Integrated chemical and biological assessment of contaminant impacts in selected European coastal and offshore marine areas

Ketil Hylland[1], Craig D. Robinson[2], Thierry Burgeot[3], Concepción Martínez-Gómez[4], Thomas Lang[5], Jörundur Svavarsson[6], John E. Thain[7], A. Dick Vethaak[8], Mattew J. Gubbins[2]

[1] *Department of Biosciences, University of Oslo, PO Box 1066, Blindern, N-0316 Oslo, Norway*

[2] *Marine Scotland Science, Marine Laboratory, 375 Victoria Road, Aberdeen, AB11 9DB, UK*

[3] *IFREMER, Laboratory of Ecotoxicology, Rue de l'Ile d'Yeu, B.P. 21105, 44311 Nantes Cédex 03, France*

[4] *Instituto Español de Oceanografía (IEO), Oceanographic Centre of Murcia, Varadero 1, PO BOX 22, 30740 San Pedro del Pinatar (Murcia), Spain.*

[5] *Thünen Institute of Fisheries Ecology, Deichstr. 12, 27472 Cuxhaven, Germany*

[6] *University of Iceland, Askja – Natural Science Building, Sturlugata 7, 101 Reykjavík, Iceland*

[7] *Cefas Weymouth Laboratory, Barrack Road, The Nothe, Weymouth, Dorset, DT4 8UB, UK*

[8] *Deltares, Marine and Coastal Systems, P.O. Box 177, 2600 MH Delft, The Netherlands / VU University Amsterdam, Amsterdam Global Change Institute, Institute for Environmental Studies, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands*

Communicating author: Ketil Hylland, ketilhy@ibv.uio.no, phone +4722857315/+4741451694

Abstract

This paper reports a full assessment of results from ICON, an international workshop on marine integrated contaminant monitoring, encompassing different matrices (sediment, fish, mussels, gastropods), areas (Iceland, North Sea, Baltic, Wadden Sea, Seine estuary and the western Mediterranean) and endpoints (chemical analyses, biological effects). ICON has demonstrated the use of a framework for integrated contaminant assessment on European coastal and offshore areas. The assessment showed that chemical contamination did not always correspond with biological effects, indicating that both are required. The framework can be used to develop assessments for EU directives. If a 95% target were to be used as a regional indicator of MSFD GES, Iceland and offshore North Sea would achieve the target using the ICON dataset, but inshore North Sea, Baltic and Spanish Mediterranean regions would fail.

Keywords: ICON, contaminants, European seas, biological effects, assessment

Introduction

Thousands of tonnes of waste are released into European seas every minute, containing chemicals that have the potential to accumulate in marine organisms and/or affect their health. As discussed in Borja et al. (2010), it is crucial in this context to have a clear understanding of how it can be determined whether organisms or populations in an area are affected by pollution and if so, the extent to which they are impacted. With regards to chemicals, this implies quantifying chemical-specific effects on marine organisms or processes. In addition to a required knowledge of effects, there are reasons why it may also useful to have information about concentrations of chemicals in organisms or abiotic matrices: (i) to link observed effects to specific chemicals for regulatory purposes, (ii) to ensure concentrations are not above limits set for human consumption, and finally (iii) to document the presence of chemicals that may or may not cause effects. As support for effects, it is the exposure of organisms to chemicals that matters. For persistent bioaccumulating substances, exposure can be estimated through measuring the concentration of chemicals or their metabolites in the tissues of the target organism (e.g. Hylland et al., 2009) or in other matrices such as passive samplers (Utvik & Gärtner, 2006), sediments or non-target organisms in the same habitat, e.g. blue mussels. Some polluting chemicals may however be quickly degraded or present at concentrations below the detection limit of routine chemical analyses, but still cause impacts, e.g. many endocrine disrupting substances, organophosphate pesticides and pharmaceuticals. In this case, biological responses will be the most sensitive method by which to detect their presence, e.g. through the inhibition of acetylcholinesterase as a result of organophosphate exposure (Bocquené et al., 1993) or increased plasma concentrations of vitellogenin in juvenile fish as a result of oestrogen exposure (Allen et al., 1999). To understand the possible environmental consequences and regulate inputs of contaminating chemicals, we therefore need to know both the concentrations of contaminants in appropriate matrices as well as how they affect organisms. The two types of measurements, chemical and biological, should ideally be combined in an integrated assessment (cf. Davies & Vethaak, 2012). Any monitoring programme underpinning such an assessment will however produce a very extensive and complex data matrix, which will require some sort of aggregation procedure prior to being used for regulatory decisions. Such aggregation procedures are generally termed "indicators", see e.g. Rees et al. (2008). Indicators have previously been developed separately to aggregate or combine chemical analyses (see e.g. OSPAR, 2010) or biological responses, e.g. the health assessment index, HAI (Adams et al., 1993), biological assessment index, BAI (Broeg et al., 2005), an expert system (Viarengo et al., 2000; Dagnino et al., 2007), the integrated biological response, IBR (Devin et al., 2014), the biomarker response index (BRI) (Hagger at al., 2008) or the integrative biomarker Index, IBI (Marigómez et al., 2013). In addition, there are some practical examples of integrating or combining chemical analyses and biological responses, such as in the UK Fullmonti project, including chemical analyses, benthic community status and fish health (described in Thain et al., 2008) or by using a weight-of-evidence approach (see e.g. Chapman et al., 2002). In some national programmes, the interpretation of fish health is aided by taking account of contaminant levels in addition to confounding factors such as size and gender, and environmental factors such as temperature and season (see e.g. Sandström et al., 2005; Hylland et al., 2008, 2009; Vethaak et al., 2008). The main difference between the framework used here (described in Vethaak et al., this issue-a) and other indices is that the current framework is based on internationally agreed threshold criteria for biological responses and tissue residues of chemicals, identifying responses above background, responses that indicate probable impacts at the population level and concentration of chemicals above thresholds (see Robinson et al., this issue). In addition, the framework includes more matrices than most other indices and is flexible in the species included, as long as criteria exist for core methods.

Over the last decade, Europe has implemented two directives that largely direct the management of the environmental conditions of coastal and offshore marine areas, the Water Framework Directive (WFD, 2000/60/EC) and Marine Strategy Framework Directive (MSFD, 2008/56/ EC). Particularly descriptor 8 of MSFD, 'Concentrations of contaminants are at levels not giving rise to pollution effects", is clearly relevant for the assessment described here for the ICON project (International workshop on marine integrated contaminant monitoring, see Hylland et al., this issue-a, for a full description). Using biological responses to provide the information required for descriptor 8 has been suggested in e.g. Bourlat et al. (2013), Giltrap et al. (2013), Hagger et al. (2008), Lehtonen et al. (2014) and Lyons et al. (2010). As outlined in Lyons et al. (2010), the framework described in Vethaak et al. (this issue-a) and applied to the ICON project will output a metric that can be used to determine Good Environmental Practice (GES) in MSFD.

The current paper reports on an integrated assessment of the results from the ICON (International workshop on marine integrated contaminant monitoring) project, using results reported in Burgeot et al. (this issue), Carney Almroth et al. (this issue), Hylland et al. (this issue-b), Kammann et al. (this issue), Lang et al. (this issue – a,b), Lyons et al. (this issue), Martinez-Gomez et al. (this issue –a, b), Robertson et al. (this issue), Vethaak et al. (this issue-b).

As described in Vethaak et al. (this issue-a), this indicator of status for each determinant can then be combined at different levels: matrix, site and region, and expressed with varying levels of aggregation to graphically represent the proportion of different types of determinants (or for each determinant, sites within a region) exceeding assessment criteria. Such an approach has several advantages: (i) the combination of data can be done for selected levels depending on the type of assessment required and the monitoring data available, (ii) the representation maintains all the original information and it is straightforward to identify determinants that exceed the assessment criteria, (iii) 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 (cf. Jennings et al., 2008). In contrast to some other integrating indicators, e.g. IBI and BRI, there is no weighing of the methods included in the current framework. The approach is based on the OSPAR regional assessment tool developed for contaminants (OSPAR, 2010).

Methods

The assessment criteria used with chemical components of the framework were OSPAR Background Assessment Criteria (BACs) and Environmental Assessment Criteria (EACs) or EU Environmental Quality Standards (EQSs); EC food safety regulation limits were used where EACs or EQSs are not available (OSPAR, 2008). Food safety regulation limits are not necessarily protective for the environment. Assessment criteria for biological responses (biomarkers) were from Davies & Vethaak (2012). Initial comparisons (step 1 below) would decide whether the concentration or response for any species or matrix at any site was less than BAC, between the BAC and EAC, or above EAC. As described in detail in Hylland et al. (this volume – a) and Vethaak et al. (this volume – a), biological responses were grouped in either "exposure" or "effect", subject to whether there is available data showing adverse effects corresponding to that particular response.

The sites included in the ICON project are described in Hylland et al. (this issue - a). They comprised sites from the Mediterranean in the south to Iceland in the north, encompassing the Seine estuary, Wadden Sea, a range of coastal, estuarine and offshore sites in the North Sea and one site in the Baltic (Table 1). The two coastal and two offshore sites on Iceland were included as reference sites.

The matrices chosen for ICON were sediment, haddock (*Melanogrammus aeglefinus*), dab (*Limanda limanda*), flounder (*Platichthys flesus*), red mullet (*Mullus barbatus*), gastropod (*Nucella lapillus*) and mussels (*Mytilus edulis* or *M. galloprovincialis*) (cf. Hylland et al., this issue-a). The chemical analyses performed in ICON were for PAHs, PCBs, Cd, Hg and Pb (Robinson et al., this issue). The biological responses included for fish were (exposure indicators): red blood cell micronucleus frequency, genotoxicity (comet assay), cytochrome P4501A activity (EROD), bile PAH metabolites (by HPLC), plasma vitellogenin (VTG) and intersex, and (effect indicators): lysosomal membrane stability (LMS), acetylcholinesterase inhibition (AChE), bile PAH metabolites (by synchronous scanning fluorometry, SFF), DNA adduct concentration, external fish disease, hepatic neoplasms and liver histology. The two methods for PAH metabolite analyses can be converted one to the other, but only SSF data has been linked directly to adverse effects in experimental studies, hence the grouping in "exposure" and "effect". Effect responses for mussels were acetylcholinesterase inhibition (AChE), stress-on-stress (SoS), scope for growth (SfG), metallothionein (MT), histopathology (histo), lysosomal membrane stability (LMS), and for gastropods imposex (VDSI). The reader is referred to Davies & Vethaak (2012) and the relevant chapters of that volume for more detail on background data and the motivation for selecting methods. The selection of methods follows on from discussions in the ICES working group on biological effects of contaminants (WGBEC) over the past two decades (see e.g. ICES, 2010). The original list of recommended methods were further refined by the ICES/OSPAR working group SGIMC (ICES, 2011), taking into account additional issues such as cost-benefit and availability of analytical techniques in different countries. The final selection largely corresponds to the methods chosen by HELCOM for the Baltic (CORESET) (Lehtonen et al., 2014). The data from the individual studies in ICON (reported in this special issue) were compiled and subjected to a five-step procedure, eventually resulting in an overall assessment of the sites included in ICON. The assessment strategy is transparent and, depending on the objectives of an assessment, it may be desirable to stop after steps two, three or four.

Step 1: Assessment of monitoring data against BAC and EAC

All measurements performed within ICON were compared with the relevant BAC and EAC for that specific endpoint and species and expressed as a colour depending on whether the value exceeded the BAC or EAC. Details of calculations can be found in Davies & Vethaak (2012) and in Vethaak et al. (this volume –a). A red classification would indicate that the value was above EAC, blue indicated values below the BAC, while green indicated concentrations or effect responses between the BAC and EAC. The method for determining whether a response is in either category can be found in Vethaak et al. (this issue-a). For all biological responses it is possible to identify a level at which the investigated population would be classified as being exposed to contaminants, i.e. with values above the background assessment concentration (BAC), but for only some of the methods will there be data available that can link the response to e.g. increased mortality in some life stage of the same species at that concentration, providing the environmental assessment concentration (EAC).

Step 2: Integration of determinants by matrix for a given site

For each of the matrices the results of the individual assessments were aggregated into three main categories: contaminants, exposure indicators and effects indicators. For sediment/water, passive sampling and bioassays were done for some sites (see Vethaak et al., this issue-a). Exposure indicators are biological responses that are not predictive of "significant" effects, i.e. exceeding EAC, and can hence only be blue or green. It was found necessary to split the biological effects measurements into two categories depending on whether an EAC was set for that specific response or not. Otherwise aggregated information on the proportion of determinants exceeding the separate AC would 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.

In future projects with aggregation/integration of the above indicators across matrices for a specific site, bioassays will be considered ‘effects indicators’ as EACs become available. It will be possible to include data from passive sampling and *in vitro* bioassays in both the water and sediment components in the framework whenever assessment criteria become available.

The integration by matrix and category of determinant are expressed by three- or four-coloured bars showing the proportions of determinants that exceed the BAC and EAC. To indicate a lack of results for core methods or lack of data, grey has been used. Each method for contaminant, effect or exposure assessment carries the same weight, within matrix, in the integration. All determinants carry the same weight in the assessment as they are perceived to have equivalent significance. That is to say all determinants either represent a contaminant concentration or effect that is either above or below background (BAC), or likely to cause (contaminant EAC) or be indicative of (effect EAC) significant detrimental effects to individuals or populations of marine organisms.

Step 3: Integration of matrices for a site assessment

In order to express the results of assessment for any particular site, assessments were aggregated across matrices and expressed by determinant category. To achieve this, results from passive sampling from sediment and water categories were integrated into the contaminant indicator graphic and bioassays and gastropod intersex/intersex integrated into 'effects indicators'. Thus the outcome of assessment of all determinants from all matrices can be expressed for a whole site. Practically, the process adopted is to sum the percentages of each colour in, say, the "contaminants" columns for each matrix, and then to scale the sums to a total of 100%.

For some assessments, this will be the highest level of aggregation required. However, for assessments covering larger geographical areas where assessments need to be undertaken across multiple sites, a further level of integration is required (steps 4 and 5).

For transparency, each determinant group 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 sub-set of the monitoring matrices.

Step 4: Regional assessment across multiple sites

A regional assessment can be done at different levels, i.e. aggregation of data at the sub-regional, regional and national levels, in different ways to express both the overall assessment of proportion of determinants (across all matrices) exceeding both assessment thresholds (BAC/EAC) and by determinant for the region, showing the proportion of sites assessed in the region that exceed the thresholds. Both approaches show the overall proportion of determinant/site that exceeds the threshold for each method.

Step 5: Overall assessment

The assessment by region can be aggregated further into a single schematic showing the proportion all determinants across all sites that exceed BAC and EAC. This can be used for the purposes of an overall assessment. The overall assessment can be easily “unpacked” through the steps above to determine which sites and determinants (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 the desired status for a region.

The assessment criteria for fish were grouped in three categories: concentrations of selected contaminants, biomarkers of exposure (e.g. PAH metabolites and cytochrome P4501A (EROD) activity) and biomarkers of effect (e.g. DNA damage, fish disease). For each category the response at each location was then scored.

Results

Assessments were performed by matrix (sediment, mussels, gastropods and fish), by site and by region.

Assessment results by matrix

Contaminant concentrations measured did not exceed EAC values at any of the offshore sites for sediments, yet at two of these sites (Iceland SE and Firth of Forth offshore) sediment bioassay results exceeded EAC values, suggesting effects may be being caused by contaminants not measured in sediment samples (Figure 1). Iceland SE is adjacent to areas with high volcanic activity, which could result in elevated concentrations of e.g. metals not analysed for. At inshore sites, concentrations of the trace metals mercury and lead exceeded EAC values at the Wadden Sea site, the Baltic Sea site and the Cartegena site in Spain, while mercury also exceeded EAC values in the Seine estuary and the Firth of Firth, where PAH concentrations also exceeded EAC. In the Wadden Sea, sediment bioassay results exceeded EACs, indicating significant effects, presumably resulting from the high trace metal concentrations recorded.

The mussel data assessment for Bjarnarhöfn (Iceland) and Palos Cape (SE Spain) showed good relationship between chemical analytical results and biological responses, with contaminant concentrations generally below BAC and little biological effects (Figure 2). The results also showed a response of the mussels that corresponded with the less contaminated station in Le Moulard (France) and the more contaminated site in Le Havre (France), both in the Seine estuary. At one site (Cartagena, SE Spain) there were elevated lead concentrations in the mussels, which did not appear to result in biological effects. In contrast, a high stress response (LMS) was observed at two sites (Firth of Forth in Scotland, Wadden Sea in the Netherlands) where concentrations of the measured contaminants were below EAC thresholds, suggesting alternative environmental stressors (not measured here) as the cause of the response. More focused monitoring would be required to determine the cause of the effects observed at those two sites.

The imposex response of gastropods to environmental concentrations of organotins has been integrated in the scheme by incorporating results from adjacent shoreline populations (Figure 3). Only a single site (Le Havre in the Seine estuary) had a level of imposex of concern, above EAC.

The fish species included in the assessment were dab (LL), flounder (PF), haddock (MA) and red mullet (MB). Two of the species were found at some sites, e.g. dab and haddock in the Firth of Forth and the two Iceland sites and dab and flounder in the Seine estuary and the Baltic site (Figure 4). Concentrations of PCBs in dab, flounder and haddock exceeded EACs at some sites and fish at all sites except red mullet at Cartagena had elevated concentrations of Cd. Furthermore, there was evidence of exposure of dab, flounder and haddock to PAHs at many sites, including Hvassahraun, Firth of Forth, German Bight, Wadden Sea, Seine sites and the Baltic site. There was good correspondence between results for the two methods used to quantify PAH metabolites, but no clear relationship between the elevated PAH metabolite concentrations at many locations and responses such as EROD and measures of genotoxicity (comet, DNA adducts). There were however values above EAC for both LMS and AChE at three sites, including Ekofisk, Dogger Bank and the Baltic site (all dab), and for one of them at Iceland (dab), Firth of Forth (dab), the Seine estuary (flounder) and the Baltic (flounder). Histology also suggested a range of sites were somewhat affected, i.e. dab at both Iceland sites, dab at Ekofisk, flounder at all Firth of Forth sites, dab at Firth of Forth, Dogger Bank and the German Bight.

Assessment by site

To allow region-wide assessments, data are combined by matrix and site. Such an assessment could include selected regions, e.g. Iceland, North Sea coastal and offshore, the Baltic and the Mediterranean. Figures are only shown for North Sea offshore to demonstrate what such an assessment may look like. Sites at Iceland included both coastal (Bjarnarhöfn, Hvassahraun) and offshore (Iceland SE, Iceland SW) locations. All determinants for the coastal sites were below EAC, whereas contaminants (PCB in haddock liver) and effects (AChE and DNA adducts in fish and bioassays of whole sediments) were above EAC for one or more of the two offshore sites sampled. Most of the exposure responses were at or below background levels. Both contaminants and effects were above EAC at some coastal sites in the North Sea. Although coastal North Sea sites comprised the greatest data contribution to the overall assessment, there were biological responses lacking, particularly for exposure. Contaminant concentrations were largely below EAC levels in North Sea offshore sites, except for PCBs in fish liver at Firth of Forth and German Bight (Figure 5). At most sites there was evidence of exposure of fish to genotoxic compounds. At the sites Ekofisk, Firth of Forth and Dogger Bank there were significant levels (>EAC) of toxicant-induced physiological stress. At the single site surveyed in the Baltic there was evidence of contamination above background levels for PAH and heavy metals (Cd) with some heavy metals (Pb, Hg) exceeding EAC thresholds in sediment and PCBs exceeding EAC in dab livers. Dab was found to be exposed to PAH, and both flounder and dab showed significant effects through LMS (and AChE for flounder) effects indicators.

Regional assessments

Results of the assessments conducted above can be further aggregated into regional assessments by representing the proportion of determinant/matrix/site in each assessment category (blue, green, red). This can be visualised for contaminants, exposure and effects indicators as in Figure 6 or by combining the three in Figure 7.

For an area or region, Figure 7 shows that we have a simple aggregated assessment for all matrices, determinants and sites in a region with the relative proportion of all observations exceeding BAC and EAC. When considering suitable environmental targets for contaminants and their effects and the wording of Descriptor 8 in the Marine Strategy Framework Directive (MSFD), Good Environmental Status might be taken to mean that concentrations of contaminants and measurements of their effects should always be less than EAC. It should be borne in mind that when very large numbers of observations are made there is always the possibility that outliers are present and it would not be reasonable in such circumstances to have a 100% compliance target (or “one out all out”). Therefore SGIMC (ICES, 2011) proposed a pragmatic approach that

95% of measurements should be less than EAC (allowing for a 5% error rate). This target is represented as a horizontal red line in Figure 7.

Discussion

The assessment of the results from the ICON project shows that the framework provides a good and transparent reporting tool that makes it possible to present complex environmental monitoring datasets on contaminants concentrations and biological responses across multiple matrices, sites and seas. The key to the assessment is the development of the method- and species-specific criteria, which allows for the setting of thresholds of assumed equal significance for contaminants, exposure indicators and effect indicators, eventually allowing the different data types to be combined in a common indicator (cf. Vethaak et al., this issue-a). The flexibility and transparency is more extensive than frameworks proposed earlier, not least because contaminant concentrations and biological responses could be combined in a final assessment of environmental status. In addition, the ICON sampling campaign in European coastal and offshore areas provided a large dataset that resulted in a comprehensive and comparative evaluation of the state of selected European coastal and offshore marine areas.

The core methods included in the scheme were selected as the minimum set of contaminants and biological effects techniques that would need to be applied in order to determine whether contaminants are impacting on 'ecosystem health'. They achieve this by covering the main contaminant groups likely to cause such effects and that may be routinely monitored, as well as covering the main toxicity endpoints that are reasonably measurable in sentinel species, i.e. general toxicant stress, neurotoxicity, genotoxicity (Hylland et al., this issue-b), carcinogenicity (Lang et al., this issue-b), endocrine disruption (Burgeot et al., this issue), energetic costs (Martinez-Gomez et al., this issue-a) and mortality, as well as biomarkers of exposure to groups of compounds likely to have such effects. This core set of methods is not identical to, but similar to those suggested by under HELCOM (Lehtonen et al., 2014), but more extensive than methods suggested in e.g. Giltrap et al. (2013) and Hagger et al. (2008). Sediment bioassays are not mandatory in the OSPAR framework, but should comprise more than one method (as reported here). Sediment toxicity was addressed using different methods in Vethaak et al. (this issue – b).

There are environmental factors that may modulate biological responses, e.g. season. Data used to derive BAC and EAC were from studies where ICES guidelines for sampling have been adhered to, i.e. sampling outside the reproductive period. Criteria have been developed for selected species using hundreds and thousands of analyses as a basis, but there is an underlying assumption in this strategy that a species will respond to contaminant exposure in a similar fashion throughout its geographical range, all else being equal.

The biological responses selected for the framework comprise a range of methods that are sensitive to contaminant stress, including some that are specific to important contaminant groups and some that provide responses to a wide range of substances, including cumulative effects and effects from chemicals not directly monitored for. The integrated nature of the approach also identified instances where high concentrations of contaminants of concern were recorded, but where effects were not detected at a significant level. In these instances, contaminant availability may be limited and concentrations of limited concern as a result. In this case, the lack of effects in the assessment will down-weigh the importance of the contaminant result in an overall assessment. If the 95% target were to be used as a regional indicator of MSFD GES, Iceland and offshore North Sea would achieve the target using the ICON dataset, but inshore North Sea, Baltic and Spanish Mediterranean regions would fail.

Through applying the integrated assessment framework to the ICON dataset, several issues were identified that will need to be considered or spawn further research to improve the robustness of the framework. Because the assessment approach largely aggregates the results of applying thresholds to monitoring data at various levels of organisation and spatial scales, all data are treated equally in the assessment process and missing data will necessarily introduce less robustness into the overall assessment. Similarly, the introduction of additional data, for example from multiple matrices of the same type, e.g. multiple species of fish at the same site, can skew the assessment result. The ICON project has demonstrated that even on the scale of a large project with more than 20 partner institutions, data are likely to be missing from an assessment. In the current report, this has been dealt with by the use of 'grey' in the figures, so that the uncertainty of an assessment can be identified. It is further recommended that a 'robustness indicator' be developed in order to be able to quantify the quality of site assessments (see Martinez-Gomez et al., this volume – b). Such an indicator would be based on the relevance and completeness of the range of determinants comprising an assessment. Finally, the outcome of any integrated assessment has the potential to be strongly influenced by the selection of sites for the programme. At present there are no guidelines recommending a minimum number of sampling sites per region, appropriate statistical power for monitoring using this approach or how to account for hotspot or inshore sites in a wider scale regional assessment. Those are issues that need to be addressed to ascertain relevant and efficient marine monitoring in the future.

Conclusions

The ICON project has provided one of the most comprehensive integrated monitoring datasets of its kind and was found to be suitable for assessment using the framework developed within ICES and OSPAR. The approach is considered suitable for the determination of GES for Descriptor 8 under the MSFD.

The ICON project has shown that it is feasible to apply the OSPAR framework for integrated chemical and biological monitoring. The results show that Iceland has locations less impacted by contaminants than other locations in Europe, followed by offshore locations in the North Sea, with coastal locations being most clearly impacted.

The framework can be applied to datasets with missing data and determinants, but the validity of the assessment decreases with increasing missing data. Further guidance on minimal requirements for an integrated assessment and the development of a robustness indicator is suggested.

Assessment criteria for passive sampling techniques and *in vitro* bioassays need further development before they can be included in the integrated assessment framework.

There is a need to evaluate some assumptions in the OSPAR framework, e.g. that different populations of a species with a wide geographical coverage will respond similarly to contaminant exposure.

Acknowledgements

The authors wish to acknowledge the work by colleagues in ICES and OSPAR working group, i.e. WGBEC, WKIMON, SGIMC, as well as the cruise leaders, cruise participants and crews of R/V Walther Herwig III (Germany), R/V Scotia, R/V Alba na Mara (Scotland), R/V Gwen Drez (France) and R/V Endeavour (UK). The French participation was funded by IFREMER and ONEMA.

Literature references

Adams, S.M., Brown, A.M., Goede, R.W., 1993. A quantitative Health Assessment Index for rapid evaluation of fish condition in the field. Transactions of the American Fisheries Society, 122, 63-73.

Allen, Y., Scott, A. P., Matthiessen, P., Haworth, S., Thain, J. E., Feist, S. 1999. Survey of estrogenic activity in United Kingdom estuarine and coastal waters and its effects on gonadal development of the flounder *Platichthys flesus*. Environmental Toxicology and Chemistry, 18, 1791–1800.

Bocquene, G., Galgani, F., Burgeot, T., Le-Dean, L., Truquet, P., 1993. Acetylcholinesterase levels in marine organisms along French coasts. Marine Pollution Bulletin, 26, 101–106.

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

Bourlat, S.J., Borja, A., Gilbert, J., Taylor, M.I., Davies, N., Weisberg, S.B., Griffith, J.F., Lettieri, T., Field, D., Benzie, J., Glöckner, F.O., Rodríguez-Ezpeleta, N., Faith, D.P., Bean, T.P., Obst, M., 2013. Genomics in marine monitoring: New opportunities for assessing marine health status. Marine Pollution Bulletin, 74, 19–31.

Carney Almroth, B., Hultman, M., Wassmur, B., Sturve, J., Is oxidative stress evident in dab (*Limanda limanda*) in the North Sea? (this issue)

Chapman, P.M., McDonald, B.G., Lawrence, G.S., 2002. Weight-of-evidence issues and frameworks for sediment quality (and other) assessments. Human and Ecological Risk Assessment 8, 1489–1515.

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

Davies, I.M., Vethaak, A.D. (Eds.), 2012. Integrated monitoring of chemicals and their effects. ICES Cooperative Research Report 315, 227 pp.

Devin, S., Burgeot, T., Giamberini, L., Minguez, L., Pain-Devin, S., 2014. The integrated biomarker response revisited: optimization to avoid misuse. Environmental Science And Pollution Research, 21, 2448-2454.

Giltrap, M., Ronan, J., Hardenberg, S., Parkes, G., McHugh, B., McGovern, E., Wilson, J.G., 2013. Assessment of biomarkers in *Mytilus edulis* to determine good environmental status for implementation of MSFD in Ireland. Marine Pollution Bulletin 71, 240–249.

Hagger, J.A., Jones, M.B., Lowe, D., Leonard, D.R.P., Owen, R., Galloway, T.S., 2008. Application of biomarkers for improving risk assessments of chemicals under the Water Framework Directive: A case study. Marine Pollution Bulletin, 56, 1111–1118.

Hylland, K., Beyer, J., Berntssen, M., Klungsøyr, J., Lang, T., Balk, L. 2006. May persistent organic pollutants affect fish populations in the North Sea? Journal of Toxicology and Environmental Health, Part A, 69, 125-138.

Hylland, K., Gubbins, M.J., Robinson, C., Burgeot, T., Martínez-Gómez, C., Lang, T., Svavarsson, J., Thain, J.E., Vethaak AD. Integrated chemical and biological assessment of contaminant impacts in selected European coastal and offshore marine areas (this issue-a)

Hylland, K., Ruus, A., Grung, M., Green, N., 2009. Relationships between physiology, tissue contaminants and biomarker responses in Atlantic cod (*Gadus morhua* L.). Journal of Toxicology and Environmental Health, Part A, 72, 226-233.

Hylland, K., Skei, B.B., Gubbins, M.J., Lang, T., Brunborg, G., Le Goff, J., Burgeot, T., Genotoxicity in dab (*Limanda limanda*) and haddock (*Melanogrammus aeglefinus*) from European seas (this issue-b)

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

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

Kammann, U., Akcha, F., Budzinski, H., Burgeot, T., Gubbins, M.J., Lang, T., Le Menach, K., Vethaak, A.D., Hylland, K. PAH metabolites in fish bile: from the Seine Estuary to Iceland (this issue)

Lang, T., Feist, S.W., Stentiford, G.D., Bignell, J., Vethaak, A.D., Wosniok, W. Diseases of dab (*Limanda limanda*): analysis and assessment of data on externally visible diseases, macroscopic liver neoplasms and liver histopathology at offshore sites in the North Sea, Baltic Sea and off Iceland (this issue-a)

Lang, T., Kruse, R., Haarich, M., Wosniok, W. Methylmercury in dab (*Limanda limanda*) from the North Sea, Baltic Sea and Icelandic waters: relationship to host-specific variables (this issue-b)

Lyons, B., Thain, JE, Stentiford, GD, Hylland, K, Davies, I, Vethaak, AD. 2010. Using biological effects tools to define Good Environmental Status under the European Union Marine Strategy Framework Directive. Marine Pollution Bulletin, 60, 1647-1651.

Lyons B.P., Bignell, J.P., Stentiford, G.D., Bolam, T., Rumney, H.S., Bersuder, P., Barber, J., Askem, C.W., Maes T., Thain, J.E. Determining Good Environmental Status under the Marine Strategy Framework Directive: case study for descriptor 8 (chemical contaminants) (this issue)

Marigómez, I., Garmendia, L., Soto, M., Orbea, A., Izagirre, U., Cajaraville, M.P., 2013. Marine ecosystem health status assessment through integrative biomarker indices: a comparative study after the Prestige oil spill "Mussel Watch". Ecotoxicology, 22, 486–505.

Martínez-Gómez C., Burgeot T., Robinson, C.D., Gubbins, M.J., Halldorsson, H.P. Albentosa, M., Bignell J.P., Hylland, K., Vethaak A.D. Lysosomal membrane stability and Stress on Stress in mussels as common Pan-European contaminant-related biomarkers (this issue-a)

Martínez-Gómez C., Fernández B., Robinson, C.D., Campillo J.A., León V.M., Benedicto J., Hylland, K., Vethaak A.D. Assessing the good environmental status (GES) of the Cartagena coastal zone (W Mediterranean) using an integrated framework of chemical and biological effect data: a practical case study (this issue-b)

OSPAR, 2008. Draft Agreement on CEMP Assessment Criteria for the QSR 2010, , MON 09/8/1/6 Add.1. OSPAR Commission, London.

OSPAR, 2010. Quality Status Report 2010. OSPAR Commission, London, 176 pp.

Robinson, C.D., Webster, L., Martínez-Gómez, C., Burgeot, T., Gubbins, M.J., Thain, J.E., Vethaak, A.D., McIntosh, A.D., Hylland, K. Assessment of contaminant concentrations in sediments, fish and mussels sampled from the North Atlantic and European regional seas within the ICON project (this issue)

Sandstrom, O., Larsson, A., Andersson, J., Appelberg, M., Bignert, A., Ek, H., Forlin, L., Olsson, M., 2005. Three decades of Swedish experience demonstrates the need for integrated long-term monitoring of fish in marine coastal areas. Water Quality Research Journal of Canada, 40, 233–250.

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

Tornero, V., d' Alcalà, M.R., 2014. Contamination by hazardous substances in the Gulf of Naples and nearby coastal areas: A review of sources, environmental levels and potential impacts in the MSFD perspective. Science of the Total Environment 466, 820–840.

Vethaak, A.D., Davies, I.M., Thain, J.E., Gubbins, M.J., Martínez-Gómez, C., Robinson, C.D., Moffat, C.F., Burgeot, Maes, Wosniok, Giltrap , M., Lang, T., Strand, J., Hylland, K. Integrated indicator framework and methodology for monitoring and assessment of hazardous substances and their effects in the marine environment (this issue-a)

Vethaak, A.D., Hamers, T., Martínez-Gómez, C., Kamstra, J.H., de Weert , J., Leonards, P., Smedes, F. Toxicity profiling of marine surface sediments: a case study using rapid screening bioassays of exhaustive total extracts, elutriates and passive sampler extracts (this issue-b)

Vethaak, A. D., Jol, J. G., Martínez-Gómez, C., 2011. Effects of cumulative stress on fish health near freshwater outlet sluices into the sea: a case study (1988–2005) with evidence for a contributing role of chemical contaminants. Integrated environmental assessment and management, 7, 445-458.

Figure captions

Figure 1. Assessment of sediment data against BAC (background assessment criteria) and EAC (ecotoxicological assessment criteria); blue - below BAC, green - between BAC and EAC, red - above EAC, grey – data lacking; FoF = Firth of Forth.

Figure 2. Assessment of mussel data against BAC (background assessment criteria) and EAC (ecotoxicological assessment criteria); blue - below BAC, green - between BAC and EAC, red - above EAC; grey cells indicate core analyses not performed.

Figure 3. Assessment of imposex data (as VDSI) against BAC (background assessment criteria) and EAC (ecotoxicological assessment criteria); blue - below BAC, green - between BAC and EAC, red - above EAC; grey cells indicate analyses not performed.

Figure 4. Assessment of contaminant concentrations (liver), exposure and effects in fish from Iceland, the North Sea, Baltic Sea, Seine estuary (two sites) and Mediterranean Sea; LL – dab, PF – flounder, MA – haddock, MB - red mullet; blue - below BAC, green - between BAC and EAC, red - above EAC; grey cells indicate core analyses not performed; see Davies & Vethaak (2012) and relevant chapters for individual methods.

Figure 5. Assessment of contaminants, exposure and effects for the indicated locations in the North Sea (offshore); grey cells indicate core analyses not performed.

Figure 6. Assessment of contaminants, exposure and effects for each of the five areas. From left: Iceland (4 sites), coastal North Sea (10 sites), offshore North Sea (5 sites), German Baltic Sea (1 site) and Spanish Mediterranean Sea (2 sites). Numbers indicate data for each category.

Figure 7. Integrated assessment for each of the five areas. From left: Iceland (4 sites), coastal North Sea (10 sites), offshore North Sea (5 sites), German Baltic Sea (1 site) and Spanish Mediterranean Sea (2 sites). Numbers indicate data for each category; red line = 95% threshold.