Contaminants in marine ecosystems: developing an integrated indicator framework using biological-effect techniques

John E. Thain, A. Dick Vethaak, and Ketil Hylland

Thain, J. E., Vethaak, A. D., and Hylland, K. 2008. Contaminants in marine ecosystems: developing an integrated indicator framework using biological-effect techniques. – ICES Journal of Marine Science, 65: 1508–1514.

Input of contaminants is an important pressure in most urbanized coastal areas, but establishing appropriate indicators of their presence and effects has been challenging. Such indicators would, at the very least, have to integrate chemical and biological data. One difficulty has arisen because the measurements provide information on different levels of biological organization (gene level up to community), although it is not obvious how this information could be conceptually linked. In addition, there are complicating factors, such as the differing ecological relevance of measurements, natural variation, confounding factors, and knowledge of background level or responses for each method. The challenge of how to take these issues forwards is discussed in light of current scientific thinking and of meeting international obligations. First, an integrated approach must be developed to using biological-effect techniques with chemical measurements, and second, assessment tools are required. Proposals for both of these have been initiated by ICES and OSPAR working groups and workshops. Concomitantly, steps have been taken to develop integrated assessment tools on a national basis. These show promise but highlight the difficulties of using biological-effect measures as indicators of ecosystem health.

Keywords: assessment criteria, biological effects, contaminants, integrated approach.

Received 23 December 2007; accepted 21 May 2008; advance access publication 28 July 2008.

J. E. Thain: Centre for Environment, Fisheries and Aquaculture Science, Weymouth Laboratory, Barrack Road, The Nothe, Weymouth, Dorset DT4 8UB, UK. A. D. Vethaak: Deltares, Marine and Costal Systems, PO Box 177, 2600 MH Delft, The Netherlands. K. Hylland: Department of Biology, University of Oslo, PO Box 1066, Blindern, N-0316 Oslo, Norway. Correspondence to J. E. Thain: tel: +44 1305 206620; fax: +44 1305 206601; e-mail: john.thain@cefas.co.uk.

Introduction

Over the past decade, an increasing number of techniques to measure the biological effects (e.g. bioassays, biomarkers, and disease) of contaminants has been incorporated into national and international monitoring activities, including the JAMP/CEMP (Joint Assessment Monitoring Programme/Co-ordinated Environmental Monitoring Programme) of the Oslo–Paris Commission (OSPAR). In relation to hazardous substances, the JAMP/CEMP monitoring activities are carried out to measure and monitor the quality of the marine environment and to assess progress towards realizing OSPAR objectives, viz.:

(i) What are the concentrations in the marine environment and the effects of the substances on the list of chemicals of concern (hazardous substances) identified by OSPAR? Are they at, or approaching, background levels for naturally occurring substances and close to zero for man-made substances?

(ii) Are any problems emerging related to the presence of hazardous substances in the marine environment? In particular, are any unintended/unacceptable biological responses being caused by exposure to hazardous substances?

OSPAR and the Helsinki Commission (HELCOM) have agreed on an ecosystem approach to managing the marine environment, which tries to understand and assess the interactions between, and the impact of, human activities on biota, so that appropriate measures can be taken. This development requires an integrated approach using suitable indicators of marine ecosystem health.

Many strategies and approaches have been advocated to assess ecosystem health in the recent scientific literature (EEA, 2001; IOC, 2002; Jorgensen *et al.*, 2005). Biological-effect methods are important elements in environmental monitoring programmes, because they can indicate links between contaminants and ecological responses. Biological-effect monitoring can thus be used to indicate the presence of substances, or combinations of substances, that had not been identified previously as being of concern, but also to identify regions of decreased environmental quality or reduced ecosystem health. In this context, environmental quality could be regarded as comparing something (i.e. a measurement, value, or score) against or relative to a standard (e.g. background or reference value), whereas ecosystem health would involve some measure of "well-being". However, usually, aligning the science with the aspirations of those demanding this type of "product/output" is not always straightforward or indeed easily obtainable. Any assessment of ecosystem health assumes that the basic requirements for the assessment are in place: data are available, assessment tools are available, and the components within the ecosystem are interrelated. In most instances, however, the linkages between different components of the ecosystem are not well understood or have at best been poorly addressed by science. This is no less so in developing indicators of marine ecosystem health. Some good datasets may exist in certain marine science disciplines, but in general, there are a number of shortcomings: limited availability of techniques and data; limited geographic coverage and few temporal datasets; and often where datasets do exist, mechanistic links between components of the ecosystem are poorly understood or recognized; and there are few criteria available for assessing data. In the context of biological effects, these are certainly important issues and need to be addressed.

We describe an approach currently being pursued by the ICES Working Group on Biological Effects of Contaminants (WGBEC; ICES, 2006) and the workshop on Integrated Monitoring of Contaminants and their Effects in Coastal and Open-Sea Areas (WKIMON; ICES, 2007c) to integrate chemical and biological effects in a common framework that assesses the pressures from contaminants, using relevant ecosystem components, i.e. sediment, water, and biota.

What are biological effects?

A biological effect may be defined as the response of an organism, a population, or a community to changes in its environment, man-made or natural. The usefulness of any biological-effect method will obviously depend on how well it is able to separate anthropogenic stressors from the influence of variability in natural processes. The output should be a measure of the well-being or health of some ecosystem component. Vital aspects determining health relate to questions such as: are organisms, populations, or communities viable; can they reproduce; do individuals or populations fulfil their potential for growth under relevant natural ecosystem conditions; do they behave normally, and are they physiologically competent? The penultimate two processes are more subtle and may be measures of the homeostatic response of an organism, but can in some instances be a measure of detrimental genetic, biochemical, or physiological changes that may be predictive of higher-level effects. The most widely used biological-effect tools are measures of the biochemical and/or physiological state of selected organisms, such as mussels or fish.

These techniques have most commonly been used to identify the effects of contaminants, whether at point discharges, of diffuse inputs, after accidental spills, or during field monitoring to assess the health status of marine habitats. It is necessary to establish cause-and-effect relationships between the presence of contaminants and biological-effect responses for the methods to be useful tools for informing management and directing environmental policy. A good example here would be tributyltin (TBT)-induced imposex in dogwhelk. Once population effects that could be directly related to concentrations of organotins in the marine environment had been observed (Gibbs and Bryan, 1996), management actions were taken to reduce TBT emissions and introduce international policies through the EU and International Maritime Organisation (IMO), which has resulted in a decrease in the prevalence and severity of imposex (Birchenough *et al.*, 2002). The link between a specific contaminant and a biological-effects response is not always so clear-cut, however, because many biological effects are not specific to individual contaminants and will therefore provide more a measure of general malaise. More than 100 000 known man-made chemicals are released into the environment in one way or the other. Because chemical analytical techniques will only be available for a fraction of all those chemicals, and because of the high cost of chemical analyses and the presence of contaminants as mixtures, it is essential to use biological-effect techniques that can provide measures of the health of ecosystem components. Such techniques need to encompass the synergistic and antagonistic effects of contaminant mixtures when used in a field setting.

Current status and developments

WGBEC (ICES, 2007a) has reviewed the status of biological-effect techniques regularly and recommended in its reports those techniques for fish and invertebrates that are in the research phase, look promising, and require development and analytical-quality control, or are available for use and take-up in national and international monitoring programmes. Some of the recommended methods have been included in OSPAR guidelines for contaminant-specific or general monitoring (JAMP, 1998a, b) and have, after a process of quality assurance, been included in CEMP. The updated list (Table 1) includes information on the current position of each technique relative to these guidelines.

In the past, OSPAR has focused on the tasks of detecting and monitoring specific contaminants, or classes of contaminants, that were known to cause problems [e.g. organotins, polycyclic aromatic hydrocarbons (PAHs), and selected metals]. Although it is possible to design a programme that specifically addresses the effects of individual contaminants, as has been the aim of JAMP (1998b), observed effects will rarely be caused by one agent only. Methods that relate directly to specific contaminants (e.g. fish bile metabolites of PAHs to PAH exposure, or imposex to TBT exposure) are exceptions rather than the rule. Examples to the contrary are response markers to PAH-exposure, most of which will also be affected by other planar contaminants, such as non- and mono-orthopolychlorinated biphenyls (PCBs) and dioxins (dibenzofurans and dibeno-*p*-dioxins). It is therefore important to integrate chemical and biological methods in environmental-contaminant monitoring, and biological-effect methods should not be used in isolation but as part of a suite.

Biological-effect techniques range from responses measured at the subcellular level (e.g. metallothionein and DNA adducts) to whole-organism responses (e.g. scope for growth and disease occurrence). It remains difficult, however, to attach a scoring value to most of those measurements as reflecting the well-being of an individual or population, let alone to make statements about ecosystem health. Clearly, responses at tissue or organ level are more important to the health of an individual than are subcellular responses. Nevertheless, if lower-order responses can be linked to higher-order effects, predictive capability will increase. Relating lower-level effects to ecosystem health remains a challenge. As the level of biological complexity increases, so too does the ecological relevance of any contaminant effect, but this in turn is mirrored by decreasing responsiveness, detectability, and mechanistic understanding.

A further confounding issue is natural variability, such as caused by genetic differences, adaptability to different habitats, or fluctuations in food availability, temperature, salinity, and water quality parameters. Even more important in this respect is that the measures of higher-level effects tend to have a high inherent natural variability (Table 2). To address this issue, considerable effort has gone into standardizing measurement and assessment techniques and into designing sampling programmes that follow prescribed guidelines. Standardization requires several steps. First, the technique itself must be scientifically sound, well tried, and therefore well cited in the scientific literature. Second, the procedure must allow standardization and knowledge of confidence and detection limits using reference

Table 1. OSPAR status of biological-effect techniques for invertebrates and fish (JAMP).

Method	In JAMP	CEMP category	Rec. by WGBEC	QC
Mussels				
Whole sediment bioassays	Yes	II	Yes	B
Sediment pore water bioassays	Yes	II	Yes	B
Sediment seawater elutriates	Yes	II	–	–
Water bioassays	Yes	II	Yes	B
In vivo bioassays	No	–	Yes	B (some)
In vitro bioassays	No	–	Yes	–
Lysosomal stability	No	–	Yes	B-a
Multidrug resistance (MXR/MDR)	No	–	Yes	–
Scope for growth	No	–	Yes	B-a
AChE	No	–	Yes	–
Metallothionein (MT)	No	–	Yes	B-a
Histopathology	No	–	Yes	–
Imposex/intersex in gastropods	Yes	I (M)	Yes	Q
Benthic community analysis	Yes	–	Yes	B
Fish				
AChE in muscle	No	–	Yes	–
Lysosomal stability	Yes	II	Yes	B-a
Externally visible diseases	Yes	I (V)	Yes	B
Reproductive success (eelpout)	Yes	II	Yes	B-a
Metallothionein	Yes	II	Yes	B-a
ALA-D	Yes	II	Yes	B-a
Oxidative stress	Yes	II	–	–
CYP1A-EROD	Yes	II	Yes	Yes
DNA-adducts	Yes	II	Yes	B-a
PAH metabolites	Yes	II	Yes	Q
Liver neoplasia/ hyperplasia	Yes	I (V)	Yes	B
Liver nodules	Yes	I (V)	Yes	B
Liver pathology	Yes	I (V)	Yes	B
Vitellogenin in cod	No	–	Yes	Yes
Vitellogenin in flounder	No	–	Yes	–
Intersex in male flounder	No	–	Yes	B-a

CEMP category: II, method suitable for marine monitoring purposes; I, method suitable and analytical-quality control (AQC) is available; M, mandatory method in place, with AQC and assessment criteria established. Quality control: V, voluntary method in place, with AQC but conducted voluntarily. Recommendations for inclusion by WGBEC (ICES, 2007a) and information on existence of quality control [QC: B, Biological Effects Quality Assurance in Monitoring Programmes (www.bequalm.org); B-a, available online at http://www.bequalm.org; Q, Quality Assurance of Information for Marine Environmental Monitoring in Europe, available online at http://www.quasimeme.org].

materials. Third, the reproducibility of analyses performed by independent researchers and laboratories must be known (as part of the quality assurance programme). Fourth, and arguably most important, can the measured response, taking into account natural variability, be uniquely related to the presence of contaminants?

Table 2. Qualitative scoring (L, low; M, medium; H, high) of ecological relevance (Rel), level of natural variability (Var, including the influence of confounding factors), and specificity (Spec), and ease of standardization of techniques (Stand: D, difficult; E, easy), in relation to level of organization.

Level of organization	Rel	Var	Spec	Stand
Population/community	H	H	L	D
Individual (bioassays)	L/M	L	L	E
Subcellular health (biomarkers)	L/M	L/M	M/H	E

For some biological-effect measurements, variability can be reduced by correcting for confounding factors and by standardizing procedures. The inclusion of methods into national JAMP activities has provided insight into these improvements. Confounding factors addressed and their solution include:

(i) *Reproductive status*. Because the stage of gametogenesis may affect the response considerably (e.g. for EROD, 7-ethoxyresorufin-*O*-deethylase activity, and lysosomal stability), organisms should be sampled when gametogenesis is latent or has limited effect on the response (Eggens *et al*., 1995).

(ii) *Sex*. Because differences in the response may be orders of magnitude different between males and females, particularly in relation to their reproductive state, this factor needs to be included for many techniques (Förlin and Haux, 1990).

(iii) *Size or age*. Because internal organs, such as the liver, vary in size relative to reproductive state and food availability (Vethaak and Jol, 1996), concentrations of some contaminants may build-up over time; and because size can affect biological effects (Ruus *et al*., 2003), size and age of sampled organisms have to be standardized to ensure like-for-like comparison of data.

(iv) *Origin*. Because different races or subpopulations may be present at the same location at different times of the year, care should be taken that measurements refer to a group with the same genetic origin (wherever possible).

These factors are important for ensuring data quality from any monitoring survey, but become even more important for interpreting year-to-year information from one location as well as spatial information across wide geographical areas. Ideally, when multiple biological-effect methods are used, they should be applied to the same organism whenever possible (e.g. EROD, DNA-adducts, and histopathological biomarkers, all from the same liver; Feist *et al*., 2004). Such standardization would give confidence in understanding apparent cause-and-effect relationships and would also provide a good example of how biological-effect responses could be integrated.

A strategy for integration

Developing an indicator of ecosystem health using measures of biological effects is a great challenge. As the level of biological complexity increases, so too does the level of ecological relevance of any contaminant effect. Conversely, this increase is mirrored by a decrease in immediate responsiveness, detectability, and mechanistic understanding. The challenges are therefore to

Table 3. Preliminary assessment criteria for three levels of response (background, elevated, high = cause for concern) in relation to biological effects [LOD, limit of detection; from WKIMON (ICES, 2007c)].

Biological effect	Qualifying comments	Background	Elevated	High
VTG in plasma (μg ml^{-1})	Cod	LOD to 2	–	–
	Flounder	LOD to 2	–	–
Reproduction in eelpout (mean frequency in %)	Malformed larvae	0–1	>1–2	>2
	Late dead larvae	0–2	>2–3	>3
	Growth/retarded larvae	0–4	>2–6	>6
EROD (pmol min^{-1} mg^{-1} protein)	Cod	≤80	–	–
	Dab	≤40	–	–
	Flounder	≤10	–	–
DNA-adducts (nmol adducts per mol DNA)	Dab	≤7.9	–	–
	Haddock	≤6.8	–	–
	Saithe	≤7.9	–	–
Bioassays (% mortality)	Sediment *Corophium*	0–30	>30 to <100	100
	Sediment *Arenicola*	0–10	>10 to <100	100
	Water bivalve embryo	0–20	>20 to <100	100
	Water copepod	0–10	>10 to <100	100
	Water echinoderm	0–10	>10 to <100	100
Lysosomal stability (min)	Cytochemical	>20	≤20 to ≥10	<10
	Neutral red retention	>120	≤120 to ≥50	<50

develop (i) an integrated approach to using the biological-effects measures, and (ii) tools for assessing the data in such a way that the contribution of each measure to the health of the system can be determined, both by itself and as part of the overall effect. This conceptual task is made more difficult because the data available for contaminants and ecological effects only provide information for some components of the ecosystem, and because uncertainties exist about how this information should be conceptually linked. Contaminant concentrations and biological-effect data mostly provide information at a relatively low level of molecular or organismic integration, i.e. at the sub-lethal or individual level.

Over the past 2 years, some progress on how to meet these challenges has been made through the work of WGBEC and WKIMON (i) by starting to develop assessment criteria, and (ii) by developing a strategy for integrative sampling and assessment.

Developing assessment criteria

The development of assessment criteria was addressed by WGBEC (ICES, 2006, 2007a) and further considered by WKIMON (ICES, 2007c). WGBEC (ICES, 2006) distinguished between biological-effect methods for which it would be appropriate to establish a global background level, and hence the ability to derive a general assessment criterion (deviation from normal) and methods for which there would be a need for a reference location with which to compare populations at affected locations. The latter category included methods that were thought difficult to control for external, non-contaminant factors (e.g. sex, temperature, etc.). However, the present view is that it may be possible to derive global values for all methods if a limited number of dominant external factors can be controlled for (ICES, 2007c). Hepatic cytochrome P4501A activity (measured by EROD) is one example for which temperature and/or gonad development corrections may in fact be sufficient. Response levels above background for EROD, bile metabolites, DNA-adducts, and VTG have yet to be determined. Preliminary values are currently available for some methods (Table 3), but this list still has to be expanded and improved.

A second criterion should distinguish between moderate and strong effects. Proposals have been made for effects on reproduction in eelpout (Table 3). The ICES Working Group on Diseases and Pathology of Marine Organisms (WGDPMO) has developed a fish disease index (FDI), which includes an assessment of externally visible lesions and parasites, macroscopic liver neoplasms, and histopathological liver lesions (ICES, 2007b). In general, however, more work remains to be done to validate the assessment criteria developed so far and to establish elevated response levels, as well as those that are a cause for concern.

Developing an integrated approach

WGBEC (ICES, 2006) made a preliminary proposal, endorsed by WKIMON (ICES, 2007c), for combining methods in a programme of integrated chemical- and biological-effect monitoring that contains three ecosystem components: water, sediment, and biota (restricted so far to fish, bivalves, and gastropods; Figure 1). Although this is an obvious simplification in terms of ecosystem assessment, the design rests on existing knowledge, and the methodology available could easily be extended if more information becomes available. Compared with the current CEMP methodology, the assessment of contaminants in water has been removed from a cost–benefit perspective, because of the high variability in water sample analysis and the associated large number of analyses required. Another discussion has arisen about benthic-community analyses, standard components in many national monitoring programmes. In the past, community analyses (termed "benthic ecology" in Figure 1) have been seen as not providing sufficient information about contaminant effects to merit their inclusion. On the other hand, such analyses would provide information about other environmental processes,

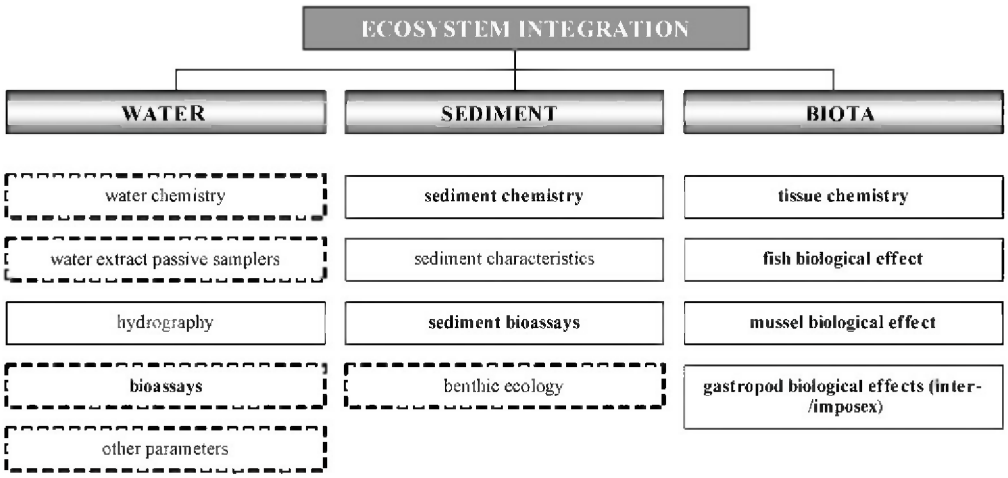

Figure 1. Overview of main components proposed for an integrated monitoring programme for effects of contaminants. Solid-lined boxes, prioritized core components; bold text, components in CEMP; broken-lined boxes, additional optional components proposed.

such as eutrophication and hypersedimentation, that may interact with contaminants in their environmental impacts.

Assessment criteria for effects on gastropods (intersex/imposex) have been described extensively (JAMP, 1998b; ICES, 2007c), leaving biological effects on fish and mussels as the most complex issues to deal with in the framework. The proposal for fish has taken account of methods already in CEMP but has suggested deletions and additions in line with current scientific understanding (Figure 2a). Cost–benefit considerations have also been taken into account. PAH should not be measured in fish tissues because most PAHs are efficiently metabolized, so measuring metabolites in bile is preferred. Although a CEMP method, reproductive success in eelpout is not suggested as a standard component because of its limited distribution. The presence of liver nodules (diameter >2 mm) is viewed as an imprecise category, the relevant components of which are included in liver neoplasms. Metallothionein (a metal-binding protein) has been a component in both UK and Norwegian monitoring programmes, but measurements do not appear to be easily interpreted for toxic metal stress for the two species (Atlantic cod, dab) for which the method has been applied. At relevant exposure levels, metallothionein appears to reflect the status of Cu and Zn in the liver, at least in cod. Although found to be useful in the Norwegian JAMP, δ-aminolevulinic acid dehydratase (ALA-D) has not been used by other countries and is therefore not recommended for inclusion in a core programme. DNA-adduct formation is currently in CEMP and has been retained as a proposed method, although analyses are expensive. Important reasons for inclusion include the possibility to develop highly specific assessment criteria and its relevance in interpreting effects at the tissue level (neoplasia). Two methods not in CEMP have been suggested for inclusion in the core programme: plasma vitellogenin and acetyl cholinesterase inhibition (AChE) in muscle. Plasma vitellogenin can be used to assess impacts of environmental estrogens and has been widely tested in a range of marine species.

Mussels are important monitoring species, not least because of their ecological importance and wide geographical distribution (basically the entire range of North Atlantic, Baltic, and Mediterranean coastal areas). A range of techniques for blue mussels have been recommended by WGBEC, but there are also other promising methods on the list that are viewed as appropriate for integrated chemical- and biological-effect monitoring, the core methods suggested being scope for growth, histopathology, lysosomal stability, AChE inhibition, and micronucleus formation in hematocytes (Figure 2b).

There are processes under way to establish assessment criteria for the biological-effect methods. A range of strategies has been proposed for the integration of results, including multivariate approaches and graphical representation of a range of components (Choi *et al.*, 2005). Obvious criteria will be transparency, reproducibility, and appropriate weighting of individual components towards a final indicator.

Developing a framework for indicators of health

Some attempts to develop a framework for indicators of health have been trialled. In the UK, a strategy has been proposed known as the Fullmonti (Fully Integrated Monitoring Strategy; Defra, 2007). Three main components (chemistry, benthic community, and fish) were assessed separately for each of a range of locations along the UK coasts to produce “traffic-light”-type indicators (good, moderately impacted, and strongly impacted). Each indicator was in turn computed from a range of weighted subcomponents, the weighting being based on expert judgement. For biological effects, the weighting was based on the perceived ecological relevance of the response measured so that, for example, a reproductive or disease response would have a greater weighting than a general biomarker response, such as enzyme induction. For community assemblages, indices such as the Shannon–Wiener, AZTI marine biotic index, and infaunal trophic index had a greater weighting than species number. For chemistry, a greater weighting was given to persistent and bioaccumulating compounds. Attempts to combine the three indicators into a single indicator were unfruitful because so much useful information is watered down or is even lost. An important benefit of the Fullmonti approach is that partial datasets can be used (e.g. sites at which only a restricted set of biological effects or contaminants have been measured). Although more comprehensive expert systems could be devised, the Fullmonti approach shows promise.

Different strategies to develop a holistic assessment have been chosen in other European countries. In Norway, biological-effect techniques have been implemented in the national JAMP, so

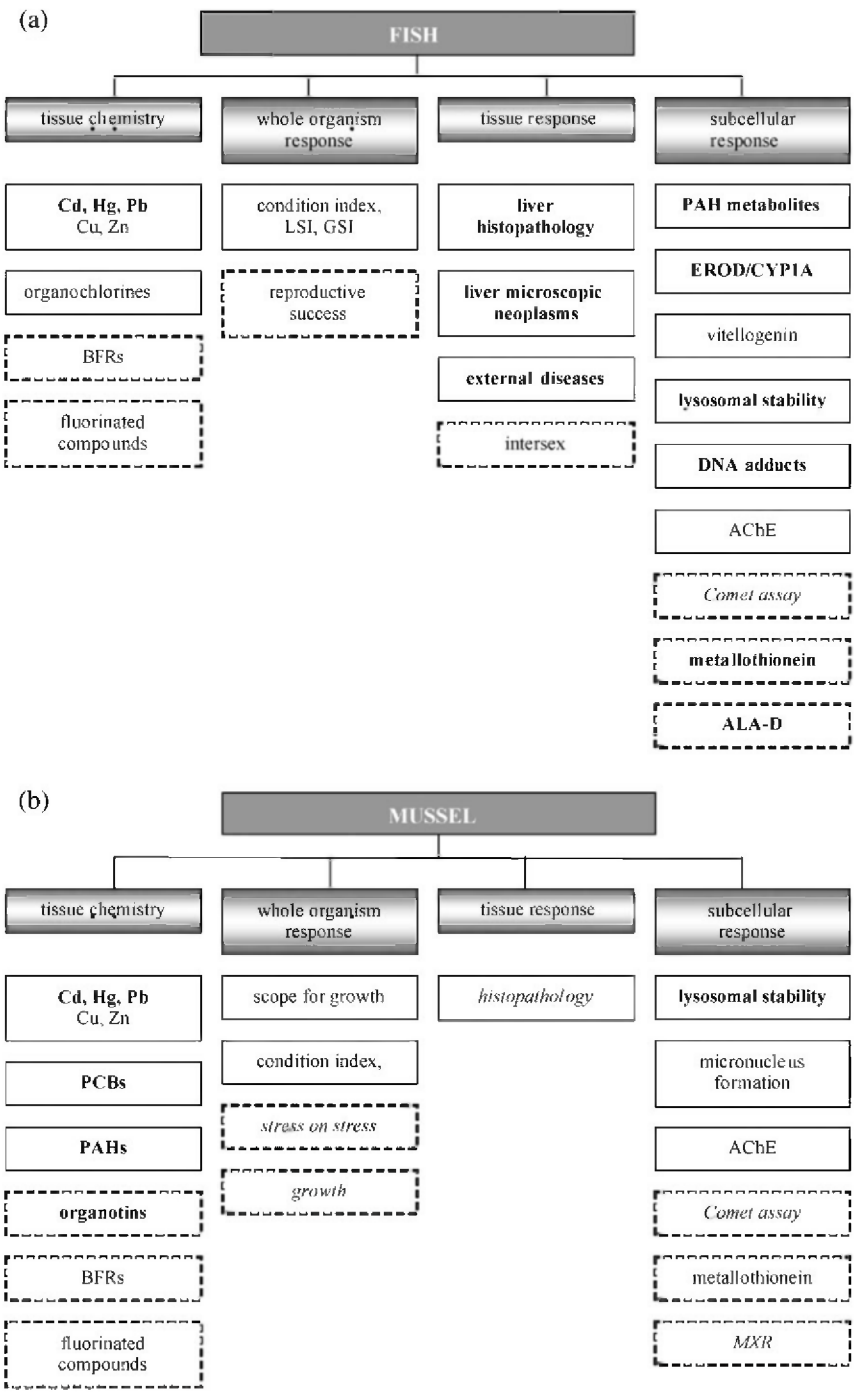

Figure 2. Overview of biological-effect methods to be included in an integrated monitoring programme [modified from WGBEC (ICES, 2006) and WIKIMON (ICES, 2007c)] for (a) selected fish species and (b) blue mussel. Legends as in Figure 1, except broken-line boxes with regular text refers to recommended methods, and italicized text to promising methods (WGBEC; ICES 2007a).

that both contaminant concentrations and effect parameters (biomarkers) were measured in the same individuals. The aim was to improve the interpretation of both effects and to accumulate additional information on physiological parameters (liver somatic index, size, age, and gender), as well as temperature (Ruus *et al.*, 2003). Although some links could be established between contaminant concentration in tissue and effects, the results highlighted some other issues: (i) species differences were the rule rather than the exception (sampling included Atlantic cod, flounder, and dab); (ii) inclusion of physiological parameters was essential; (iii) generally, no direct relationship should be expected between effects and accumulating contaminants (e.g. PCBs); and (iv) there were strong site-dependent factors that could not be identified (Ruus *et al.*, 2003). Nevertheless, the data did provide an overall picture of how and whether or not contaminants affected fish at the sites. A similar approach has been in operation in Sweden since the 1980s (Hansson *et al.*, 2006). Briefly, a range of physiological and biochemical parameters has been measured alongside contaminants and population-level assessments for perch in the Baltic and (less extensively) eelpout in the Skagerrak. Because of the long time-series and knowledge of baseline levels for contaminants, it has been possible to pinpoint specific pollution events, as well as assessing long-term changes in these fish populations. In the Dutch national JAMP, fish-disease monitoring with dab and flounder has been integrated with residue-contaminant measurements (including exposure biomarkers such as biliary PAH metabolites) and contaminants in sediment, and supporting biological and hydrographical data (Bovenlander and Langenberg, 2006). The integrated approach allowed evaluation of one facet of coastal and estuarine ecosystem health, but at the same time demonstrates that migration patterns play a critical role in explaining the distribution of chronic diseases such as liver neoplasms in flatfish (Vethaak *et al.*, 2008).

The resources involved for contracting parties in implementing a core monitoring programme as proposed will be substantial. Although it is envisaged that the criteria and assessment protocols required will be established soon and be part of an OSPAR–JAMP guideline on integrated monitoring, the guideline should preferably be tried out in, and if necessary adjusted based on, a demonstration programme before its implementation by contracting parties.

Further work

Although important steps have been taken to develop a process for integrated chemical- and biological-effect monitoring, the following is needed to finalize the work:

(i) final agreement on components to be included;

(ii) finalize assessment criteria for the methods to be included;

(iii) final agreement on strategy and method to combine results at the ecosystem level;

(iv) validation of assessment criteria;

(v) validation of the framework through a set of case studies and a field validation activity (currently under way).

References

Birchenough, A. C., Evans, S. M., Moss, C., and Welch, R. 2002. Re-colonisation and recovery of populations of dogwhelks *Nucella lapillus* (L.) on shores formerly subject to severe TBT contamination. Marine Pollution Bulletin, 44: 652–659.

Bovenlander, R. W., and Langenberg, V. T. 2006. National evaluation report of the joint assessment and monitoring programme of the Netherlands 2004. RIKZ Report, 2006.002. RWS/National Institute for Coastal and Marine Management, The Hague. 45 pp.

Choi, J. S., Frank, K. T., Petrie, B. D., and Legget, W. C. 2005. Integrated ecosystem assessment of a large marine ecosystem: a case study of the devolution of the Eastern Scotian Shelf, Canada. Oceanography and Marine Biology: an Annual Review, 43: 47–67.

Defra. 2007. Department of Environment, Food and Rural Affairs, Development of a Fully Integrated Monitoring Strategy. Department of Environment, Food and Rural Affairs, contract reports. http://www.defra.gov.uk.

EEA. 2001. Environmental Signals Indicator Report. European Environment Agency. Environmental Assessment Report, 8. 112 pp.

Eggens, M., Bergman, A., and Vethaak, A. D. 1995. Seasonal variation of hepatic EROD activity in flounder (*Platichthys flesus*) in the Dutch Wadden Sea. Marine Environmental Research, 39: 122–124.

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

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

Gibbs, P. E., and Bryan, G. W. 1996. Reproductive failure in the dogwhelk *Nucella lapillus* associated with imposex caused by tributyltin pollution: a review. *In* Organotin: Environmental Fate Effects, pp. 259–284. Ed. by M. A. Champ, and P. F. Seligman. Chapman and Hall, London.

Hansson, T., Lindesjo, E., Förlin, L., Balk, L., Bignert, A., and Larsson, A. 2006. Long-term monitoring of the health status of female perch (*Perca fluviatilis*) in the Baltic Sea shows decreased gonad weight and increased hepatic EROD activity. Aquatic Toxicology, 79: 341–355.

ICES. 2006. Report of the Working Group on Biological Effects of Contaminants (WGBEC). ICES Document CM 2006/MHC: 04. http://www.ices.dk/iceswork/wgdetail.asp?wg=WGBEC.

ICES. 2007a. Report of the Working Group on Biological Effects of Contaminants (WBGEC). ICES Document CM 2007/MHC: 03. http://www.ices.dk/iceswork/wgdetail.asp?wg=WGBEC.

ICES. 2007b. Report of the Working Group on Pathology and Diseases of Marine Organisms (WGPDMO). ICES Document CM 2007/MCC: 04. http://www.ices.dk/iceswork/wgdetail.asp?wg=WGPDMO.

ICES. 2007c. Report from the ICES/OSPAR Workshop on Integrated Monitoring of Contaminants and their Effects in Coastal and Open-Sea Areas (WKIMON). ICES Document CM 2007/ACME: 01. http://www.ices.dk/iceswork/wgdetailacfm.asp?wg=WKIMON.

IOC. 2002. A Reference Guide on the Use of Indicators for Integrated Coastal Management. Intergovernmental Oceanographic Commission of the UNESCO, Manuals and Guides, 45. 127 pp.

JAMP. 1998a. JAMP guidelines for general biological effects monitoring. Joint Assessment and Monitoring Programme. Oslo and Paris Commissions. 38 pp.

JAMP. 1998b. JAMP guidelines for contaminant-specific biological effects monitoring. Joint Assessment and Monitoring Programme. Oslo and Paris Commissions. 38 pp.

Jorgensen, S. E., Xu, F-L., and Costanz, R. (Eds). 2005. Handbook of Ecological Indicators for Assessment of Ecosystem Health. CRC Press, Boca Raton, FL. 439 pp.

Ruus, A., Hylland, K., and Green, N. 2003. Joint Assessment and Monitoring Programme (JAMP). Biological Effects Methods, Norwegian Monitoring 1997–2001. NIVA Report, 869/03, TA-1948/2003. 139 pp.

Vethaak, A. D., and Jol, J. 1996. Diseases of flounder (*Platichthys flesus*) in Dutch coastal waters, with particular reference to environmental stress factors. Part 1. Epizootiology of gross lesions. Diseases of Aquatic Organisms, 16: 81–97.

Vethaak, A. D., Pieters, J., and Jol, J. 2008. Long-term decline in fish disease prevalence reflects water quality improvements in the North Sea. Deltares Report, Z4612.10/08, Delft. 41 pp.

doi:10.1093/icesjms/fsn120