that this approach is tested with real monitoring data and could provide a suitable approach for the assessment of GES for Descriptor 8 of the MSFD.

Annex 26: Additional database parameters

Details of changes required to DOME to allow capture of data and supporting parameters for a range of biological effects measurements.

A number of bioassays now fulfil (or are close to fulfilling) OSPAR CEMP requirements and should be made available for reporting to the ICES database DOME. In order for this to be possible, it is necessary that DOME is configures appropriately. A series of conversations were held between SGIMC and the ICES DataCentre to identify the additional capability that will be required of DOME in order to accept data on:

a) DR-Luc

b) COMET Assay

c) Micronucleus assay

d) Stress on stress

e) Mussel condition

f) Histology

g) Parasites in organisms

DR-Luc is a biological technique which measures the effect of dioxins and furans in sediment on in vivo cell lines by measuring light production expressed in pg/g/TEQ. The technique has been described in a TIMES document.

- Some bioassays use *in vivo* organisms, bacterium or cell lines for testing toxicity. *In vivo* organisms and bacterium can already be reported in the Environmental Reporting Format (ERF3.2) method record 21: field SPECI. DR-LUC uses cell lines. Cell lines should be added to options for 21:SPECI.
- The following should be added to RECO:
  - Cell lines
    - yeast lac-Z ERα cDNA
    - T47 D human receptor
    - H411E
    - Bg1LUC4E2
  - Parameter: light production
  - Measurement unit: picogram/gramme/toxicity equivalent (TEQ)
  - METOA options: “destructive” and “non-destructive”
- DATSU checks:
    - Light production can only be reported with MUNIT picogram/gramme/toxicity equivalent (TEQ)
    - Only sediment matrices SEDtot, SED2000 or SED63 are allowed
    - Dry Weight basis must be reported for light production at picogram/gramme/toxicity equivalent (TEQ)
    - METOA “destructive” or “non-destructive” is mandatory for light production.
    - Parameter “Organic carbon” is needed for calculations and is therefore mandatory when the DR-LUC parameter is reported

COMET assay. TIMES not available. When available, the following will be added:

- Field description expansion for METFP to include lysis
- The following should be added to RECO:
  - METFP: PH xx (for alkalinity)
  - MATRX:
  - PARAM: % DNA in tail
  - Possible PARAMs for legacy data:
    - Tail length
    - Tail moment
- DATSU checks:
  - MUNIT % for %DNA in tail

Micronucleus assay

- The following should be added to RECO:
  - MATRX:
    - erythrocyte (ER)
    - haemolymph
  - PARAM:
    - micronuclei (MNC) linked to Pargroup B-MBA
    - Cofactor micronuclei-number of cells counted (MNC-QC-NR)
      - ER,MNC,3,NR/1000 cells,
      - ER,MNC-QC-NR,4000,NR,
  - MUNIT: number per thousand cells (nr/1000 cells)
- DATSU checks:
  - PARAM: TEMP to be reported in 92 record – mandatory
  - MNC-QC-NR >1000

Stress on stress response in mussels

- The following should be added to RECO:
  - PARAM:
    - LT50 Lethal threshold linked to B-TOX
    - TMM time to maximum mortality
- DATSU checks:
  - MATRX – WO for both parameters
  - MUNIT – days to be reported with TMM
  - LT50 MUNIT=D
    - WO,LT50,D,15,
    - WO,TMM,D,20
    - TMM>LT50

Mussel condition index

- The following should be added to RECO:
  - PARAM:
    - Internal shell volume in ml

        - Mussel condition index A = 100 x (dry flesh wt/live whole wt)
        - Mussel condition index B = 100 x (dry flesh wt/wet flesh wt)
        - Mussel condition index C = 100 x (dry flesh wt/internal shell volume)
        - Mussel condition index D = ratio of length:width/dry wt
    - MUNIT unitless–ratio?
- DATSU checks:
    - Length – LNMIN/LNMEA/LNMAX with MUNIT mm
    - Wet weight mandatory

What technique? Histology in general?

- The following should be added to RECO:
    - PARAM: Apidogranular cells in vesicular connective tissue
        - DATSU checks
            - Stage 0–4
            - MATRX Gonad
    - PARAM:
        - Degenerative lesions
        - Connective-to-diverticular ratio (CTD)
            - DATSU checks
                - MATRX Digestive Gland
                - MUNIT ratio?
    - PARAM: Digestive tubule epithelial atrophy and thinning
        - Measured as mean proportion of tubule width (MPTW) as follows
        - Mean luminal radius (MLR) / mean epithelial thickness (MET) µm/µm
        - MET / mean diverticular radius (MDR) µm/µm
        - Digestive tubule epithelial atrophy and thinning is an index from 0–4
    - PARAM: Lysosomal alterations (Matrix digestive gland)
        - Lysosomal enlargement (VvLys) µm3/µm3
        - Surface to volume ratio of lysosomes (S/VLys) µm2/µm3
        - Intracellular neutral lipid accumulation (INLA) measured as: Volume density of neutral lipids (VvNL µm3/ µm3

B-GRS codes Parasites in organisms

- WGPDMO want raw data
- Matrix whole organism for mussels
- The following should be added to RECO:
    - PARAM with MUNIT stage:
        - Trematode sporocyst infection – need "infection"?
        - Intracellular ciliates and protistans - Why as one?
            - DATSU checks
                - MATRX=WO

- ♦ Stage 0–4
- ♦ Allow stages for these 2 diseases
- ♦ Remove NODIS requirement based 03SPECI – if SPECI=MYTILUS

- PARAM with MUNIT affected number:
  - *Martelia refringens*
  - *Mytilicola intestinalis*
  - *Ancistrum mytili*
  - *Ciliophora* sp
  - *Steinhausia mytilovum*
  - *Haplosporidia* sp
  - *Bucephalid* sp
  - *Rickettsia* sp
  - *Chlamydia* sp
    - DATSU checks for diseases reported with MUNIT AFNR
      - ♦ MATRX=WO
      - ♦ MUNIT: Disease units are reported as AFNR
      - ♦ Remove NODIS requirement based 03SPECI – if SPECI=*MYTILUS*
      - ♦ Pool check: If R04.NOINP>1 and R10.MUNIT is "AFNR" then R10.VALUE must be <= "NOINP"

?? WGPDMO : Intersex in fish needs parameter code under B-HST called gonadal histology. Intersex to be reported as % prevalence in a population of individuals. Matrix is gonad. Female tissue on male gonads – *ovitestis*.

Annex 27: Update of SGIMC Workplan from 2010/2011

Review of progress against Work Programme for SGIMC from January 2009 to March 2011 (from SGIMC 2010 report).

EFFECT	TASK	RESPONSIBLE MEMBER	WHEN	REPORT TO	STATUS JANUARY 2010
EROD	Organising Workshop	Ian Davies	For Oct 2009	SGIMC 2010	Completed
EROD	Update Background Document and develop improved approach to Background Response assessment criteria	Ian Davies and others	Oct 2009	SGIMC 2010	Completed
PAH bile metabolites	Update Background Document	Dick Vethaak and Ketil Hylland	Mar 2009	WGBEC 2009	Completed
PAH bile metabolites	To develop Background Response assessment criteria	Ketil Hylland (and Ian Davies)	Oct 2009	SGIMC 2010	Completed
DNA adducts	Update Background Document	Brett Lyons (and Ian Davies)	Oct 2009	SGIMC 2010	Completed
DNA adducts	To develop Background Response and EAC-equivalent assessment criteria	Brett Lyons (and Ian Davies)	SGIMC 2011	To be reviewed by SGIMC 2011	Completed
DR-CALUX	Prepare Background Document	Dick Vethaak	September 2010	SGIMC 2011	Completed
DR-CALUX	To develop Background Response and EAC-equivalent assessment criteria	Dick Vethaak	Oct 2009	SGIMC 2011	Completed
DR-CALUX	Complete TIMES series method document	Dick Vethaak	March 2010	SGIMC 2011	In progress; will be sent to editor spring 2011

EFFECT	TASK	RESPONSIBLE MEMBER	WHEN	REPORT TO	STATUS JANUARY 2010
Liver nodules (neoplasm)	To develop Background Response and EAC-equivalent assessment criteria	Thomas Lang and Dick Vethaak	Oct 2009	SGIMC 2010	Completed
Extraction procedures for bioassay methods	Complete TIMES series method document	Dick Vethaak + John Thain	Imminent	WGBEC 2011	Completed
VTG	Establish BAC in monitoring species	Ian Davies and Dick Vethaak	For SGIMC 2010	SGIMC 2010	Completed
VTG	Develop EAC equivalent for monitoring species	Ian Davies and Dick Vethaak	SGIMC 2011	SGIMC 2011	SGIMC 2011 decided not currently possible
Intersex in fish	Review Background document	Steve Feist	WGBEC2011	SGIMC 2011	Completed
Intersex in fish	To develop Background Response and EAC-equivalent assessment criteria	Steve Feist (and Ian Davies)	SGIMC 2010	SGIMC 2011	Completed
Fish Disease Index.	No action required by SGIMC		No action required by SGIMC		
Reproductive success (eelpout).	Review BG document and TA		SGIMC 2010	SGIMC 2010	Completed
Background document on Supporting parameters in fish: condition indices, SLI and SGI	develop BD	John Thain, Dick Vethaak	WGBEC 2011	SGIMC 2011	Completed
Lysosomal stability (Neutral Red)	Organising training workshop Draft proposal Permission from ICES/OSPAR	Concepcion Martínez-Gómez	June 2010	SGIMC 2011	Completed and reviewed
Acetyl cholinesterase	Update Background Document	Thierry Burgeot	WGBEC2011	SGIMC 2010	Completed

EFFECT	TASK	RESPONSIBLE MEMBER	WHEN	REPORT TO	STATUS JANUARY 2010
Acetyl cholinesterase	To develop Background Response assessment criteria	Thierry Burgeot	WGBEC 2011	SGIMC 2011	Completed
Mussel histopathology	ICES Times manuscript incl uding BAC in preparation	Steve Feist + Miren Cajaraville	WGBEC 2011, WGPDMO 2011	SGIMC 2011	Reviewed and to be submitted for publication
Micronucleus assay + comet assay	Background document and draft BAC	Brett Lyon	WGBEC2011	SGIMC 2011	Completed
MT & ALA-D	Develop BC using recent data	Ketil Hylland	Feb 2011	SGIMC 2011	Completed
New Chapter	In vitros YES/YAS, ER CALUX	JT / DV	WGBEC 2011	SGIMC	Task ransferred to WGBEC
Chapter 8	Add Sed & SW elutriate bioassays for invert embryos. Further validate others as more data becomes available	Ricardo Beiras	2010	WGBEC 2011	Completed
Chapter 9	As above with copepods	As above	As above	WGBEC 2011	Completed
Chapter 10	Update BG and ass cri for Whole sediments with amphipods as more data becomes available	As above	As above	WGBEC 2011	Completed

Annex 28: Technical minutes by RGMON1

- Report of the view Monitoring Review Group (RGMON1) on the Report of the Study Group on Integrated Monitoring of Contaminants and Biological Effects (SGIMC), 2011.
- deadline 2 May 2011
- Members RGMON1: Jacob de Boer, The Netherlands (Chair), Les Burridge, Canada, Pekka J. Vuorinen, Helsinki, Finland.

Introduction

The task of this review group was to evaluate the Report of the Study Group on Integrated Monitoring of Contaminants and Biological Effects (SGIMC), 2011 and to assess if it meets the request of OSPAR on this matter, worded as follows:

To complete the development of JAMP guidance for integrated monitoring of chemicals and their biological effects through preparing technical annexes on:

(i) Survey design. The purpose is to provide guidance on the selection of representative stations, taking into account requirements under the Water Framework Directive and the proposed Marine Strategy Directive, and for the selection of stations for integrated monitoring. This work should build on work by WGSAEM 2007 relating to the spatial design of monitoring programmes and should take into account the approach taken by the UK in re-designing their station network;

(ii) Groups of biological effects methods to be deployed to address specific questions. This should provide guidance on recommended packages of chemical and biological effects for monitoring on determinand basis to ensure that chemical and biological methods were well matched and that chemical analysis underpinned biological effects monitoring.

Apart from evaluating the Study Group Report against the request of OSPAR, the evaluation was focused on the technical correctness and scientific quality of the Report.

Method of evaluation

The three members of the review group have all read the main part of the report and the Annexes 1–3. The rest of the annexes were equally distributed over the three members. After an e-mail consultation, this review report was prepared by the chair and agreed upon by the two other members.

Review results

The Study Group Report is a comprehensive, good quality report that gives advice on integrated, simultaneous monitoring of contaminants and biological effects in marine water, sediments and biota. It is the outcome of scientific discussions during several years and various meetings of the Study group. The documents are generally well written and technically sound. It is based on a large number of scientific publications and reports, which are all mentioned in the various annexes. There are 27 annexes, a number of which contain a detailed description of a biological effects method. Other annexes deal with assessment criteria, recommendations and several other, more general documents. A number of annexes have no clear recommendations, but are either technical background documents or reflections of meeting activities.

As regards the first item of the OSPAR request, the review group concludes that this aspect was not taken into account. There is very little information on this topic in the report. To meet the request of OSPAR, this information on survey design and selection of sampling stations still needs to be carried out.

The Study Group Report deals extensively with the second part of the OSPAR request, i.e. recommending biological methods and groups of biological methods to be deployed and recommending combinations of chemical analyses and biological effects methods. A large number of suggestions are given on the application and combination of such methods. The Report also gives information on the extent of coverage of biological effects methods by chemical analysis. Clearly, not all methods that are available and not all methods mentioned in the Report are fully validated. For a number of biological methods further work on calibration and validation is needed. The Report is, however, conservative enough in recommending biological methods and combinations of these with chemical analysis.

There is considerable overlap in text and figures of the Annexes 20 and 21. It would help the reader of the report when the overlap was taken out and the two parts would be integrated. In chapter 6 (page 14) it is recommended to send the two updated annexes 9 and 10 to OSPAR with a request to update the older versions. It is unclear if the Annex 21 already includes the updated versions of the text and figures. Therefore, it is recommended to take out Annex 20, adjust the text of Chapter 6 and ensure that the overarching document Annex 21 contains the most actual information.

As regards the organic contaminants mentioned in Annex 20, it may be worth considering for the longer term to follow the developments in the Stockholm Convention on Persistent Organic Pollutants. This would lead to international harmonization and ensure that the most persistent contaminants are being monitored within ICES and OSPAR. List of official persistent organic pollutants (POPs) is regularly updated (ca. once per two years) and currently covers most of the organic contaminants mentioned in Annex 20.

The Review group considers the report very useful and recommends it to be sent to OSPAR, noting that the survey design has not been taken into account. The Review group found a number of specific comments, which are given below.

Detailed comments: (these comments have been accommodated in the final version of the report).

Title page: The Recommended format for purposes of citation says: 2044, this should presumably be 2011.

Page i: 8.1. Concentrations (plural).

Page iii: Annex: no title? Annex 26...: Bring in line.

Page 7, 2nd paragraph: sentence unclear, see suggestion or improvement in Track change version of report.

Page 13, 4th paragraph, it is mentioned that lysosomal membrane stability method (LMS) would be in use in monitoring in Finland. However, in Finland actual biological monitoring is not yet conducted, but the method is being tested in ongoing projects. The same is true for Sweden, too. This was checked at the Finnish Environment Institute.

Annex 4. pages 37–47

In this protocol the units should be written in conformity throughout, and preferably like ml, l, ml/l etc. It has been corrected in the report with track changes.

Page 39, fish bile samples 2, comment-column: it should be beta-glucuronidase-arylsulfatase

Page 43, 3.3. One additional characteristic to be tested for a solvent should be if it e.g. potentiates effects or toxicity. In testing acute lethal toxicity of a slimecide to fish we had to use acetone as a solvent, which we, however, noticed to increase toxicity.

Annex 5. pages 48–56

Page 50, 3rd paragraph, in order to better understand the technique and calculation of the OSI, some more text from Bateman *et al*., 2004 should be copied and inserted: "In order to assess the distribution of oocytes throughout the testis, all specimens examined, WERE STEP-SECTIONED LONGITUDINALLY AT 0.2-MM INTERVALS THROUGHOUT TISSUE at a thickness of 3 to 5 μm, mounted…"

Page 53, 2nd paragraph, would it have been possible to write some more about relationship between plasma vitellogenin concentration and intersex than is mentioned in the sentence: "Complications in specifically linking the presence of a chronic marker (such as intersex) with more acute phase markers (such as VTG), or…".

Intersex examination is a laborious job and, as judged, needs quite a large sample size (59 individuals). Thus a comparison of sample size to that needed for other biomarkers would be warranted.

Annex 6. pages 57–65

Page 57, Heading A. Should be HSI, not HIS.

Page 57, 1st paragraph, on sixth line it is said that gonad weight is related to whole body weight that is contradictory with the GSI-formula on page 59.

Page 57, summary table, I don't see a reason for a suggestion to determine age in ten fish, while the other parameters are measured in individuals. Measuring age from ten fish in this context doesn't increase information or provide any explanations because growth of fish in the same population may fluctuate from year to year, and age-size relationships should be available from longer time periods and for different populations with which a single observation could be compared.

Page 60, Measurement of LSI, apparently one should record liver weight and not gonad weight in this context. Also in the formula liver weight, not gonad weight should be subtracted.

Page 60, 19. Determination of age, instead of otolith, age can be more easily determined in scales or some other bone depending on fish species.

Page 60, last line, should be "relevant" and the sentence "Ideally age should be ..." only partially describes the situation and could be replaced by "Ideally age-size (length and weight) relationship should be known for several populations of fish species for longer time periods, because growth of a fish species may vary in different populations/sites and from year to year."

Page 61, 20. Interpretation of data, in comparison to the preceding test it would be more logical to write LSI (HSI).

Annex 7. pages 66–74

As such this background document is again well compiled. However, a viewpoint of crucial importance is missing, and it is of concern of all, especially enzymatic biomarkers in poikilothermic animals. Many enzymatic biomarkers originate from human medicine, and optimized test kits are commercially available. These necessarily don't function properly for samples from poikilothermic animals, because enzyme activity is determined at room temperature (25°C) or even at a higher temperature. Poikilothermic aquatic animals, however, face fluctuating environmental conditions, e.g. seasonally changing temperature, and the animals physiologically adapt to environmental changes. This means, for example, synthesis of iso-enzymes of different characteristics to cope needs in a changed environment. Of course it is convenient, e.g. to measure enzymic activity in a constant temperature that also makes possible to compare results and organize intercalibrations between laboratories. Results from such measurements, however, necessarily don't tell about real status of the animal. If, for instance, seasonal cycle in enzymic activity would be measured in a constant assay temperature and at the environmental temperature, at which the animal was sampled, the resultant seasonal activity cycles could be quite different. What makes it still more complicated is that different iso-enzymes may have different optimal substratum concentrations, different pH optima, etc. All the above should be considered if AcHE-activity measurement is used to detect minor changes, but the method as presented apparently works with environment accidents and higher level pollution occasions, when enzymic inhibition is large.

Annex 8. pages 75–87

Page 76, 2nd paragraph, it could be mentioned that pathogens may also make organisms more vulnerable to pollutant effects.

Page77, last line and 78, first line, special characters are not printed properly.

Annex 9. pages 88–109

Page 90, 6. Target tissues, the second sentence is inaccurate and should be corrected. Of the listed tissues, only kidney and specifically head or cephalic kidney (as correctly referred in Barsiene *et al*., 2006a), is a haematopoetic tissue, and others are not. Thus the sentence could be written as: "There are other studies (albeit limited) available describing the use of blood cells of fish in other tissues, such as liver, kidney and gills (Baršienė *et al*., 2006a; Rybakovas *et al*., 2009), and also other cells like fin cells (Archipchuk, Garanko, 2005)."

Page 102, 2nd paragraph, last line, should there be referred to Table 5 and not Table 1? Where is Table 4?

Annex 10. pages 110–118

This is well-balanced good background information. It clarifies shortcomings of techniques. Some minor comments in the text are shown with track changes in the full report. References appear to be recent and relevant, although there was not enough time to check if all references were cited or if all citations were referenced.

Annex 11. pages 119–126

Many of the references appear to be older. This may reflect the current state of knowledge.

The report seems balanced, also identifying shortcomings of techniques.

The report would be greatly improved with a bit more background information in the form of an introduction so that the uninitiated could get some basic understanding of the procedures. For example, in pt #6 a clearer description of isolation of exposure water would make it much easier to understand the difficulties and complications involved.

Annex 12. pages 127–129

This annex is rather short and lacks a reasonable introduction. Other annexes/background reports provide more detail and are, therefore, more informative. Authors say that UNEP and EPA have standard methods. Some details could be provided.

Annex13. pages 130–133

The background document also seems quite short and lacking in detail compared to others. The authors but does identify need for a standard test organism in Europe. The most recent reference is 1998.

Annex 14. pages 134–143

This is a well written background report. It is very informative, technically sound and well referenced. The paper is almost exclusively about PAHs and other carcinogens are not given much focus. The authors should address this or simply say they will be discussing PAH –induced DNA adducts.

Annex 15. pages 144–152

This is a thorough and well-written document. There seems to be a discrepancy between text on page 148 and the Table 1. Page 152 regarding the derivation of an AC for DR-luc.

Annex 16. pages 153–159

Good background report. Balanced and well-referenced. The author's discussion of differences between various analytical methods would be better if they clearly recommended a technique. The table legends should include the analytical technique employed.

Annex 17. pages 160–178

This a fairly lengthy background report. It is technically sound and well referenced but it could be shortened without losing much information. It is also clear that much of the report was cut and pasted from other documents. One paragraph has actually been placed in the document twice.

The authors do provide clear recommendations, which is commendable.

Page 160: The only ref. To the BEQUALM guidelines is a website. The contents of a website may change in future. Therefore, an additional reference would be useful. The TIMES Series document can maybe be used as soon as it appears (has it appeared already – 2010?).

Figure 2 should be re-worked with a descriptive title. It isn't very clear what point is being made by the authors.

Annex 18. pages 179–194

Page 179: 2nd par. Ref. to BEQUALM work should be provided; 4th par: modifications made after 2009 workshop?

Page 180: 2nd par: Some modifications: should be made more specific.

Page 185: 1st par: details and availability of the CD ROM should be mentioned.

Annex 19. page 195

This annex shows an agreement of the Study Group on publishing this report in the ICES CRR Series, which seems a good idea.

Annex 20. pages 196–214

See text at the beginning of this review report. The overlap in Annexes 20 and 21 is rather confusing. Page 197, Annex 20: the numbers of the Annexes 9 and 10 under Annex 20 may lead to confusion.

Annex 21. Pages 215–226

See comment on Annex 20.

Annex 22. page 227

Supporting documentation summary.

Annex 23. pages 228–232

Assessment criteria for biological effects measurements. References to how these were derived are given.

Annex 24. pages 233–245

Again, overlap is found with Annex 21. Also the reference ANNEX 7 at the top of the page is confusing.

Annex 25. Pages 247–253

An assessment framework is presented which integrates contaminant data and biological effects and allows assessments to be made across matrices, sites and regions. The authors recommend further testing. Apparently, this is not at a stage that it could be included in the guidelines for biological effects monitoring.

Annex 26. pages 254–257

Page 250: question marks at 3rd line from bottom: where do they refer to?

Annex 27. pages 258–261

Overview of Workplan. No comments.