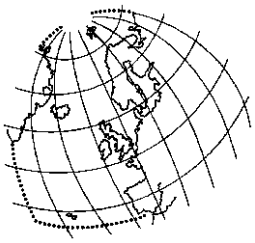
OSPAR CONVENTION FOR THE PROTECTION OF THE MARINE ENVIRONMENT OF THE NORTH-EAST ATLANTIC



JAMP Guidelines for Contaminant-Specific Biological Effects (OSPAR Agreement 2008-09)[1]

[1] Technical annexes revised in 2007. This agreement replaces agreement 2003-10 (addition of Technical Annex 4).

1. Introduction

Biological effects monitoring is an important element in programmes which aim to assess the quality of the marine environment since such monitoring can demonstrate links between contaminants and ecological responses. Biological effects monitoring can thus be used to indicate the presence of substances, or combinations of substances, not previously identified as being of concern and can also be used to identify regions of decreased environmental quality. Over the past few years research has enabled the development of a broad range of techniques for measuring biological effects. These have been assessed by ICES and a suite of tests suitable for incorporation into monitoring programmes has now been identified. To date relatively few biological effects measurements have been undertaken in international programmes and those which have are largely independent of any chemical monitoring. It is now recognised that a coordinated chemical and biological effects monitoring programme is essential for identifying areas of concern.

The Joint Assessment and Monitoring Programme (JAMP) includes a substantial biological effects monitoring component and this should be fully integrated with the use of diagnostic chemical analysis. The broad objectives of the programme will be to identify where contaminants are causing biological effects, to predict the highest organisational levels at which these effects occur (*i.e.* cell, individual, community or ecosystem) and to determine whether these effects result in harm to living resources and/or marine ecosystems, or otherwise interfere with other legitimate uses of the sea. This will provide a basis for:

a. the identification of areas of concern and the reasons for the concern;
b. an understanding of the causes of the observed biological effects;
c. decisions to be made concerning measures for controlling the pollution;
d. an assessment of the effectiveness of the measures in chemical and/or biological terms.

2. Purposes

Measurements of biological effects are carried out for, *inter alia*, the following purposes:

a. to establish the spatial distribution and extent of biological effects of specific contaminants in marine organisms;
b. to establish temporal trends in biological effects in order to estimate the magnitude of changes in biological effects with time.

3. Quantitative objectives

Monitoring activities are traditionally categorised into activities aimed at identifying variation in space (spatial monitoring) and those aimed at identifying variation in time (temporal monitoring). The same distinction is used in these guidelines, although there is nothing to prevent the two activities being carried out simultaneously, as long as this is incorporated into the design of the programme. For both types of monitoring it should be realised that objectives and sampling design are so closely related that they should be developed simultaneously. Moreover, to design a monitoring programme, some *a priori* knowledge of the system to be monitored is required. Alternatively, the objective of the monitoring activity can be to obtain such information, so that more focused monitoring can be done subsequently. In practice, this will always be an iterative process where information and experience obtained at a particular step will be used at the next. Three broad categories of objectives can be recognised; these are illustrated below with example objectives.

3.1 Exploratory sampling

The purpose of exploratory sampling is to estimate the level of a particular biological effect measurement at a particular time or place, to describe the normal range of values of the measurement and its spatial variability.

Example 1 to estimate the level of a particular biological effect at all points in an area, with a specified precision.

Example 2 To estimate a parameter or parameters (*e.g*. mean, median, 95 percentile) to describe a population of interest within an area, with a specified precision.

Example 3 to estimate a gradient in a biological effect measurement from a point source, with a specified precision.

3.2 Areas of concern

The area of concern must be defined for each particular programme. The purpose of such monitoring would be to locate an area or areas of concern and/or the measurement of the extent of such an area.

Example 4 to locate all areas of concern of a certain size in an area with a specified probability of success.

Example 5 to estimate the extent of a known area of concern with a specified precision.

3.3 Detection of change

Monitoring to detect either temporal or spatial changes over an area.

Example 6 To estimate the change in the level of a biological effect measurement (as described in Example 2) over a specified period of time, with a specified precision.

Example 7 to estimate the change in the spatial extent of an area of concern over a specified period of time, with a specified precision.

4. Sampling strategy

A stepwise approach is proposed for the application of biological effects monitoring techniques:

a. an assessment of contaminant concentrations in the Maritime Area, to identify areas for biological effects monitoring;

b. an identification of specific biological effects at contaminated sites and comparison with the situation at reference sites;

c. determination of the severity of damage or effects on organism health by the application of a suite of techniques, at different levels of biological organisation, which are collectively diagnostic of contaminant exposure and effects. The sampling strategy should enable integrated biological and chemical sampling and, where significant biological effects are measured and detailed information on contaminant concentrations are not available, an additional programme of analytical measurements should be undertaken to determine the scale and extent of local contamination;

d. initiation of a temporal trend sampling programme, for both polluted and reference sites to include chemical analyses and biological effects monitoring. Information on the variability of the biological markers at contaminated and reference sites obtained in the previous two stages will be used to design the strategy.

5. Quality assurance

Quality assurance (QA) is recognised as an essential component of the biological effects monitoring programme which is currently absent, or only partially in place, for most of the recommended techniques. Quality assurance for chemical analysis is well documented and provides a framework for establishing such procedures in biological assessments. However, due to the variability of biological end-points, the details of the procedures will necessarily be different from those used for chemical analyses. Quality assurance is the total management scheme required to ensure the consistent delivery of quality controlled information. It should include periodic inspection of test facilities, QA documentation, inter-laboratory performance testing exercises and a scheme of analytical quality control (within laboratory analytical check) applied to each test method. It is unlikely that full QA procedures can be established in the short-term in laboratories conducting biological assessments for OSPAR's purposes. Nevertheless, a minimum requirement, to ensure that comparable quality data is produced, must be established. This should include:

a. the adoption of only those methods which are referenced and have both a standard operating procedure (SOP) and associated analytical quality control (AQC) preferably following internationally agreed protocols;

b. staff trained to an agreed level of competence to conduct the test;

c. regular internal (within laboratory) calibration, including where possible, the introduction of blind samples during normal analysis and strict adherence to AQC procedures for each test method;

d. inter-laboratory performance assessment with the periodic circulation of samples for analysis by participating laboratories;

e. an action plan to respond to breaches of acceptable limits (limits established in AQC procedures or agreed for inter-laboratory performance assessment).

Full QA procedures require quality management plans and the appointment of independent QA officers. This imposes strict procedures for verifying and authorising data with the production of QA reports. It is necessary

to identify the level of quality that is required. The minimum scheme outlined could satisfy immediate requirements for reporting and for environmental quality status reports without further action. Where possible laboratories should adopt the same method and associated SOP/AQC procedures. Where this is not possible an exercise to demonstrate that the different methods are compatible and deliver the same end-point will be necessary if the reported data are to be compared. Quality assurance procedures can also be applied to sampling techniques although adoption of agreed procedures for sampling and their implementation by trained staff will be acceptable for immediate purposes. It is essential that all participating organisations agree a programme to introduce, at least, minimum QA procedures and make a commitment to their strict adherence.

Technical Annex 1

Metal-specific biological effects monitoring

1. Sampling strategy

The bioavailability of different metals in the marine environment will depend to varying extents on factors such as salinity, turbidity, dissolved organic matter, particulate organic matter, particle size distribution, organic content, redox conditions, Fe and Mn concentrations and biological activity. In addition, trophic status and feeding preferences will strongly affect the relative impact of sediment-associated metals on different benthic organisms. For benthic fish there are additional reasons for a lack of correspondence between sediment concentrations and concentrations in tissues, such as migration and individual feeding strategies. Metal levels in tissues are a reflection of both biologically active metal and metal irreversibly bound to inorganic and organic matrices. Therefore to establish the impact of metals on flatfish it is necessary to focus on the biological responses in the fish, although sediment concentrations are the obvious means to identify relevant sampling areas where impacts are likely to occur.

A strategy at four levels is proposed for the application of biological effects techniques in combination with past and current knowledge from chemical surveys including:

Stage 1

Evaluation of previous knowledge of contamination by metals. The first stage in the tiered approach will require a systematic assessment of metal contamination in sediments in the Maritime Area.

Stage 2

Identification of metal-specific effects from exposure to bioavailable metals at contaminated and reference sites. For the second stage there are only two currently available techniques that are regarded as specific for metals; metallothioneins (MT) in liver and δ-amino levulinic acid dehydratase (ALA-D) inhibition in blood. These techniques should be deployed using selected fish species.

Stage 3

Focused studies on areas with recognised metal pollution at concentrations capable of causing biological damage. For the third stage techniques for the measurement of 'antioxidant enzymes' and the products of oxidative damage in selected fish and the blue mussel, are recommended. These techniques provide an index of radical formation that can lead to carcinogenesis or cell death. It must be emphasised that these biological effect indices are not entirely limited to metals in their response. However, they are the only suitable techniques that are available to be used at this time.

Stage 4

Initiation of a temporal trend-sampling programme, for both polluted and reference sites and to include chemical analyses and biological effects monitoring. For the fourth stage the above techniques will be deployed at polluted sites and reference locations for temporal trend monitoring. Information on the variability of the biological markers at contaminated and reference sites obtained in stages two and three will be used to design the strategy.

To date there are no well proven biological effect techniques available for metals that could be used to assess specific effects at the population and community level although work in Norway has indicated a possible cause-effect relationship between metal-rich discharges and changes in benthic community structure (Rygg, 1985). The techniques proposed in these guidelines are only applicable at the level of the individual.

2. Sampling locations

The metal-specific biological effects monitoring programme is used to assess both local effects near point sources and larger scale effects within the Maritime Area. A variety of sources for metal inputs to the maritime area can be recognised, some are local *e.g*. offshore and fertiliser industries, others have a wider impact range *e.g*. estuaries and atmospheric inputs. The spatial distribution survey should employ a general strategy for defining biological effects of metals at representative industrialised estuaries and at offshore locations to allow for the assessment of biological effects of metals throughout the Maritime Area. Significant point sources identified during an assessment of monitoring data, may require additional, more intensive investigation to describe local elevations in metal concentrations. Within the framework of the spatial distribution survey a smaller number of sites should be selected for annual monitoring in order to assess long-term improvements or deterioration of environmental quality.

3. Methods to be used

3.1 Biological indicators of metal exposure

At present there are only two methods that are reasonably specific for one or more metals and that have been evaluated to any extent for aquatic organisms - metallothionein (MT; Cu, Zn, Cd and inorganic Hg) and δ-amino levulinic acid dehydratase (ALA-D; Pb).

Metallothionein

Metallothionein is present in the tissues of all vertebrates and many invertebrates. The major role of the protein under normal metabolism is presumed to be the homeostatic regulation of intracellular Cu-availability and/or Zn-availability. Metallothionein is induced by and binds non-essential metals (Cd, inorganic Hg) in addition to the two essential metals Cu and Zn. Hepatic MT in fish has been thoroughly evaluated as a marker for metal stress under both field and laboratory conditions (Olson and Haux, 1986; Sulaiman *et al*., 1991; Hogstrand and Haux, 1990a and 1991 and Hogstrand *et al*., 1991). The potential of MT-measurement in clarifying the relative effects of Cu, Zn and Cd was demonstrated during the IOC/ICES workshop at Bremerhaven (Hylland *et al*., 1992), where it was shown that MT in male dab was elevated due to hepatic Cu and Cd - presumably caused by environmental contamination, whereas MT in female dab was primarily affected by Zn - an endogenous, pre-spawning effect. Without the determination of MT this information would have been lost.

Elevated concentrations of hepatic MT in fish from contaminated sites compared to reference sites indicate that fish at the contaminated sites have been affected by environmental contamination by Cu and/or Zn and/or Cd. To quantify the relative contribution of each metal and biological factor hepatic concentrations of Cu, Zn and Cd together with fish size and gonadosomatic index (GSI) should be included as factors in a multiple regression model with MT as the dependent variable (this analysis should be done for each sex).

Inorganic Hg has been shown to be a strong inducer of MT in various *in vitro* studies (fish cell cultures), but its *in vivo* influence has not been properly evaluated mainly because the majority of Hg in tissue is present in the form of organic-mercury complexes. Therefore, at present, the relationship between Hg and MT in tissues is not clear and so MT is not currently recommended as a marker for contamination by Hg. Although metal-inducible metallothioneins have been identified in various marine invertebrates, *e.g*. mussel, shore crab, lobster and sea urchin, those proteins have only been used to a limited extent in field studies. At present MT in invertebrates is not recommended for wider adoption in monitoring programmes.

ALA-D

Lead causes a dose-dependant inhibition of δ-amino levulinic acid dehydratase (ALA-D) which is an enzyme essential for the synthesis of haemoglobin in the haemopoietic tissue. ALA-D inhibition is therefore a good indicator of Pb exposure which is maximally inhibited before other signs of Pb toxicity become apparent (Haux *et al*., 1986). The inhibition of ALA-D by Pb is quite specific but has only been used to monitor Pb exposure in fish, birds and mammals. The effects of Pb on ALA-D activity in fish blood cells have been evaluated in both laboratory and field studies in Canada (Hodson, 1976; Schmitt *et al.*, 1984) and Sweden (Larsson *et al.*, 1985; Haux *et al*., 1986). Although most published studies concern freshwater fish the technique has also been successfully applied to marine fish. The methodology is simple, inexpensive and precise and can easily be included as a standard procedure in a monitoring programme.

Inhibition of ALA-D in fish from contaminated sites compared to fish from an appropriate reference site indicates that fish at the contaminated sites have been affected by environmental contamination by Pb. A negative relationship is to be expected between ALA-D and Pb in blood cells.

3.2 Evidence for deleterious effects

Formation of oxygen-derived radicals is a natural process in living cells. This process is enhanced by high levels of metal pollution in a way that is harmful to living organisms. Increased activity in a suite of enzymes such as superoxide dismutase (SOD), catalase (CAT) and glutathione peroxidase (GPX) which in combination are indicative of oxidative stress can provide an index of heavy metal pollution. Lipid peroxidation is a process induced by highly reactive oxygen species such as hydrogen peroxide and OH-radicals, derived from oxidative stress. The status of these biochemical indices of oxidative damage have been reviewed recently (Stegeman *et al*., 1992).

In regions where there is a recognised problem of metal contamination the activity of antioxidant enzymes and the formation of oxidised substrates measured in the liver of fish (dab, flounder and cod) and in mussels can be used to establish effects on biota. The antioxidant techniques will only be applied to the restricted number of locations that are identified by stage three as being highly contaminated by metals. To facilitate interpretation transects should be completed across hotspots with an appropriate reference site.

Increased activity of antioxidant enzymes in fish or mussels from or at contaminated sites compared to fish and/or mussels from a reference site would indicate that the stress from metals is sufficient as to increase the risk of carcinogenesis or cell death. Care must be taken in the interpretation of the data however as other environmental contaminants may also induce affect this effect.

4. Temporal trend monitoring

To design a temporal trend monitoring programme it is necessary to have prior knowledge about the variability of each parameter, (*i.e.* MT, ALA-D, antioxidant enzymes) and the programme will therefore be designed using results from stages 2 and 3. The number of samples needed to enable the identification of temporal trends will be calculated using results from stages 2 and 3. Since year-to-year variability is to be expected it is recommended that samples be taken each year.

5. Field sampling and sampling equipment

Fishing gear appropriate for providing live fish in good condition should be used. For example if using Otter trawls then short hauls of less than 30 minutes duration are most appropriate. Fish should be maintained alive prior to individual sampling in flowing or frequently replenished seawater. Dead animals or animals severely damaged during the capture process should not be sampled. Hydrographic data, including salinity and temperature, should be recorded at each site at the time of sampling.

Fish should be sampled outside the spawning season. General guidance would be that sampling should take place at least one month after spawning and that this optimum sampling window may be different for each species and for individual regions within the Maritime Area. If the detection of temporal trends is one of the objectives of the monitoring then the sites should be sampled at the same time of year (within two weeks) for each annual sampling cycle. A random sample of between 10 and 20 individuals should be used. Either male or female individuals can be used but data from males and females should not be mixed and the same sex should be chosen for all samples (areas or time points) that are to be compared.

Fish need to be of comparable and sufficient size (different for each species) to enable three liver samples of approximately one gram (or more) each. The fish used should be consistent with those specified in the Guidelines for Monitoring Contaminants in Biota and should be sampled within a species-dependant size range because of the occurrence of age-related pathology. Thus the first choice species are:

- Dab (*Limanda limanda*) in the size range 20-25cm;
- Cod (*Gadus morhua*) in the size range 30-45cm;
- Flounder (*Platichthys flesus*) in the size range 15-35cm.

Generally flounder and cod should be two-year-old fish.

Each fish should be assigned a unique specimen number and the following information recorded:

- site number and name at which the animal was captured. Sites should be identified by geographical co-ordinates (latitude and longitude);
- date of sampling;
- taxonomic identification;
- gender;
- fish weight (in g) and the total fork length (in mm);
- location and description of grossly visible anomalies, lesions and parasites;
- the tissues collected and for what purpose.

The inclusion of fish with grossly visible external lesions and parasites may increase the variance of some biological effects measurements. For some purposes (*e.g*. the detection of temporal trends) it may be appropriate to only use apparently healthy fish in the data analysis.

A blood sample for ALA-D determination should be taken, prior to or within five minutes of killing the fish, from the caudal vein using a heparinised syringe and needle.

Fish should be sacrificed individually by severing the spinal cord just posterior to the brain. The liver should be dissected out, taking care not to contaminate it with bile or other tissues, weighed if practicable and then kept ice cold (0-4°C) during all subsequent processing. The liver should be trisected according to Technical Annex 2, (Figure 1); one part for measurement of MT, one part for the determination of antioxidant enzymes and one part for the analysis of Cu, Cd and Zn. Within any particular programme the dissection of the liver should be standardised so that the same part of the liver should always be used for any particular determinand. All tissues should be kept ice cold (0-4°C) during processing and should be freeze clamped in liquid nitrogen within 5 minutes of killing the fish. The gonad should be dissected out and weighed. The stage of maturation of the gonad, gonad somatic index (GSI) and liver somatic index (LSI) should be recorded. Otoliths should be removed for ageing the fish.

6. Storage and pre-treatment of samples

The blood sample should either be centrifuged directly after sampling (2000 x g for 5 min) and the blood cells (pellet) frozen in liquid nitrogen or the whole blood frozen directly in liquid nitrogen. Liver tissues should be freeze-clamped in liquid nitrogen. Both liver and blood tissue for biochemical analysis are most conveniently stored in cryovials at -70°C or in liquid nitrogen for up to one year before further processing. Tissues for trace metal analysis (Cu, Cd, Zn and Pb) should be stored and processed according to the Guidelines for Monitoring Contaminants in Biota (Technical Annex 2 Section 4).

7. Analytical procedures

7.1 Metallothionein

Methodology

Three methods are proposed:

a. immunochemical (RIA/ELISA). The recommended method is that of Hogstrand and Haux (1990b);

b. differential pulse polarography. The recommended method is that of Olafson and Sim (1979);

c. spectrophometric assay (adapted from Viarengo *et al.*, 1997). The described method was established for molluscs, but has also been evaluated for fish tissues. The following equipment is needed to perform the assay: homogenising equipment, a refrigerated centrifuge with fixed-angle and swing-out rotors, a spectrophotometer (or plate reader). Calf liver RNA is available from Sigma, St. Louis, MO, USA. The remaining reagents should all be of analytical grade and are available from most major suppliers of laboratory chemicals.

   Liver tissue should be homogenised and centrifuged at 10 000g to yield an S9 fraction. One ml of the resulting S9 supernatent is transferred to a reaction tube, 1,05 ml cold (-20°C) ethanol and 80 ml chloroform added (Kimura *et al.*, 1979). The mixture is then centrifuged at 6000g for 10 min (4°C). The supernatant is collected, mixed sequentially with 1 mg RNA, 40 ml 37% HCl and cold ethanol to a final concentration of 87%. The reaction mixture is kept at -20°C for 1 hr, then centrifuged in a swing-out rotor at 6000g for 10 min. The MT-containing pellet is washed in 87% ethanol/1% chloroform in homogenising buffer, re-centrifuged at 6000g for 10 min, the supernatant removed and the MT-containing pellet dried under a flow of nitrogen.

   The assay uses the reaction of DTNB with sulfhydryl groups described by Ellman (1958, 1959). The MT-containing pellet is resuspended in 150 ml 250 mM NaCl. Subsequently 150 ml 1N HCl with 4 mM EDTA is added. A volume of 4.2 ml of 2M NaCl, pH adjusted to 8.0 by 0.2M Na-phosphate, with 0.43 mM DTNB (5.5-diothiobis2-nitrobenzoic acid), was then added to the MT-containing solution at room temperature. Following mixing and centrifugation at 3000g for 5 min the absorbance of the supernatant was read at 412 nm. The concentration of MT is calculated from a standard curve.

Quality assurance

Reference material from common pools of minced liver containing high levels (induced by injection of Cd) and low levels of metallothionein for selected species (proposed species are cod, dab and flounder) should be distributed to participating laboratories. The material should be frozen in liquid nitrogen in a small volume of buffer (with protease inhibitors and reducing agent) and shipped to participating laboratories on dry ice. Analyses should be performed within 3-4 weeks.

Metallothionein standards for the selected species (proposed species are cod, dab and flounder) should be distributed to participating laboratories.

7.2 δ-amino levulinic acid dehydratase inhibition in blood (ALA-D)

Methodology

The recommended method is based on those described by Hodson (1976). The δ-amino levulinic acid dehydratase enzyme catalyses the formation of porphobilinogen from δ-aminolevulinic acid in the haem synthesis pathway. The determination of ALA-D activity is carried out using lysed red cells, measuring the amount of porphobilinogen formed from added δ-amino levulinic acid. Porphobilinogen is assayed spectrophotometrically after reaction with Ehrlich's reagent. The method to be followed is described in Johannson-Sjobeck and Larsson (1978). The protein content of the red cells used in the assay should be

determined and the ALA-D activity should be expressed in units of porphobilinogen (PBG) formed per unit time and protein (nmol PBG/mg protein/min).

Quality assurance

Reference material from common pools of heparinized whole blood with high and low (injected with Pb) levels of ALA-D activity for selected species (proposed species are cod, dab and flounder) should be distributed to participating laboratories. Material should be frozen in liquid nitrogen and shipped to participating laboratories on dry ice. Analyses should be performed within 3-4 weeks.

7.3 Oxidative stress

Methodology

The methods for tissue preparation and the determination of superoxide dismutase (SOD), glutathione reductase, glutathione peroxidase (GPX) and malonedialdehyde, catalase (CAT) and tissue lipid peroxides is given in Livingstone *et al*., (1993).

Quality assurance

Reference material from common pools of minced liver for selected species (proposed species are cod, dab and flounder) should be distributed to participating laboratories. Material should be frozen in liquid nitrogen in a small volume of distilled water or buffer and shipped to participating laboratories on dry ice. Analyses should be performed within 3-4 weeks.

8. Reporting requirements

Data reporting should be in accordance with the requirements for National Comments and with the latest ICES reporting formats. The following data are required

8.1 Contaminants

- Cu, Zn, Cd in liver ($\mu g\ g^{-1}$ liver wet weight);
- Pb in blood ($\mu g\ ml^{-1}$);
- Cu, Cd, Pb and Zn in sediments.

8.2 Biological effects measurements

- metallothionein ($\mu g\ g^{-1}$ liver wet weight) and the standard used (*e.g*. rabbit MT-1 Sigma);
- ALA-D activity in units of porphobilinogen (PBG) formed per unit time and protein (ηmol PBG mg protein^{-1} min^{-1}). The protein assay and standard (*e.g*. IgG, BSA) used should also be reported;
- malonaldehyde (lipid peroxide) ηmol g^{-1} wet weight;
- matalase mmol min^{-1} g^{-1} wet weight;
- glutathione peroxidase μmol min^{-1} g^{-1} wet weight;
- superoxide dismutase Units g^{-1} wet weight. Where 1 Unit is the amount of activity resulting in 50% inhibition of cytochrome c reductase as described in Livingstone *et al.,* (1993).

8.3 Supporting parameters

- fish identifier;
- site code and name;
- taxonomic identification;
- temperature and salinity at each collection site;
- date of sample collection;
- gender;
- location and description of grossly visible anomalies and lesions;
- liver somatic index;
- gonad somatic index;
- stage of gonadal maturation;
- fish weight (in g) and length (total fork length in mm);
- age of the fish.

9. References

Ellman, G.L., 1958. A colorimetric method for determining low concentrations of mercaptans. Arch. Biochem. Biophys. 74:443-450.

Ellman, G.L., 1959. Tissue Sulfhydryl Groups. Arch. Biochem. Biophys. 82: 70-77.

Haux, C., Larsson, A., Lithner, G. and Sjobeck, M.-L., 1986. A field study of physiological effects on fish in lead-contaminated lakes. Environ. Toxicol. Chem. 5:283-288.

Hodson, P.V., 1976. δ-Aminolevulinic acid dehydratase activity of fish blood as an indicator of harmful exposure to lead. J. Fish. Res. Bd. Can. 33:268-271.

Hogstrand, C. and Haux, C., 1990a. Metallothionein as an indicator of heavy-metal exposure in two subtropical fish species. J. exp mar. Biol. Ecol. 138:69-84.

Hogstrand, C. and Haux, C., 1990b. A radioimmunoassay for perch (*Perca fluviatilis*) metallothionein. Toxicol. Appl. Pharmacol. 103:56-65.

Hogstrand, C. and Haux, C., 1991. Binding and detoxification of heavy metals in lower vertebrates with reference to metallothionein. Comp. Biochem. Physiol. 100C:137-142.

Hogstrand, C., Lithner, G. and Haux, C., 1991. The importance of metallothionein for the accumulation of copper, zinc and cadmium in environmentally exposed perch, *Perca fluviatilis*. Pharmacol. Toxicol. 68:492-498.

Hylland, K., Haux, C. and Hogstrand, C., 1992. Hepatic metalothionein and heavy metals in dab *Limanda limanda* from the German Bight. Mar. Ecol. Prog. Ser. 91:89-96.

Johansson-Sjobeck, M.-L. and Larsson, Å., 1978. The effect of cadmium on the hematology and on the activity of δ-amino levulinic acid dehydratase (ALA-D) in blood and haemopoetic tissues of the flounder *Pleuronectes platessa* L. Environ. Res. 17: 191-294.

Kimura, M., Otaki, N. and Imano, M., 1979. Rabbit liver metallothionein: tentative amino acid sequence of metallothionein B. In: Metallothionein. Kägi, J.H.R. and Nordberg, M. (Editors), Experientia Suppl. Birkhauser, Basel, Vol. 24: 163-168.

Larsson, A., Haux, C. and Sjobeck, M., 1985. Fish physiology and metal pollution: results and experiences from laboratory and field studies. Ecotox. Environ. Saf. 9:250-282.

Livingstone, D.R., Lamaire, P., Matthews, A., Peters, L., Burke, D. and Law, R.J., 1993. Pro-oxidant, antioxidant and 7-ethoxyresorufin *O*-deethylase (EROD) activity responses in liver of dab (*Limanda limanda*) exposed to sediment contaminated with hydrocarbons and other chemicals. Mar. Pollut. Bull. 26 (11): 602-606.

Olafson, R.W. and Sim, R.G., 1979. An electrochemical approach to quantitation and characterisation of metallothioneines. Anal. Biochem. 100:343-351.

Olsson, P.-E. and Haux, C., 1986. Increased hepatic metallothionein content correlates to cadmium accumulation in environmentally exposed perch (*Perca fluviatilis*). Aquat. Toxicol. 9:231-242.

Rygg, B., 1985. Effect of sediment copper on benthic fauna. Mar. Ecol. Prog. Ser. 25:83-89.

Schmitt, C.J., Dwyer, F.J. and Finger, S.E., 1984. Bioavailability of Pb and Zn from mine tailings as indicated by erythrocyte δ-aminolevulinic acid dehydratase (ALA-D activity in suckers) (Pisces: Catostomidae). Can. J. Fish. Aquat. Sci. 41:1030-1040.

Stegeman, J.J., Brouwer, M., DiGiulio, R.T., Forlin, L., Fowler, B.A., Sanders, B.M. and Van Held, P.A., 1992. Molecular responses to environmental contamination: enzyme and protein systems as indicators of chemical exposure and effect. In: Biomarkers: Biochemical, Physiological and Histological Markers of Anthropogenic Stress. Huggett, R.J., Kimmerle, R.A., Merhle, P.M. Jnr. and Bergman, H.L. (Editors). Michigan, USA: Lewis Publishers.

Sulaiman, N., George, S. and Burke, M.D., 1991. Assessment of sublethal pollutant impact on flounders in an industrialised estuary using hepatic biochemical indices. Mar. Ecol. Prog. Ser. 68:207-212.

Viarengo, A., Ponzano, E., Dondero, F., Fabbri, 1997. A simple spectrophotometric method for metallothionein evaluation in marine organisms: an application to Mediterranean and Antarctic molluscs. Mar. Environ. Res. 44 (1) 69-84.

Technical Annex 2[2]

PAH-specific biological effects monitoring

1. Sampling strategy

The combination of five biomarker techniques is used to describe the impact of PAH compounds on biota at the biochemical and at the cellular/tissue level. The biomarker techniques selected are CYP 1A activity, bulky aromatic-DNA adducts PAH metabolites in bile, liver histopathology and macroscopic liver neoplasms (= liver nodules). These indicators can be considered as an interconnecting series since planar PAHs are effective in inducing CYP 1A enzyme and some of their members are metabolised to reactive epoxides forming DNA and protein adducts which are linked to mutagenesis and carcinogenesis as well as other potentially important deleterious, e.g. pathological, effects (*e.g*. the formation of neoantigens).

The suite of biomarkers give a measure of exposure and biological effects of which the observation of DNA adducts, liver histopathology and macroscopic liver neoplasms can be marked as a deleterious effect.

Increase of CYP 1A is a biomarker of exposure and to some extent also of biological effects as this response can also be linked to a number of toxicological effects. The induction of CYP 1A is a biochemical change within the organism which often precedes the onset of more serious cellular and physiological changes such as hepatic damage, reproductive toxicity, immunotoxicity *etc.* The CYP 1A enzyme system is known to activate certain PAHs to reactive, more toxic species which may otherwise be generally cytotoxic, mutagenic or carcinogenic. Most interest to date has centred around a role for CYP 1A in increasing certain types of chemical carcinogenesis. The Ah receptor involved with induction of CYP 1A has also been linked to various types of chemical toxicity *e.g*. cataract formation, retinal degradation, atherosclerosis and immune function (see references in Ioannides and Parke, 1993; Bock, 1993; Payne *et al*., 1994).

PAH metabolites are not indicative of biological effects *per se* but can provide a sensitive marker of exposure to bioavailable levels of PAHs in the environment. PAH bioavailability can vary markedly in different fish species living in environments similarly contaminated with PAHs.

The covalent binding of reactive PAH metabolite species to DNA results in the formation of bulky DNA adducts. Such adducts are believed to be important in the initiation of carcinogenesis by many chemicals. The measurement of DNA adducts in humans is currently being used to assess their exposure to genotoxic compounds in the workplace and general environment (see review and references therein by Jones and Parry, 1992; Phillips *et al*., 1993). DNA adducts have also been used to assess exposure of fish and other wildlife to environmental genotoxins including recently in the North Sea (Fisser and Blömeke, 1992). The results of field and laboratory studies show that hepatic DNA adducts detected in fish are associated with increased exposure to environmental PAH (Stein *et al*., 1993, 1994).

Exposure to PAHs and the resulting formation of toxic DNA adducts are known to induce a range of neoplastic and non-neoplastic liver lesions, particularly in bottom-dwelling fish such as flatfish species (Feist *et al.*, 2004). Therefore, PAH-specific biological effects monitoring should include studies on histopathological changes with a focus on those linked to carcinogenesis considered to be an endpoint of PAH exposure. Histopathological liver changes, including non-specific, early toxicopathic non-neoplastic, pre-neoplastic and neoplastic lesions, can be detected by applying a routine histopathological study in a random sample of fish. Another approach is to quantify the occurrence of macroscopic liver neoplasms (benign and malignant tumours) by recording the prevalence of grossly visible liver nodules in a sample of fish and a subsequent confirmation of the neoplastic nature of the nodules by histopathology.

Biological effects attributable to PAH are more difficult to assess in shellfish, therefore finfish are selected as the principal monitoring species. The choice of fish species to be monitored will reflect availability throughout the Maritime Area. Dab (*Limanda limanda*) is common throughout the North Sea, the English Channel and the Irish Sea but is not generally available in the southern part of the Maritime Area. The dragonet (*Callionymus lyra*) and plaice (*Pleuronectes platessa)* are suitable alternative species. For estuarine

2 This technical annex was last updated in 2007 to reflect the recommendation from ICES in 2005 that monitoring of PAH-specific biological effects as part of the CEMP should include the monitoring of macroscopic liver neoplasms (=liver nodules) (which are now only included in the CEMP General Biological Effects Monitoring component) because this technique detects the prevalence of macroscopically visible neoplastic liver lesions (histologically confirmed liver tumours). Neoplastic liver lesions are one of the best documented biological effects of exposure to PAHs. In the monitoring, the quantification of liver nodules can easily be combined with the studies on liver histopathology (which are already part of the PAH-specific Biological Effects Monitoring component of the CEMP). Further amendments are based on recommendations made in the CEMP background on externally visible diseases, liver nodules and liver pathology including liver neoplasia/hyperplasia.

monitoring the flounder (*Platichthys flesus*) is recommended. Other species may be chosen according to particular regional concerns. However, for the purpose of diagnosing specific PAH-mediated effects, demersal species are preferred.

2. Sampling locations

The PAH-specific biological effects monitoring programme is used to assess both local effects near point sources and larger scale effects within the Maritime Area. A variety of sources for PAH inputs to the Maritime Area can be recognised, some are local *e.g*. oil production; others have a wider impact range *e.g*. estuaries and atmospheric inputs of combustion products. The spatial distribution survey should employ a general strategy for defining biological effects of PAHs at representative industrialised estuaries and at offshore locations to allow for the assessment of biological effects of PAHs throughout the Maritime Area. Significant point sources may require additional, more intensive investigation to describe local elevations in PAH concentration.

3. Methods to be used

Detailed information on the recommended methodology is given in Section 7.

3.1 Cytochrome P4501A

Cytochromes P450 linked mono-oxygenases play a central role in the metabolism of endogenous substrates such as steroid hormones and vitamins as well as foreign compounds. CYP 1A are unique in that markedly increased levels as well as variant forms are commonly found in the tissue of animals exposed to different types of inducing compounds. Inducers of environmental importance include PAHs, PCBs and dioxins which effect induction via a cytosolic protein (the so-called Ah receptor) binding to DNA. EROD (7-ethoxyresorufin *O*-deethylase) enzyme activity reflects the presence of induced cytochrome P450IA (EROD), the primarily inducible form in fish (see reviews and references therein by Payne *et al*., 1987; Goksøyr and Forlin, 1992; Stegeman *et al*., 1992).

Over the last few years a number of EROD studies have been carried out in the Maritime Area and in the Mediterranean Sea (La Faurie *et al*., 1989; Narbonne *et al*., 1991; Renton and Addison, 1992; Lange *et al*., 1995; Sleiderink *et al*., 1995; Kohler and Pluta, 1995, Stagg *et al*., 1995, Burgeot *et al*., 1995; Goksøyr *et al*., 1991; Beyer *et al.*, 1996).

Following the revision of EROD protocols resulting from experience gained during the North Sea Task Force Monitoring Master Plan programme and various ICES activities (*e.g*. intercalibration exercises, field experiments), it is concluded that the EROD catalytic enzyme assay is the technique recommended for monitoring CYP 1A activity. Immunoassays for CYP 1A may also be used and may give additional information in cases where contaminant exposure is sufficient to cause inhibition of catalytic activity of CYP 1A.

3.2 DNA adducts

The recommended technique for measuring DNA adducts is the P-32 post-labelling technique. The primary reason for using the P-32 technique is its high sensitivity (detection of one adduct in 109-1010 nucleotides), its requirement for small amounts of DNA and its ability to detect carcinogenic DNA adducts of unknown structure. For interpretation it is important to realise that this technique may also measure adducts from chemicals other than PAHs.

3.3 PAH metabolites

The degree to which marine organisms are exposed to PAH is difficult to assess by measuring body burdens since most biota extensively metabolise such compounds. However screening methods for fluorescent metabolites in bile have proved useful for detecting exposure to PAH. Simple direct analysis of bile or metabolic hydrolysates by fluorimetry are suitable for laboratory studies and have also been used in the field. The recommended method is a coupled HPLC/fluorimetric procedure, which is more specific using a combination of excitation/emission wavelengths to detect 2, 3 and 4-5 ring compounds.

3.4 Liver histopathology

Major histological changes in liver tissue involving such endpoints as neoplasms, , inflammation or cellular necrosis are pathological in nature. Analysis can be carried out by standard histological procedures on tissues embedded in paraffin. The categories of histopathological liver lesions to be monitored include the four histopathology categories published by Feist *et al.* (2004) plus a “new” category, the non-specific liver lesions (see Table 1), reflecting the BEQUALM categories:

- Non-specific lesions
- Early toxicopathic non-neoplastic lesions
- Foci of cellular alteration
- Benign neoplasms
- Malignant neoplasms

3.5. Macroscopic liver neoplasms

Macroscopic liver neoplasms are tumours that can be identified as grossly visible nodules on the surface of the fish liver. Their prevalence can easily be recorded during fish disease surveys and in the course of the sampling of fish liver tissue for histopathological studies (see under 3.4). The method encompasses a two-steps-procedure: (1) quantification of macroscopic liver nodules > 2 mm and (2) a subsequent histopathological confirmation of the neoplastic nature of the lesion, implying that all nodules recorded are preserved in suitable fixative. Using this method, the prevalence of benign and malignant liver neoplasms in a given sample of fish is recorded.

4. Temporal trend monitoring

Within the framework of the spatial distribution survey a smaller number of sites should be selected for annual monitoring in order to assess long-term improvements or deterioration in environmental quality.

5. Field sampling and sampling equipment

All samples taken to measure PAH-specific biological effects should be taken from live fish in good condition. Therefore, appropriate fishing gear and practises should be used. For example, if using Otter trawls, then short hauls of a maximum of 30-60 minutes duration are most suitable. Fish should immediately be sorted from the catch and be maintained alive prior to individual sampling in flowing or frequently replenished seawater of ambient water temperature. Dead animals or animals severely damaged during the capture process should not be sampled. Hydrographic data, including salinity and temperature, should be recorded at each site at the time of sampling. Temperature has *e.g.* been observed to influence EROD activity and therefore bottom water temperature between sites should be as similar as possible or appropriate controls should be incorporated into the monitoring strategy.

Fish should be sampled outside the spawning season. General guidance would be that sampling should take place at least one month after spawning and that this optimum sampling window may be different for each species and for individual regions within the Maritime Area. If the detection of temporal trends is one of the objectives of the monitoring then the sites should be sampled at the same time of year (preferably within two weeks) for each annual sampling cycle. A random sample of between 10 and 50 individuals should be used, depending on the biomarker to be measured. Either male or female individuals can be used but data from males and females should not be mixed and the same sex should be chosen for all samples that are to be compared (in terms of comparison between different areas or at different points in time within the same area).

Fish need to be of comparable and sufficient size (different for each species) to enable a sufficient number of liver tissue samples, approximately one gram (or more) each. The fish used should be consistent with those specified in the JAMP Guidelines for the Monitoring of Contaminants in Biota and for General Biological Effects Monitoring. However, they should be sampled within a narrower species-dependant size range than given in the Guidelines for the Monitoring of Contaminants in Biota because of the occurrence of age-related biological effects, *e.g.* pathology. Thus the first choice species and their size ranges are:

- Dab *(Limanda limanda)* in the size range 20 - 24 cm (≥ 25 cm for macroscopic liver neoplasms);
- Cod *(Gadus morhua)* in the size range 30 - 45 cm;
- Flounder *(Platichthys flesus)* in the size range 25 - 29 cm (≥ 30 cm for macroscopic liver neoplasms)

Each fish should be assigned a unique specimen number and the following information should be recorded:

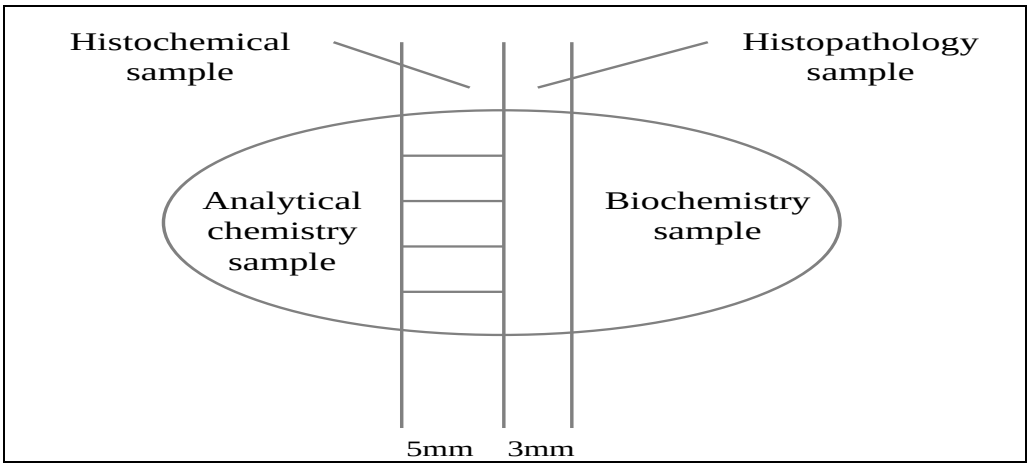
- site (station) number and name at which the animal was captured. Sites should be identified by geographical co-ordinates (latitude and longitude);
- date of sampling;
- taxonomic identification;
- gender;
- disease status;
- fish weight (in g) and the total length (to the cm below);
- location and description of grossly visible anomalies, lesions and parasites;
- the number of samples collected.

The inclusion of fish with grossly visible external lesions and parasites may increase the variance of some of the biological effects measurements. For some purposes (*e.g*. the detection of temporal trends) it may be appropriate to only use apparently healthy fish in the data analysis.

Fish should be sacrificed individually by severing the spinal cord just posterior to the brain. The liver should be dissected out, taking care not to contaminate it with bile or other tissues, weighed if practicable and then either fixed in appropriate fixatives or kept ice cold (0-4°C) during all subsequent processing. Prior to sectioning the liver it should be examined externally for the presence of liver nodules > 2 mm in diameter and any part with a nodule or lesion should be removed for pathological examination. The liver should be dissected according to Figure 1; one part for measurement of CYP1A (EROD) and DNA adducts, one part for chemistry and one part for histopathology . Within any particular programme the dissection of the liver should be standardised so that the same part of the liver should always be used for any particular determinand.

The tissue for chemistry should be processed according to the Guidelines for the Monitoring of Contaminants in Biota (Technical Annex 1 - Organic contaminants). The tissues for biochemistry should be freeze-clamped in liquid nitrogen within 5 minutes of killing the fish. The tissues for histopathology should be fixed in 10% neutral buffered formalin for paraffin embedded material or if cryopreserved tissues are to be used then freeze-clamped in liquid nitrogen within 5 minutes of killing the fish. The gall bladder should be removed intact and frozen at -20°C. Alternatively bile may be aspirated with a syringe and then frozen at -20°C. The gonad should be dissected out and weighed. The stage of maturation of the gonad, gonad somatic index (GSI) and liver somatic index (LSI) should be recorded. Otoliths should be removed for ageing the fish.

Figure 1: Sampling procedure for multiple analysis of liver



6. Storage of samples

The practical implication of the step-wise analysis of the PAH impact is that the sampling programme should be carefully planned to provide preserved material for subsequent assay. The occurrence of hepatic lesions should be recorded. Samples should be stored as follows:

- tissues for chemistry and bile should be stored at -20°C;

- tissues for EROD, DNA adduct measurements and histochemistry should be stored at -70°C or in liquid nitrogen until further processed in the laboratory;
- tissue for pathology which has been fixed in 10% neutral buffered formalin should be transferred to 70% alcohol after 24 hours for storage.

7. Analytical procedures

7.1 Cytochrome P4501A activity

Methodology

Following the revision of EROD protocols resulting from experience gained during the North Sea Task Force Monitoring Master Plan programme and various ICES activities (*e.g*. intercalibration exercises, field experiments), it was concluded that the EROD (7-ethoxyresorufin O-deethylase) catalytic enzyme assay is the recommended technique for monitoring CYP 1A activity. The reference methods to be followed are described by Stagg and McIntosh (1997). They include guidelines for sampling, methods for analysis and interpretation.

Quality assurance

Further information on interlaboratory exercises can be found in Munkittrick *et al*., (1993) and Stagg and Addison (1994). Reference material (3 x 1 g) from common pools of minced liver with high (naphthoflavone-induced) and low activities of EROD for selected species (proposed species are dab and flounder) should be distributed to participating laboratories. Material should be frozen in liquid nitrogen in a small volume of distilled water or buffer and shipped to participating laboratories on dry ice. Analyses should be performed within 3–4 weeks. Resorufin standards (from the same batch) should be distributed to participating laboratories. As a control on enzymatic degradation of EROD liver samples CYP 1A content can be measured using immunoassays (Goksøyr, 1992).

7.2 Bulky DNA adducts

Methodology

The P-32 post-labelling technique is the recommended technique for screening for bulky aromatic adducts. The primary reason for using the P-32 technique is its high sensitivity (detection of one adduct in 109-1010 nucleotides), its requirement for small amounts of DNA and its ability to detect carcinogenic DNA adducts of unknown structure. Interlaboratory comparison exercises show that agreements between laboratories are reasonable. Guidance on methodology for use in fish can be found in Stein *et al*., (1993, 1994). The basic protocol entails enzymatic hydrolysis of DNA to nucleosides followed by radioactive labelling and separation of nucleoside adducts by thin layer autoradiography.

Quality assurance

Reference material (3 x 1 g) from common pools of minced liver containing high levels (induced with PAH-contaminated sediment extract) and low levels of DNA adducts for selected species (proposed species are dab and flounder) should be distributed to participating laboratories. Material should be frozen in liquid nitrogen in a small volume of distilled water or buffer and shipped to participating laboratories on dry ice. Analyses should be performed within 4–6 weeks. Standards for measuring efficiencies of DNA hydrolysis and sample blanks ('clean' commercially available DNA) should be distributed to participating laboratories. It is critical that an external standard such as BAPDE-dG (available from NCI) is used.

7.3 PAH metabolites in bile

Methodology

The recommended methodology is the coupled HPLC/fluorometric technique, which uses a combination of excitation/emission wavelengths to detect 2,3 and 4-5 ring compounds. Guidance

on the methodology can be found in Krahn *et al*., (1984). Alternatively, bile samples for screening can be analysed directly or after hydrolysis as described in Ariese *et al*., (1993) and Beyer *et al*., (1996).

Quality assurance

Reference material (3 x 1 ml) of common pools of bile containing high levels (induced with PAH-contaminated sediment extract) and low levels of PAH metabolites for selected species (proposed species are dab and flounder) should be distributed to participating laboratories. Material should be frozen in liquid nitrogen and shipped to participating laboratories on dry ice. Analyses should be performed within 3–4 weeks. Standards (mixture of naphthalene and phenanthrene) should also be distributed to participating laboratories as internal standards.

7.4 Liver histopathology

Methodology

Sampling procedures and the inspection for gross liver lesions should follow the guidelines set out by Bucke *et al*. (1996). Specific techniques to be applied for liver histopathology are given in Feist *et al*. (2004).

All information should be recorded individually per specimen, according to the following procedure:

- collect and maintain live fish prior to individual sampling;
- select a minimum of 30-50 (or the number required to fulfil statistical requirements3) fish per station from the mid-length sizes, *e.g*. for dab, 20-24 cm, for flounder, 25-34 cm;
- sacrifice fish individually by severing spinal cord just posterior to brain;
- carefully dissect out liver avoiding damage to gall bladder and record weight. Examine and record macroscopic lesions;
- cut a 3 mm slice longitudinally from centre axis of the liver using a sharp blade (*e.g*. No. 24). For enzyme histochemistry (optional), take pieces of 5 mm^3 (see Figure 1) and place on a cold, coded chuck at refrigerated temperature. The tissue and chucks are then quenched (supercooled) in n-hexane to -70°C and stored as described in Köhler *et al*., (1992);
- record macroscopic lesions in section and place into wide mouthed container or histiology cassette with 20x fixative per sample. Leave for 24 hours at ambient temperature, then transfer to 70% alcohol for storage (optional);
- remove otoliths and store in dry container;
- process samples for light microscopy to paraffin wax blocks, section blocks at 5μm, stain with haemotoxylin and eosin (H&E);
- archive block and stained section.
- diagnosis of any lesion detected should be according to 5 histopathology categories defined as result of the BEQUALM programme listed in Table 1. Within each category, a range of specific lesions is listed the occurrence of which should be recorded.

The standard reagents required are as follows:

- fixative: 10% neutral buffered formalin (or other fixatives that will allow immuno-histochemical analysis);
- 70% alcohol (industrial methylated spirit (IMS));
- reagents for haemotoxylin and eosin staining (Bancroft and Cook, 1994).

3 Example:
for 95% confidence of detection of a 2% disease prevalence in a population: 150 specimens
for 95% confidence of detection of a 5% disease prevalence in a population: 60 specimens
for 95% confidence of detection of a 10% disease prevalence in a population: 30 specimens.

Quality assurance

Quality assurance takes place through the Biological Effects Quality Assurance in Monitoring Programmes (BEQUALM) scheme (www.bequalm.org).

7.5. Macroscopic liver neoplasms

Methodology

Sampling procedures and the inspection for macroscopic liver lesions should follow the guidelines set out by Bucke *et al*. (1996). Specific techniques to be applied for histopathological confirmation are given in Feist et al. (2004).

All information should be recorded individually per specimen, according to the following procedure:

- collect and maintain live fish prior to individual sampling;
- select a minimum of 50 (or the number required to fulfil statistical requirements4) fish per station from the large sizes, *e.g*. for dab, ≥ 25 cm total length, for flounder, ≥ 30 cm;
- if no sufficient numbers of fish of these size groups are available, fill up the sample with fish of the mid-size group (for dab, 20-24 cm total length, for flounder 25-29 cm total length);
- sacrifice fish individually by severing spinal cord just posterior to brain;
- carefully dissect out liver avoiding damage to gall bladder and record liver weight (optional). Examine and record number and size of macroscopic liver nodules > 2 mm in diameter;
- cut out the complete nodule/s or a parts of it (2-3 mm slice) if too big using a sharp blade (*e.g*. No. 24);
- place tissue sample into a wide mouthed container or a histology cassette with 20x fixative per sample, leave for 24 hours at ambient temperature, then transfer to 70% alcohol for storage (optional);
- remove otoliths and store in dry container;
- process liver tissue samples for light microscopy to paraffin wax blocks, section blocks at 5μm, stain with haemotoxylin and eosin (H&E);
- archive block and stained section;
- diagnosis of the lesion (macroscopic liver nodule) detected should be done according to the histopathology categories defined as result of the BEQUALM programme listed in Table 1. Within each category, a range of specific lesions is listed the occurrence of which should be recorded;
- only report histologically confirmed cases of liver neoplasms (including benign or malignant tumours, but no pre-neoplastic lesions, *i.e.* foci of cellular alteration)

The standard reagents required are as follows:

- fixative: 10% neutral buffered formalin (or other fixatives that will allow immuno-histochemical analysis);
- 70% alcohol (industrial methylated spirit (IMS));
- reagents for haemotoxylin and eosin staining (Bancroft and Cook, 1994).

Quality assurance

Quality Assurance takes place through the Biological Effects Quality Assurance in Monitoring Programmes (BEQUALM) scheme (www.bequalm.org).

4 Example:
for 95% confidence of detection of a 2% disease prevalence in a population: 150 specimens
for 95% confidence of detection of a 5% disease prevalence in a population: 60 specimens
for 95% confidence of detection of a 10% disease prevalence in a population: 30 specimens.

Table 1. BEQUALM categories of histopathological liver lesions in fish that should be used for 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

8. Reporting requirements

Data reporting should be in accordance with the requirements for National Comments and the latest ICES Environmental Data Reporting Formats. The following data are required:

8.1 Contaminants

- the monitoring of PAH in sediments to be consistent with the JAMP Guidelines for Monitoring Contaminants in Sediment.

8.2 Biological effects measurements

CYP1A

- 7-ethoxyresorufin *O*-deethylase activity as ηmol or ρmol of resorufin per minute per mg protein or as ηmol or ρmol of resorufin per minute per mg tissue;
- The molar absorbance of the resorufin used;
- The protein assay and standard (*e.g*. IgG, BSA) used;
- The assay conditions used (*e.g*. temperature, 7-ethoxyresorufin concentration, NADPH concentration, protein concentration in the assay).

DNA adducts

- Adducts should be expressed as the ratio of the number of adducts detected divided by the amount of DNA used in the assay (*e.g*. as ρmol DNA adducts/mol DNA or fmol DNA adducts/μg DNA);
- The level of adducts in DNA such as is obtained from testes of farmed salmon should be included as a blank;
- The efficiency of enzyme mediated transfer of ^{32}P from ATP to the DNA adducts using 7R,8S,9S-trihydroxy,10R-(N^2-deoxyguanosyl-3'-phosphate)-7,8,9,10-tetra-hydrobenzo[*a*]pyrene(BaPDGE-dG-3'p) as an external standard.

Bile metabolites

- Source and concentration of ß-glucuronidase;
- The concentration of metabolites (ηg g^{-1} wet weight) as:
  - naphthalene at the following wavelengths Ex 290 Em 335;
  - phenanthrene at the following wavelengths Ex 256 Em 380;
  - benzo[*a*]pyrene at the following wavelengths Ex 380 Em 430.
- The concentration of protein and biliverdin in the bile.

Liver histopathology

As a minimum requirement, the occurrence of the following histopathology categories should be reported:

- Non-specific lesions;
- Early toxicopathic non-neoplasticlesions;
- Foci of cellular alteration;
- Benign neoplasms;
- Malignant neoplasms;

However, it is recommended that information on the occurrence of the specific histopathological lesions assigned to the 5 histopathology categories is reported as well.

Macroscopic liver neoplasms

- The occurrence of histologically confirmed cases of macroscopic liver neoplasms (includes benign and malignant neoplasms, but no pre-neoplastic stages, i.e. foci of cellular alteration) should be reported.

8.3 Supporting parameters

- fish identifier;
- site code and name;
- taxonomic identification;
- temperature and salinity at collection site;
- date of sample collection;
- gender;
- location and description of grossly visible anomalies and lesions (externally visible ones and those on the liver or other inner organs);
- liver somatic index;
- gonad somatic index;
- fish weight and length;
- age of the fish.

9. References

Ariese, F., Kok, S.J., Verkaik, M., Gooijer, C., Velthorst, N.H. and Hofstraat, J.W., 1993. Synchronous fluorescence spectometry of fish bile: a rapid screening method for the biomonitoring of PAH exposure. Aquat. Toxicol. 26: 273-286.

Bancroft, J.D. and Cook, H.C., 1994. Manual of Histological Techniques and their Diagnostic Applications. Churchill Livingstone, UK. 457.

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

Bock, K.W., 1993. Aryl hydrocarbon or dioxin receptor: biologic and toxic responses. Rev. Physiol. Biochem. Pharmacol. 125.

Bruno, D.W. and Poppe, T.T., 1996. A Colour Atlas of Salmonid Diseases. Academic Press, London 194.

Bucke, D., Vethaak, D., Lang, T. and Mellergaard, S., 1996. Common diseases and parasites of fish in the North Atlantic: Training guide for identification. ICES No. 19, 27.

Burgeot, T., Bocquene, G., Truquet, P., Dean le L., Poulard, J.C., Dorel, D., Souplet, A. and Galgani, F., 1993. The dragonet *Callionymus lyra*, a target species used for evaluation of biological effects of contaminants on French coasts. Mar. Ecol. Prog. Ser 97:309-316.

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 *TIMES* No. 38, 42 pp.

Fisser, B. and Blomeke, B., 1992. DNA Schäden in Fischen aus der Nordess: 32P postlabelling Analyse von DNA-Addukten in Lebern von Klieschen aus der Deutschen Bucht und von der Doggerbank. Berichte Zentr. Meeres-u. Klimaforsch. Hamburg 24.

Goksøyr, A. and Förlin, L., 1992. The cytochrome P450 system in fish: aquatic toxicology and environmental monitoring. Aquatic Toxicol. 22: 287-312.

Goksøyr, A. and Husøy, A.M., 1992. The cytochrome P450 1A1 response in fish: Application of immunodetection in environmental monitoring and toxicological testing. Mar. Environ. Res. 34:147-150.

Goksøyr, A., Husøy, A.M., Larsen, H.E., Klungsoyr, J., Wilhelmsen, S., Maage, A., Brevik, E.M., Andersson, T., Celander, M., Pesonen, M. and Förlin, L., 1991. Environmental contaminants and biochemical responses in flatfish from the Hvaler archipeligo in Norway. Arch. Environ. Contam. Toxicol. 21: 486-496.

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.

Ioannides, C. and Parke, D.V., 1993. Induction of cytochrome P4501A as an indicator of potential chemical carcinogenesis. Drug Metab. Rev. 25:485-501.

Jones, N.J. and Parry, J.M., 1992. The detection of DNA adducts, DNA base changes and chromosome damage for the assessment of exposure to genotoxic pollutants. Aquat. Toxicol. 22:323-344.

Köhler, A., Deisemann, H. and Lauritzen, B., 1992. Histological and cytochemical indices of toxic injury in the liver of dab *Limanda limanda*. Mar. Ecol. Prog. Ser. 91: 141-153.

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

Krahn, M.M., Myers, M.S., Burrows, D.G. and Malins, D.C., 1984. Determination of metabolites of xenobiotics in the bile of fish from polluted waterways. Xenobiotica 14:633-646.

LaFaurie, M.J., Guidcelli, J., Carriere, S., Lemaire, P., Mathieu, A. and Negre, Y., 1989. Pollutant biotransformation in fish: use in environmental health evaluation. In: Advances in Applied Biotechnology Series Vol 5: WHO Regional Office (Europe) and United Nations.

Lange, U., Saborowski, R., Karbe, L. and Siebers, D., 1995. Towards application of EROD activity in dab (*Limanda limanda*) in biological effect monitoring of the North Sea: a model for prediction of basal EROD activity levels. ICES CM 1995/E:7.

Munkittrick, K.R., van den Heuvel, M.R., Metner, D.A., Lockhart, W.L. and Stegeman, J.J., 1993. Interlaboratory comparison and optimisation of hepatic microsomal ethoxyresorufin-*O*-deethylase activity in white sucker (*Catostomas commersoni*) exposed to bleached kraft pulp mill effluent. Environ. Toxicol. Chem. 12:1273-1283.

Narbonne, J.F., Garrigues, P., Ribera, D., Raoux, C., Mathieu, A., Lemaire, P., Salaun, J. P. and LaFaurie, M.J., 1991. Mixed function oxygenase enzymes as tools for pollution monitoring: field studies on the French coast of the Meditteranean Sea. Comp. Biochem. Physiol. 100:37-42.

Payne, J.F., Fancey, L., Rahimtula, A. and Porter, A.D., 1987. Review and perspective on the use of mixed-function oxygenase enzymes in biological monitoring. Comp. Biochem. Physiol. 86C:233-245.

Payne, J.F., Melvin, W., Mathieu, A. and Fancey, L., 1994. Biomarkers of stress in urban rivers: mixed function oxidase and acetylcholinesterase effects in brown trout in rivers in St John's, Newfoundland. Canad. Tech. Rep. Fish. Aquat. Sci. 1947.

Phillips, D.H., Castegnaro, M. and Bartsch, H., 1993. Postlabelling Methods for the Detection of DNA Adducts. Lyon: IARC Scientific Publications 124.

Renton, K.W. and Addison, R.D., 1992. Hepatic microsomal mono-oxygenase activity and P4501A mRNA in North Sea dab *(Limanda limanda)* from contaminated sites. Mar. Ecol. Progr. Ser. 91:65-69.

Sleiderink, H.M., Oostingh, I., Goksøyr, A. and Boon, J.P., 1995. Sensitivity of cytochrome P4501A induction in dab (*Limanda limanda*) of different age and sex as a biomarker for environmental contaminants in the southern North Sea. Arch. Environ. Contam. Toxicol. 28:423-430.

Stagg, R.M. and Addison, R.A., 1994. An inter-laboratory comparison of measurements of ethoxyresorufin-*O*-deethylase activity in dab (*Limanda limanda*) liver. Mar. Environ. Res. 38:93-108.

Stagg, R.M., McIntosh, A. and Mackie, P., 1995. The induction of hepatic mono-oxygenase activity in dab (*Limanda limanda* L) in relation to environmental contamination with petroleum hydrocarbons in the northern North Sea. Aquat. Toxicol. 33:245-264.

Stagg, R.M. and McIntosh, A.M., 1997. Determination of CYP1A-dependant mono-oxygenase activity in the liver of the dab (*Limanda limanda*) by the fluorimetric measurement of 7-ethoxyresorufin-*O*-deethylase (EROD) activity. ICES Techniques in Marine Environmental Sciences No. 22.

Stegeman, J.J., Brouwer, M., DiGiulio, R.T., Forlin, L., Fowler, B.A., Sanders, B.M. and Van Held, P.A., 1992. Molecular responses to environmental contamination: enzyme and protein systems as indicators of chemical exposure and effect. In: Biomarkers: Biochemical, Physiological and Histological Markers of Anthropogenic Stress. Huggett, R.J., Kimmerle, R.A., Merhle, P.M. Jnr. and Bergman, H.L. (Editors). Michigan, USA: Lewis Publishers.

Stein, J.E., Reichert, W.L., French, B. and Varanasi, U., 1993. 32P-postlabelling analysis of DNA adduct formulation and persistence in English sole (*Pleuronectes vetulus*) exposed to benzo[*a*]pyrene and 7H-dibenzo[*c,g*]carbazole. Chem. Biol. Interact. 88:55-69.

Stein, J.E., Reichert, W.L. and Varanasi, U., 1994. Molecular epizootiology: assessment of exposure to genotoxic compounds and teleosts. Environ. Health Perspect. 102:19-23.

Technical Annex 3

TBT-specific biological effects monitoring

1. Sampling strategy

To investigate the intensity of TBT effects and their geographical scale *(i.e.* the spread of effects away from point sources), the imposex (superimposition of penis and/or vas deferens on prosobranch females) and intersex condition (pathological alterations in the oviduct of littorinids and replacement of female by male organs) have been proved sensitive biomarkers for the determination of the degree of environmental organotin and especially tributyltin (TBT) pollution in coastal waters.

Imposex in the common dogwhelk *Nucella lapillus* has been used widely as an indicator of TBT contamination and is an extremely sensitive indicator. Since *Nucella lapillus* is not present in the Baltic Sea or in many areas of the southern North Sea (e.g. the coastal waters of Belgium, Germany, Portugal and with some exceptions, the Netherlands, Denmark and Sweden) at least one other widespread and highly abundant effect monitoring species is required for the Atlantic and North Sea coastal waters and the Baltic Sea, respectively. In parts of the Baltic Sea, and in some areas beyond the northern limit of *Nucella* (approximately the –1ºC winter isotherm), *Littorina* may be available as a shallow sub-tidal species. The southern limits of the ranges of *Nucella* and *Littorina* lie in the Iberian peninsular. In these and other southern areas, suitable alternative coastal monitoring organism may be *Nassarius reticulatus* (imposex).

In offshore areas the imposex response in *Buccinum undatum* can be used for the assessment of organotin contamination. In offshore areas, where *Neptunea antiqua* is more abundant than *Buccinum,* imposex response in *Neptunea* is a useful alternative. In addition there are indications that *Neptunea* is more sensitive to TBT than *Buccinum*, as imposex occurrence is more frequent in *Neptunea* in areas where the two species co-exist. Other species may be necessary for surveys in other European coastal seas, for example the Mediterranean Sea.

High degrees of development of both imposex and intersex in *Nucella* and *Littorina* have significance beyond the individual organism level, in that they have been associated with changes in the reproductive capacity of populations through lack of juveniles, poor recruitment, local extinction etc. The occurrence of imposex in *Buccinum, Neptunea* and *Nassarius* has also been linked with decreases in populations but whether this is a direct consequence of imposex is unclear. The possibility of including estimates of population structure in a monitoring programme is excluded because of difficulty in attributing cause and effect due to a range of confounding factors, such as seasonality in behaviour and differences in the character of the substrate, etc.

The occurrence and intensity of imposex and intersex in *Nucella*, *Nassarius* and/or *Littorina* are the recommended biological effects to be monitored around major marinas/shipyards/offshore installations/harbours. The occurrence and intensity of imposex in *Buccinum* and/or *Neptunea* is the recommended biological effect to be monitored in areas with high shipping traffic, such as major shipping routes.

Broad-scale surveys describing the intensity of TBT impacts in coastal waters would allow comparisons to be made between various stretches of coast and assessments of the potential for TBT-affected species to recover.

2. Sampling locations

2.1 Monitoring around point sources

Samples should be collected in coastal areas on both sides of the facility *(e.g.* marina, shipyard) in order to describe the gradients of contamination. A series of, preferably, at least six sample sites should be located on each side of the facility, covering the gradients between the highest level of effect to the local background level of effect. The sample sites should be clearly identified in order to enable repeat sampling, for example to investigate temporal trends in contamination and effects.

Coastal areas can be divided into a number of boxes, possibly following the divisions used in preparing sub-regional reports for the 1993 North Sea Quality Status Report. In areas where sub-regions have not been defined, guidance on the scale of the appropriate division may be obtained from comparison with areas where sub-regions have been used. Within each box, a number of stations should be selected along the gradient.

The sampling undertaken around marinas/shipyards/offshore installations/harbours should be such that interpolated values for the Vas Deferens Sequence Index (VDSI) are within 0.5 units of the true value with -

90% confidence in the range VDSI 2-6 for *Nucella* and 1-4 for *Nassarius*. Interpolated values of Intersex Index (ISI) should be within 0.5 units of the true value with 90% confidence.

The expressions of degree of effect at each location in a box should be combined to give indices of centrality and spread of their distribution. The values derived for each box should be compared using standard statistical techniques (e.g. t-tests or other appropriate methods) to give an expression of the relative degree of contamination of different coastal areas. The measures of effects should be used to express the general degree of inhibition of reproductive capacity of the monitored species in each area and hence the potential for these populations to act as sources of young for recolonisation of heavily affected areas.

2.2 Monitoring shipping lanes

Shipping lanes selected for study should be of defined location and width, e.g. as a result of marine traffic management/separation schemes. They should carry a significant proportion of the commercial vessel traffic in the area, have a high traffic density and preferably be some distance from other sources of TBT. The substrate should be suitable for *Buccinum undatum.* Employing these criteria, the following locations are suggested:

a. Strait of Dover;
b. German Bight - Texel T.S.S;
c. Off Ushant Island (north west France);
d. Tusker/Smalls (Irish/Celtic seas);
e. North Channel (low traffic intensity);
f. Off Berlenga/ Cabo da Roca (west Portugal);
g. Pentland Firth;
h. Skagerrak;
i. Southwest Iceland.

In each area, *Buccinum* should be collected from a minimum of three stations situated close to the shipping lane being studied. For comparison, *Buccinum* should also be collected from at least one area with low shipping traffic. This area should be far enough away from the shipping lane for the level of effect to be similar to the prevailing local background level of effect. The same strategy can be used for *Neptunea* in areas where this species is more abundant than *Buccinum*.

2.3 Regional TBT survey

The common dogwhelk *Nucella lapillus* remains absent in some coastal areas, such as the southern coast of the UK, where formerly it has been present. The whelk *Buccinum undatum* has disappeared from the Friesian area of the Wadden Sea and is now only found in the far north of the German coast, where populations are all sub-tidal. Interest in neogastropod populations away from TBT point sources is therefore centred on:

a. comparing levels of TBT contamination in different areas;
b. assessing the ability of neogastropods to survive and recolonise;
c. placing the results from point source surveys in a wider context.

As variation in TBT-induced effects exist at scales of kilometres and mapping the whole of the maritime area is not a feasible objective, broad spatial surveys could be undertaken on the basis of estimates of the degree of impact in fairly large areas. The biological effects recommended as the core of the survey are as for the monitoring of point sources - *i.e.* imposex in *Nucella lapillus* and *Nassarius reticulatus* and intersex in *Littorina littorea,* where either one, both or all three species are present. This should ensure adequate coverage of the coasts of the maritime area. It is considered premature to recommend that a similar approach should be adopted for *Buccinum* and *Neptunea* in offshore areas. This recommendation should be reviewed when the results of the investigations into effects on transects across shipping lanes are known.

There are a very large number of TBT point sources in the maritime area. The impact of individual sources may be detectable up to a few kilometres from each source. The effects of these point sources are superimposed upon a general (variable) background level of contamination in the surrounding area. The numbers of samples necessary to map the distribution of TBT effects throughout the whole maritime area are clearly very large and such an approach cannot be considered. In order to provide information on the general level of effects in large coastal regions of the maritime area it is recommended that the coastal areas are divided into a number of boxes, possibly following the divisions used in preparing sub-regional reports for the 1993 North Sea Quality Status Report. In areas where sub-regions have not been defined, guidance on the scale of the appropriate division may be obtained from comparison with areas where sub-regions have been

used. Within each box, a number of stations should be selected randomly, bearing in mind the need to avoid distortions that might be introduced by stations very close to recognised point sources of TBT.

Nucella and/or *Littorina,* as appropriate, should be sampled at each location, the numbers of samples in each box should be chosen to allow the detection of differences of one unit in RPSI or ISI between adjacent boxes using data from one sampling occasion. Contracting Parties may wish to sample more locations than are suggested by the above criteria.

3. Methods to be used

The biological effects measurements provide direct indications of the effects of TBT at the individual organism level. From these data, it is possible to make inferences regarding effects at the population level. The most important measurements of imposex/intersex are the VDSI and ISI. ISI provides a direct indication of the average reproductive capacity of the females in *Littorina* populations. VDSI is an index of imposex based on the development of both penis and vas deferens. High values of the VDSI indicate reduced breeding capability of females in populations. Observations of VDSI and ISI should be supported by calculations of the proportions of females in the populations which have been rendered sterile by exposure to TBT *(i.e.* proportion showing stage 5 or 6 imposex (only *Nucella lapillus*), or intersex stage 2 or above). The Relative Penis Size Index (RPSI) provides additional discrimination in *Nucella* populations at VDSI values of 4, the Relative Penis Length Index (RPLI) in *Nassarius* at VDSI 4 and the Average Prostate Length of Females (FPrL) similarly assists discrimination at intersex values of 3 in *Littorina littorea*. The concentrations of TBT in soft tissue allow discrimination between effects arising from current and historical contamination.

Details of the methodology for the biological effects measurements are as at Section 7. Briefly imposex and intersex assessment in *Nucella*, *Nassarius* and *Littorina* should be carried out according to the methods described in Gibbs *et al.,* (1991) Fioroni *et al.,* (1991), Stroben *et al.,* (1992a, b) and Bauer *et al.,* (1995), imposex in *Buccinum* by an adaptation of Ten Hallers-Tjabbes et *al.,* (1994) and Mensink et *al.,* (1997), and imposex in *Neptunea* according to Power & Keegan (2001) and Strand & Jacobsen (2002). The guidelines at Section 7 should be followed as they provide standardised methods unifying the differences adopted by separate scientists.

A variety of supporting data should be included to assist in the interpretation of the effects data. At all stations, the presence or absence of juvenile *Nucella*, *Nassarius* or *Littorina,* and *Nucella* and *Nassarius* egg capsules should be noted. At key stations, the size frequency distribution of the population should be determined. The concentration of TBT should be measured in pooled homogenised whole soft parts of female specimens. It has been found that while imposex and intersex are essentially irreversible, TBT residues are lost from the soft parts with half-times of around 60-80 days. The comparison of effect measurements with TBT concentrations in tissues can therefore provide indications of whether the effects represent a response to recent exposure or are a record of more historical exposure. Data should be collected on the use of the facilities being monitored. In the case of harbours, marinas, *etc.* this information will have been collected during the site selection procedure and this information should be updated as the monitoring programme progresses.

4. Temporal trend monitoring

Temporal trend studies are valuable in assessing the response of the environment to control measures, to the effects of changes in the pattern of TBT use and to possible changes in the pattern of use of the marine environment. Temporal trend studies should build upon and use similar methodology to, the spatial studies described in Section 2. The primary measurements should be of imposex and intersex in *Nucella*, *Nassarius* and *Littorina* supported by chemical analysis and population size frequency distribution. There is currently little experience of the use of *Buccinum* or *Neptunea* in temporal trend programmes, although these species may also prove to be useful. *Nucella*, *Nassarius* and *Littorina* are all quite long-lived species and it is likely that many individuals in the wild live for around five years. Both imposex and intersex are, in general, not reversible and therefore the replacement of the population on a five year scale places a clear lower limit to the duration of any temporal trend programme. Programmes of 10 years or more must be the minimum practicable length, even if attempts are made to sample a defined age class or group of age classes in the populations. A further consequence of the longevity of the animals is that annual sampling would be unnecessarily frequent. On the basis of reports of studies in Sullom Voe and elsewhere, sampling at two-year intervals was considered appropriate for temporal trend monitoring, and that a case could be made for less frequent sampling (Harding *et al.,* 1997). Single surveys of the spatial distribution of TBT effects can provide useful information on the geographical distribution of TBT contamination.

Temporal trend programmes around point sources can be assessed through the comparison of the gradients of effect detected on each survey and on the degrees of effect detected at individual sampling points. Guidance on how the gradients might be compared and assessed is not available. The data at individual points should be assessed using techniques similar to those previously established in ICES for assessing temporal trends in contaminant concentrations (e.g. Kelly *et al.,* 1994, Nicholson *et al.,* 1991).

Temporal trends in the degree of effect on coastal boxes will yield indices of centrality and spread and should be able to be assessed using procedures already available within ICES for contaminant programmes. It is considered that a significant change in VDSI would be about 1 unit and 0.5 unit in ISI. The ability of a temporal trend programme to detect changes of this scale with good (90%) confidence over a defined time period is dependent upon the ratio of the size of the change.

The sampling and analytical guidelines discussed above in relation to point sources *(cf.* Section 2.1) should be followed for temporal trends. It is suggested that the maximum frequency of sampling should be every two years. The location of sampling sites should follow the locations established in relation to point sources. As subsequent sampling exercises should use the same locations and animals at the same stage of their reproductive cycle, the initial selection of sample sites should be undertaken with care to ensure that the snail population will not be unduly affected by sampling pressure.

5. Field sampling and sampling equipment

5.1 Intertidal species

Nucella lapillus, *Nassarius reticulatus* and *Littorina littorea* should be sampled during low tide, by hand, from the substrate between spring low water and mid-tide levels with sample sizes of 40 specimens per station. Alternatively, for *Nassarius* sampling in traps baited with dead flesh or with a dredge can be considered in the shallow sub-tidal range. Especially in contaminated areas the abundance of dogwhelks is likely to decrease because of female sterilisation. In these areas it is often not possible to achieve the recommended number of 40 *Nucella lapillus* and therefore it is also imperative to analyse *Littorina littorea* or *Nassarius reticulatus.*

An attempt should be made wherever possible to exclude the largest specimens with eroded shells which are often covered by barnacles and/or sedentary polychaetes (e.g. *Spirorbis* sp.) since older specimens may exhibit an imposex or intersex intensity which reflects a former and possibly higher contamination level. The most suited age classes for the analyses are littorinids which have just become sexually mature (1,5 years or more). Normally, these specimens cover shell heights between 15 and 25 mm in *Littorina littorea* depending on the station sampled. In order to obtain an intercalibration of imposex in *Nucella lapillus* or *Nassarius reticulatus* and intersex in *Littorina littorea* on a broader database (Minchin *et al.,* 1996a,b) analyses should be performed at a number of stations where both species occur sympatrically.

The presence of juveniles or egg capsules in the field can be a valuable indication of continued reproductive activity in populations.

Penis lengths in the target gastropods are known to vary seasonally. Samples should be collected within 2-3 months at the end of the reproduction season to reduce this source of variability.

5.2 Offshore species

Wherever possible, samples of 100 mature (usually >5-6 *cm) Buccinum undatum* should be obtained by suitable methods. It is recommended that where possible samples should be collected using static gear (traps, pots, *etc.).* In areas where this is not practicable (e.g. within the shipping lanes) mobile gear (otter trawl, beam trawl *etc.)* may be towed parallel to shipping lanes, but with added uncertainty as to sampling position. Samples of 40 *Neptunea antiqua* should be collected with the same methods as used for *Buccinum.*

6. Storage of samples

In the laboratory, samples for biological measurements should be analysed as soon as possible according to the following protocol but not later than seven days after sampling. All samples containing live specimens should be transported to the laboratory in damp conditions in non-contaminating containers. If necessary, cooling facilities during transport and storage should be employed.

7. Determination of imposex and intersex

7.1 Methodology

Treatment of samples and sample analyses

Nucella, *Buccinum* and *Neptunea* should preferably be examined live, without narcotisation, to retain consistency with many previous studies. It is also likely that narcotisation will particularly affect penis length. *Littorina* and *Nassarius* should be narcotised in 7% $MgCl_2$ in distilled water (for Atlantic and North Sea samples; probably lower concentrations according to salinity at sampling sites) to achieve a maximum

relaxation, to simplify the measurement procedure and to provide more reproducible results. Before breaking the shell with a vice, individual measurements for shell heights should be measured to the nearest 0.1 mm with a vernier calliper. Specimens parasitised by trematodes or other endoparasites should be excluded from further analyses.

If, for example, for logistical reasons it is necessary to preserve samples prior to analysis, specimens may be frozen. However, freezing will affect penis length measurements and data should be corrected to make them comparable to data from live specimens. The identification of imposex and intersex stages should be unaffected by freezing.

Determination of imposex

Under the microscope in imposex-affected species the following observations and measurements (with an accuracy for organ extensions of 0.1 mm) should be undertaken:

a. the imposex stage should be determined for each female according to Figures 1a and 1b for *Nucella lapillus*, Figure 2 for *Nassarius reticulatus* and Figure 2 or 3 for *Buccinum undatum* and *Neptunea antiqua*. Generally, in *Nucella lapillus* only the imposex stages summarised in Figure 1a can be found (Gibbs *et al.,* 1987) but due to genetical variation the additional types 1-3a and b can occur in populations along the Channel coast of France and England (Oehlmann *et al.,* 1991);

b. the percentage of female *Nucella* in stages 5 and 6 (indicates proportion of females which are sterile);

c. the length of penis in *Nucella* and *Nassarius* males and females (if present). Measurement of penis length in *Nucella* and *Nassarius* should be done immediately after removing it from the shell, as the penis enlarges after its death.

On the basis of these measurements the following calculations of imposex intensities and indices of possible interference with the reproductive capability in prosobranch populations should be calculated:

a. Vas Deferens Sequence Index (VDSI) in *Nucella*, *Nassarius* and *Neptunea* as the mean value of the imposex
stages in a population:
VDSI = (Sum of imposex stage values of all females sampled) / (Number of females)
VDSI values > 4,0 indicate that at least some of the females in the sample are sterilised due to imposex development;

b. the proportion of sterile female *Nucella;*

c. for *Nucella lapillus* the Relative Penis Size Index (RPSI) should be calculated:
RPSI = ((Average length of female penis)^3/(Average length of male penis) ^{3}) x 100;

d. for *Nassarius reticulatus* the Relative Penis Length Index (RPLI) should be calculated:
RPLI = (Average length female penis) / (Average length of male penis) x 100;

e. for *Buccinum undatum* and *Neptunea antiqua* the proportion of females showing imposex (i.e. other than at stage 0) should be calculated. For *Buccinum undatum*, the Penis Classification Index (PCI) should be calculated following the descriptive scheme in Figure 3.
PCI = (Sum of imposex stage values of all females examined) / (number of females)
It may also be possible to classify imposex in this species according to Figure 2a if comparison of the developmental sequence in *Buccinum* with those in other species is desired.

Determination of intersex

Under the dissecting microscope the following measurements should be undertaken for *Littorina littorea* (with an accuracy of 0.1 mm for organ extensions), in females only:

a. intersex stage (according to Figures 4 and 5);

b. length of prostate gland (if present). Prostate tissue is normally easily identifiable in the pallial gonoduct section of females.

On the basis of these measurements the following calculations should be performed in order to determine intersex intensities and to estimate the possible interference with the reproductive capability in *Littorina* populations:

a. the Intersex Index (ISI) as the mean value of the intersex stages in a population:
ISI = (Sum of intersex stage values of all females sampled) / (Number of females);
ISI values > 1,0 indicate that at least some of the females in the sample are sterilised due to intersex development;

b. average Female Prostate Length (FPrL):
FPrL = (Sum of prostate lengths of all females sampled) / (Number of females).

7.2 Quality assurance

All determinations should be carried out by trained staff working to defined protocols. Any deviations from the protocols should be recorded and assessed by the laboratory manager for their potential to influence the results.

Laboratory standards for effects measurements should be prepared as preserved material and used as reference standards by analytical laboratories during their routine work. Control charts can be prepared from repeated measurements of penis length on single preserved specimens. The repeated measurement of penis length or prostate length in preserved specimens could be used to control biological measurements. This will control for differences in interpretation between analysts and for gross errors in microscope calibration.

International Laboratory Performance Studies of imposex and intersex measurements are available through QUASIMEME and provide a formal framework for external quality assurance.

8. Determination of organotins

8.1 Methodology

There are a wide variety of methods available for the determination of organotins in shellfish tissues[5]. It is not appropriate to specify one or more methods that should be adopted by participating laboratories. However, some guidance on performance characteristics can be given as follows:

a. the method should be designed to measure TBT (rather than, for example, total tin or total organotin);

b. the concentration range of interest is around 0.005-1.6 mg/kg dry weight or 0.002-0.6 mg/kg wet weight;

c. the detection limit should be 0.002 mg/kg dry weight (0.001mg/kg wet weight) or better;

d. reference materials should be used. Certified Reference Materials are now available for TBT in marine matrices: CRM 477 (mussel tissue), PACS-2 and BCR 646 (sediments) and a fish tissue from NIES Japans[6];

8.2 Quality assurance

TBT compounds can be purchased for the preparation of standard solutions. Reagents for the chemical analysis should be of appropriate quality to meet the needs of the determination. At present, there are very few suitable Certified Reference Materials for TBT in biota available. It should be possible to establish standard chemical Shewart control charts for the chemical analysis of biota in support of the biological effects measurements, either from Certified Reference Materials or from Laboratory Reference Materials.

All determinations should be carried out by trained staff working to defined protocols. Any deviations from the protocols should be recorded and assessed by the laboratory manager for their potential to influence the results. Normal care should be taken during chemical analysis to minimise contamination or loss of analytes and interference from other substances and to ensure accurate calibration of instruments. The necessary performance characteristics of the chemical methods are given in the Guidelines for monitoring contaminants in biota.

9. Reporting requirements

Data reporting should be in accordance with the requirements for national comments and with the latest ICES reporting formats. The following data are required:

9.1 Contaminants (Optional)

- TBT, DBT and MBT concentration in tissues (mg/kg);
- TPhT, DPhT, MPhT concentration in tissues (mg/kg) when relevant;
- wet weight or dry weight basis;

5 It is recommended that an appropriate coordinating body be commissioned to develop a complete QC/QA system for the measurement o f TBT and lipid in gastropod tissue.

6 It is thought that the EU SMT programme is preparing further materials based on cod muscle and mussel tissue.

9.2 Biological effects measurements

Imposex
- proportion of females displaying imposex (all species);
- vas deferens sequence index (all species);
- relative penis size *(Nucella)*
- average penis length of females (*Nassarius*)
- proportion of sterile females in stages 5 and 6 (*Nucella*);
- penis classification index (*Buccinum*);

Intersex
- proportion of females displaying intersex;
- intersex index;
- average length of prostate gland in females
- proportion of sterile females in stage 2 and 3.

	Nucella	*Buccinum*	*Neptunea*	*Nassarius*	*Littorina*
Proportion of females displaying imposex or intersex	1	1	1	1	1
Vas Deferens Sequence Index (VDSI)	1	1	1	1	
Proportion of sterile females	1				1
Relative Penis Size Index (RPSI)	1			1	
Relative Penis Length Index (RPLI)				1	
Intersex Index (ISI)					1
Average length of prostate gland in females (FPrL)					1
Penis Classification Index (PCI)		1			

9.3 Supporting parameters

- site code;
- taxonomic identification;
- date of sample collection;
- number of individuals in sample
- presence/absence of juveniles and/or egg capsules*;*
- population size & frequency distribution, if considered useful supplementary information.

10. References

Barreiro, R., Gonzalez, R., Quintela, M., and Ruiz, J.M., 2001. Imposex, organotin bioaccumulation and sterility of female Nassarius reticulatus in polluted areas of NW Spain. Marine Ecology Progress Series, 218, 203-212.

Barroso, C.M., Moreira, M.H., and Bebianno, M.J., 2002. Imposex, female sterility and organotin contamination of the prosobranch *Nassarius reticulatus* from the Portuguese coast. Marine Ecology Progress Series, 230, 127-135.

Bauer, B., Fiorini, P., Ide, I., Liebe, S., Oehlmann, J., Stroben, E. and Watermann, B., 1995. TBT effects on the female genital system of *Littorina littorea,* possible indicator of tributyltin pollution. Hydrobiologia 309, 15-27.

Fiorini, P., Oehlmann, J. and Stroben E., 1991. The pseudohermaphroditism of prosobranchs: morphological aspects. Zool. Anz. 226:1-26.

Gibbs, P.E., Bryan, G.W., Pascoe, P.L. and Burt, G.R., 1987. The use of the dogwhelk *Nucella lapillus* as an indicator of tributyltin (TBT) contamination. J. Mar. Biol. Ass. UK 67:507-523.

Gibbs, P.E., Pascoe, P.L. and Bryan, G.W., 1991. Tributyltin-induced imposex in stenoglossan gastropods: pathological effects on the female reproductive system. Comp. Biochem. Physiol. 100C:23 1-235.

Harding, M.J.C., Rodger, G.K., Davies, I.M. and Moore, J.J., 1997. Partial recovery of the dogwhelk *(Nucella lapillus)* in Sullom Voe Shetland from tributyltin contamination. Mar. Environ. Res. 44(3): 285-304.

Huet, M; Fioroni, P; Oehlmann, J; Stroben, E, 1995. Comparison of imposex response in three prosobranch species. Advances in Littorinid biology. Proceedings of the 4th international symposium on Littorinid biology, Roscoff, 19-25 September 1993. Hydrobiologia, 309, 29-35.

Kelly, A.G., Wells, D.E. and Fryer, R.J., 1994. Sampling strategy to detect changes in concentration of trace organic contaminants in marine sediment. Sci. Tot. Environ. 144:2 17-230.

Mensink, B.P., Van Hattum, B., Ten Hallers-Tjabbes, C.C., Everaarts, J.M., Kralt, H., Vethaak, A.D. and Boon, J.P., 1997. Tributyltin causes imposex in the common *whelk Buccinum undatum* - mechanism and occurrence. Nederlands Instituut voor Onderzoek der Zee, Rapport 1997-6.

Minchin, D., Oehlmann, J., Duggan, C.B., Stroben, E. and Keatinge, M., 1996a. Marine TBT antifouling contamination in Ireland following legislation in 1987. Mar. Pollut. Bull. 30 (10): 633-639.

Minchin, D., Stroben, E., Oehlmann, J., Bauer, B., Duggan, C.B. and Keatinge, M., 1996b. Biological indicators used to map organotin contamination in Cork Harbour, Ireland. Mar. Pollut. Bull. 32 (2): 188-195.

Nicholson, M.D., Green, N.W. and Wilson, S.J., 1991. Regression models for assessing trends in cadmium and PCBs in cod livers from Oslofjord. Mar. Pollut. Bull. 22:77-81.

Oehlmann, J., Stroben, E. and Fiorini, P., 1991. The morphological expression of imposex in *Nucella lapillus* (LINNAEUS) (Gastropoda: Muricidae). J. Moll. Stud. 57: 375-390.

Poloczanska, E.S and Ansell A.D., 1999. Imposex in the whelks *Buccinum undatum* and *Neptunea antiqua* from the west coast of Scotland. Mar. Environ. Res. 47:203-212

Power, A.J. & B. F. Keegan (2001). The significance of imposex levels and TBT contamination in the red whelk, *Neptunea antiqua* (L.) from the offshore Irish Sea. Marine Pollution Bulletin 42(9): 761-772.

Strand, J. & J.A. Jacobsen (2002). Imposex in two sublittoral neogastropods from the Kattegat and Skagerrak: the common whelk *Buccinum undatum* and the red whelk *Neptunea antiqua*. Marine Ecology Progess Series 244: 171-177.

Stroben, E., 1991. Imposex und weitere Effekte chronischer TBT-Intoxikation bei einigen Mesogastropoden und Bucciniden (Gastropoda, Prosobranchia). Göttingen: Cuvillier.

Stroben, E., Oehlmann, J. and Fioroni, P., 1992a. The morphological expression of imposex in *Hinia reticulata* (Gastropoda: Buccindae): a potential indicator of tributyltin pollution. Mar. Biol. 113: 625-636.

Stroben, E., Oehlmann, J. and Fioroni, P., 1992b. *Hinia reticulata* and *Nucella lapillus*. Comparison of two gastropod tributyltin bioindicators. Mar. Biol. 114: 289-296.

Ten Hallers-Tjabbes, C.C., Kemp, J.F. and Boon, J.P., 1994. Imposex in whelks *(Buccinum undatum)* from the open North Sea: Relation to shipping traffic intensities. Mar. Pollut. Bull. 28:311-313.

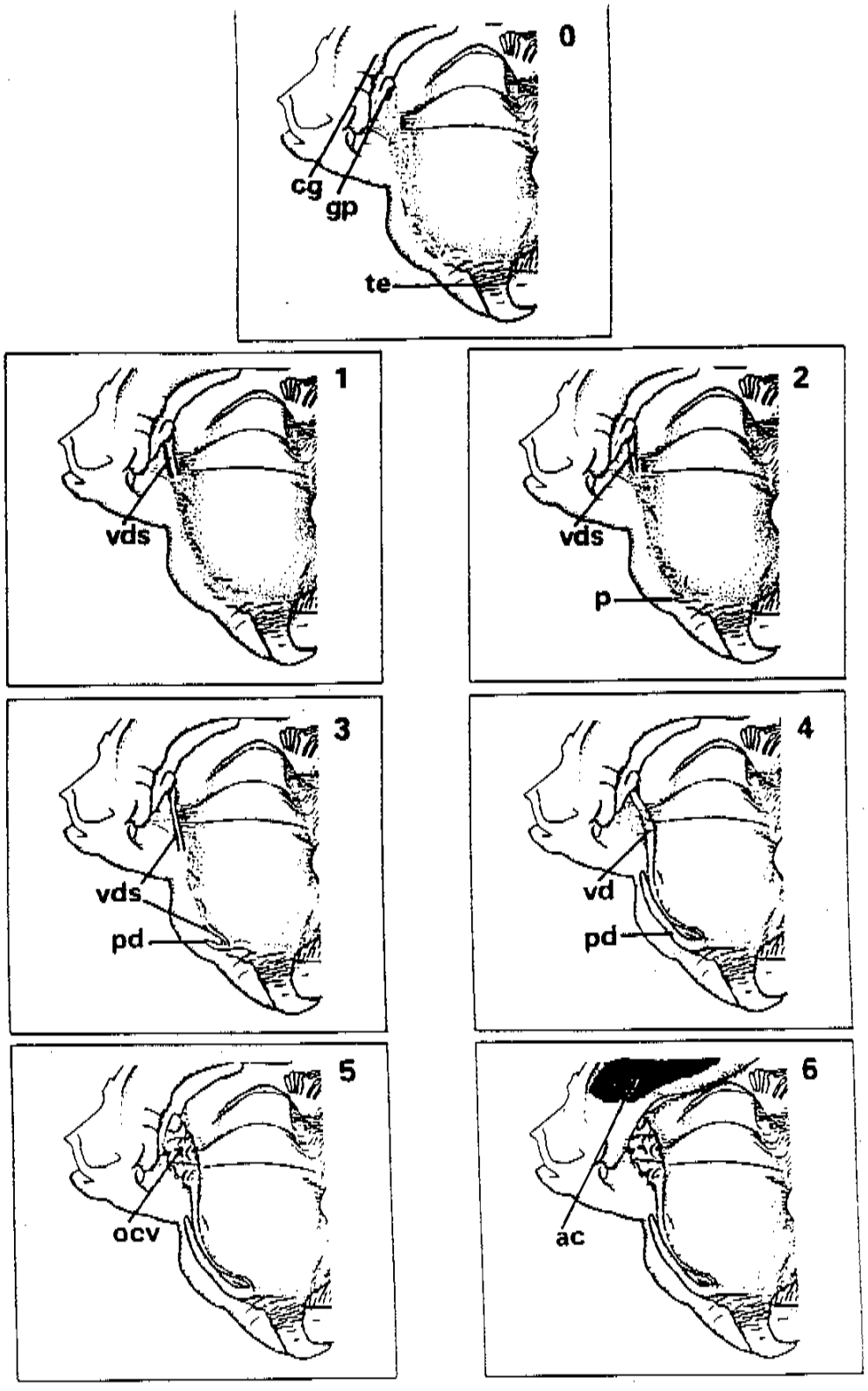


Fig. 1: Imposex development scheme in *Nucella lapillus*. ac, aborted capsules; cg, capsule gland; gp, genital papilla; ocv, occlusion of vaginal opening; p, penis without duct; pd, penis with duct; te, tentacle; vd, vas deferens; vds, vas deferens section.

Figure 1a: Imposex classification scheme for *Nucella lapillus*.

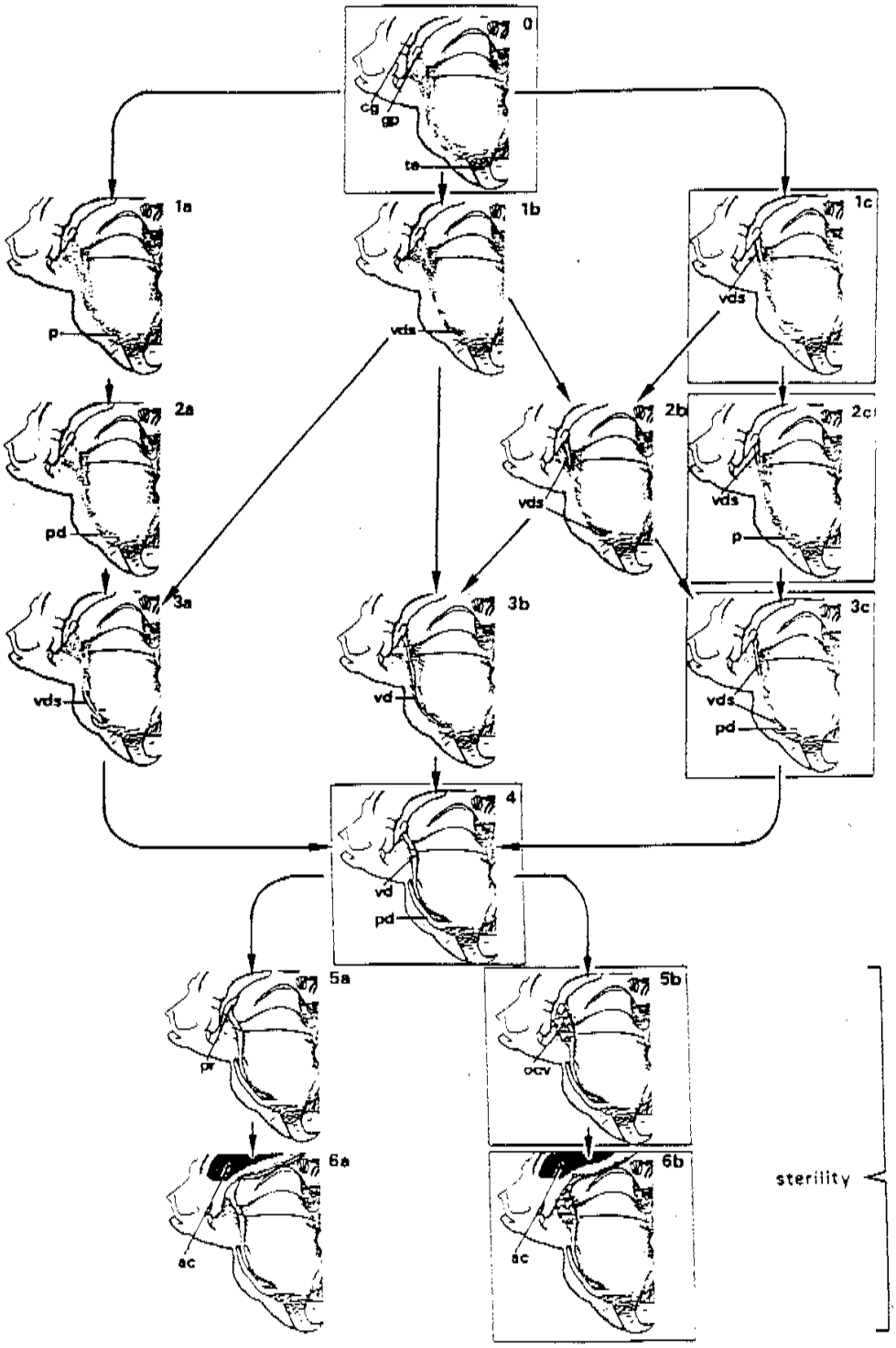


Fig. 2: Imposex development scheme in *Nucella lapillus* according to channel coast (France and England) populations. Stages of Figure 1 with frame. ac, aborted capsules; cg, capsule gland; gp, genital papilla; ocv, occlusion of vaginal opening; p, penis without duct; pd, penis with duct; pr

Figure 1b: General imposex classification scheme for European prosobranch species.

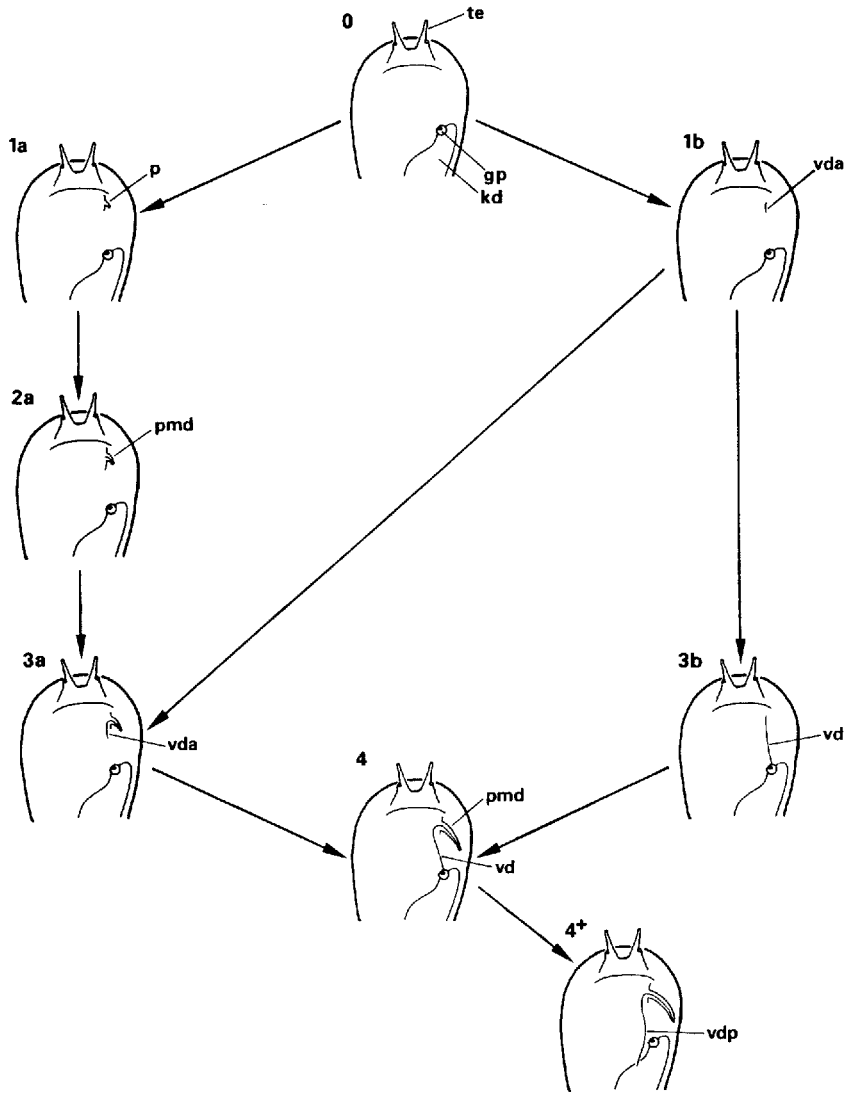


Figure 2a: Imposex classification scheme for *Nassarius reticulatus,* and *Buccinum undatum*. gp, genital papilla; kd, capsule gland; p, penis without duct; pmd, penis with duct; te, tentacle; vd, vas deferens; vda, vas deferens section; vdp, vas deferens passes vaginal opening to run into capsule gland (from Stroben, 1994). It should be noted that Huet *et al*. (1995), Barreiro *et al*. (2001) and Barroso *et al*. (2002) have reported sterilisation in the most heavily affected females, comparable to stage 6 in *Nucella*. There may therefore be scope for further harmonisation of the scoring of VDSI in *Nassarius* in relation to other species.

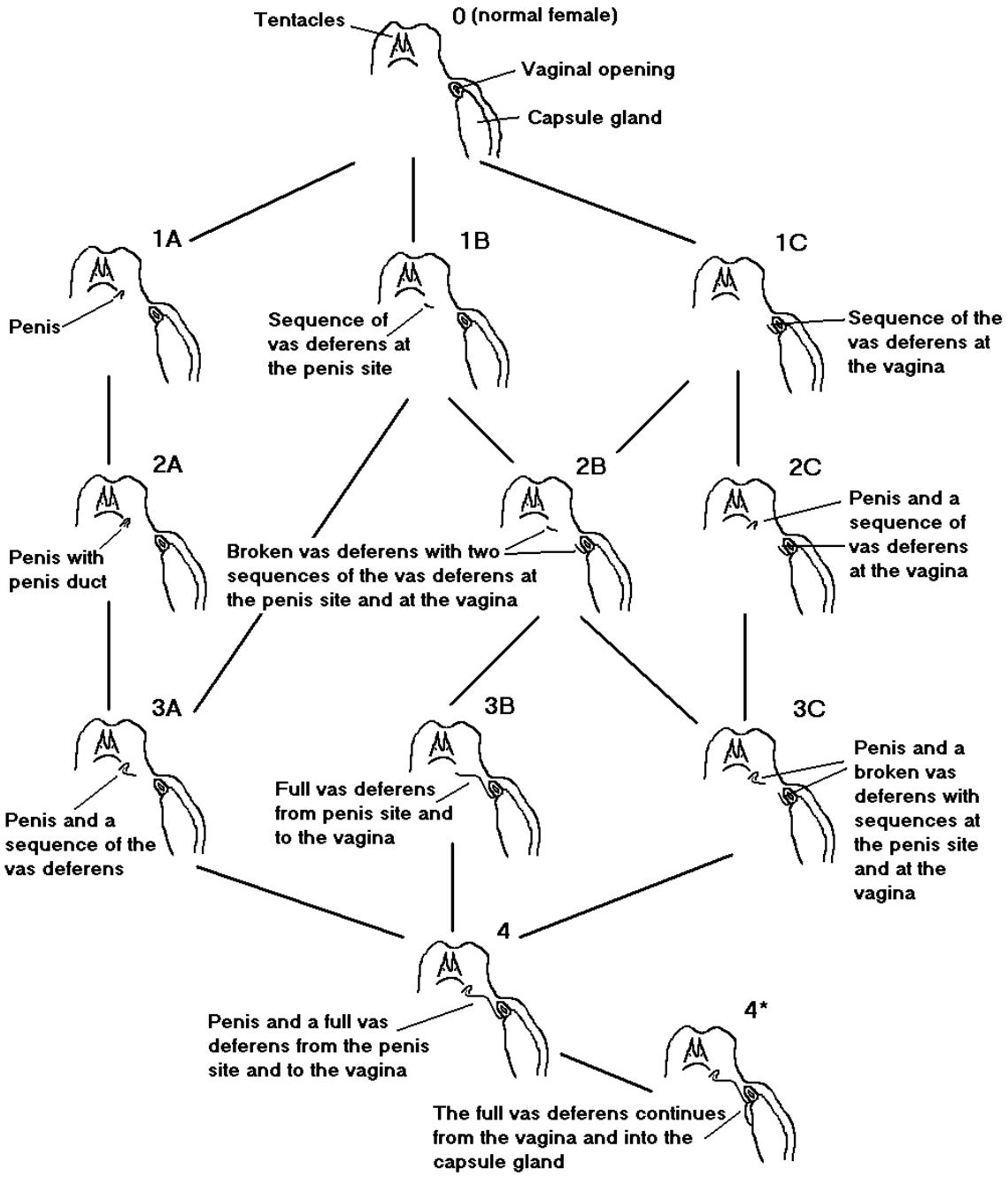


Figure 2b. VDSI imposex classification scheme for *Neptunea*

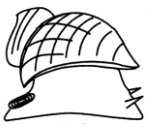
Imposex stage	seen from above	seen from the side
0		
No development of penis. Smooth epithelium at the site, where males have their penis.		
1		
Small knob at penis site indicating the development of a penis.		
2		
Small structure penis, which can wobble and be lifted up from the epithelium of the foot.		
3		
Penis is bent and with a shape which tends to look like a normal male penis.		

Figure 3: Imposex classification scheme for *Buccinum undatum* and Penis Classification Index (PCI). This is a modified classification scheme based on Mensink *et al.* (1997).
If part of, or a full vas deferens is present in addition to the penis, the value of the Penis Classification stage is increased by 0.5. If only a vas deferens is present and not a penis, the value of the Penis Classification stage is 0.5.

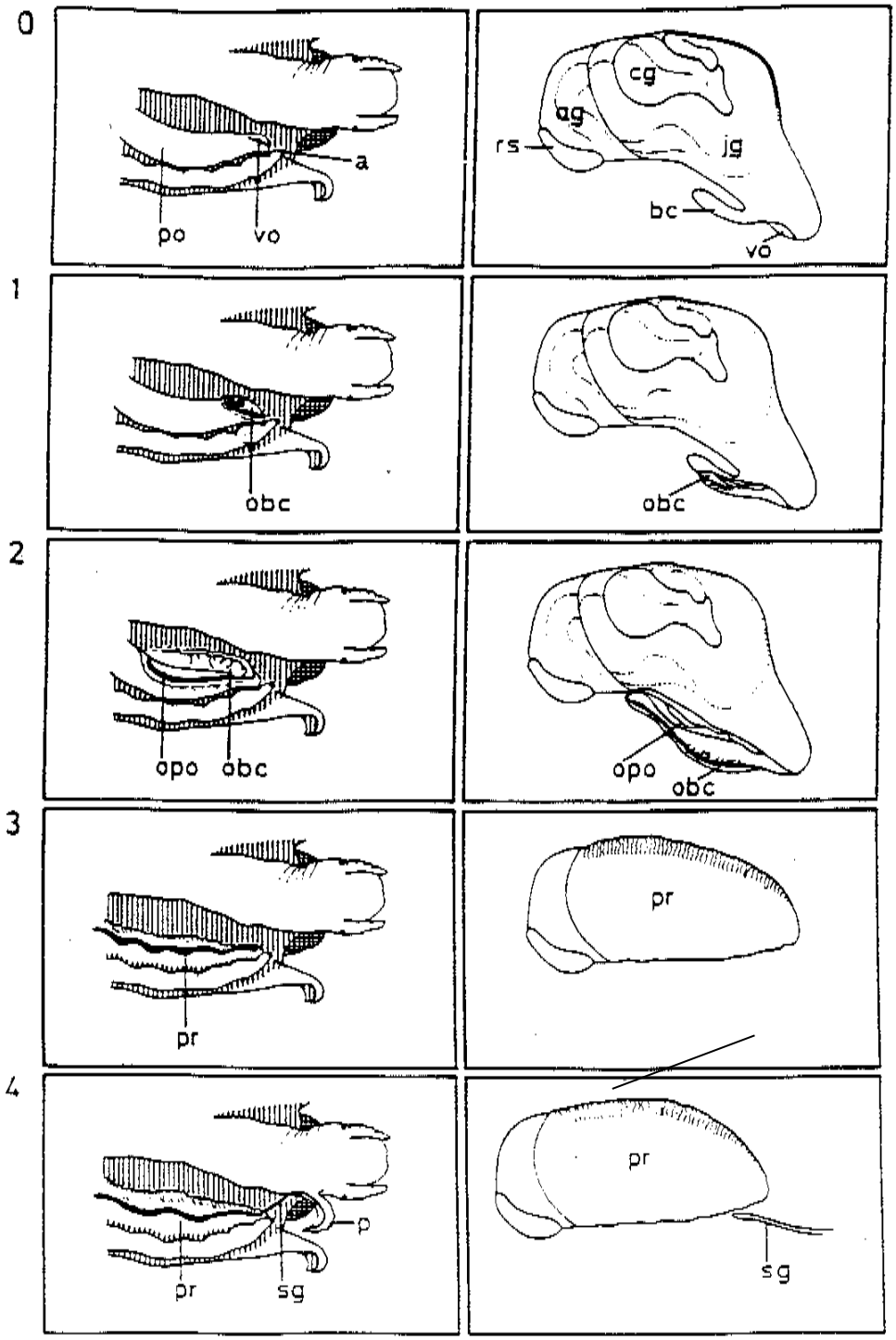
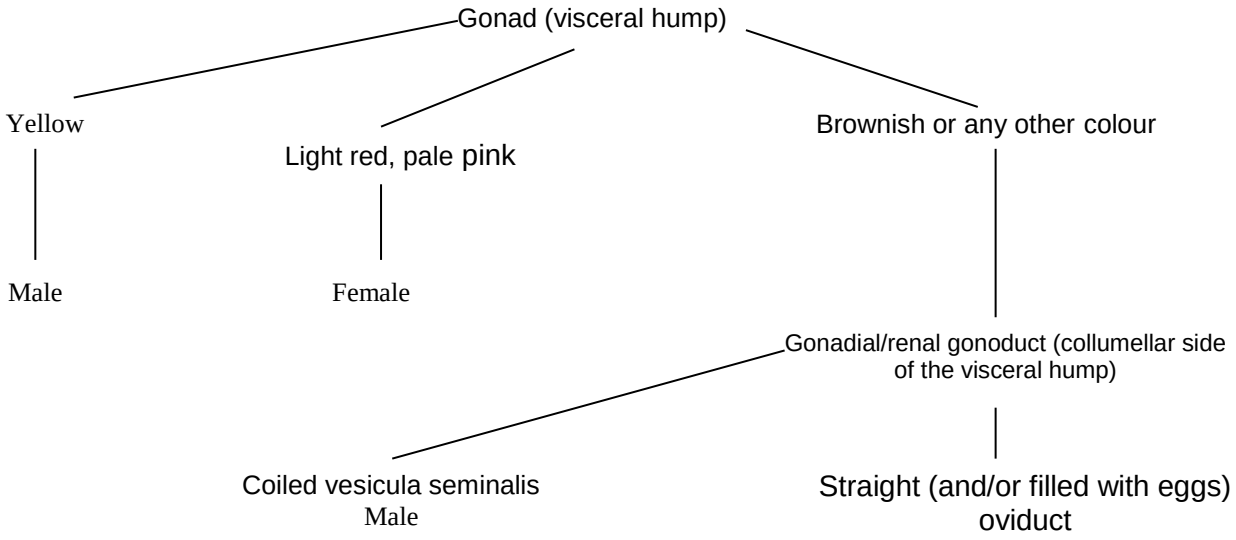


Fig. 4: Scheme of intersex development in *Littorina littorea*. Dorsal views with opened mantle cavity (left) and lateral views of pallial oviduct (right). a, anus; ag, albumen gland; bc, bursa copulatrix; cg, capsule gland; jg, jelly gland; obc, open bursa copulatrix; opo, open pallial oviduct; p, penis; po, pallial oviduct; pr, prostate; rs, receptaculum seminis; sg, sperm groove; vo, vaginal opening.

Figure 4: Imposex classification scheme for *Littorina littorea*.

a) Separation of the sexes



b) Identification of the intersex stages

Identification showing how to distinguish between the sexes (a) of *Littorina littorea* and the intersex stages (b).

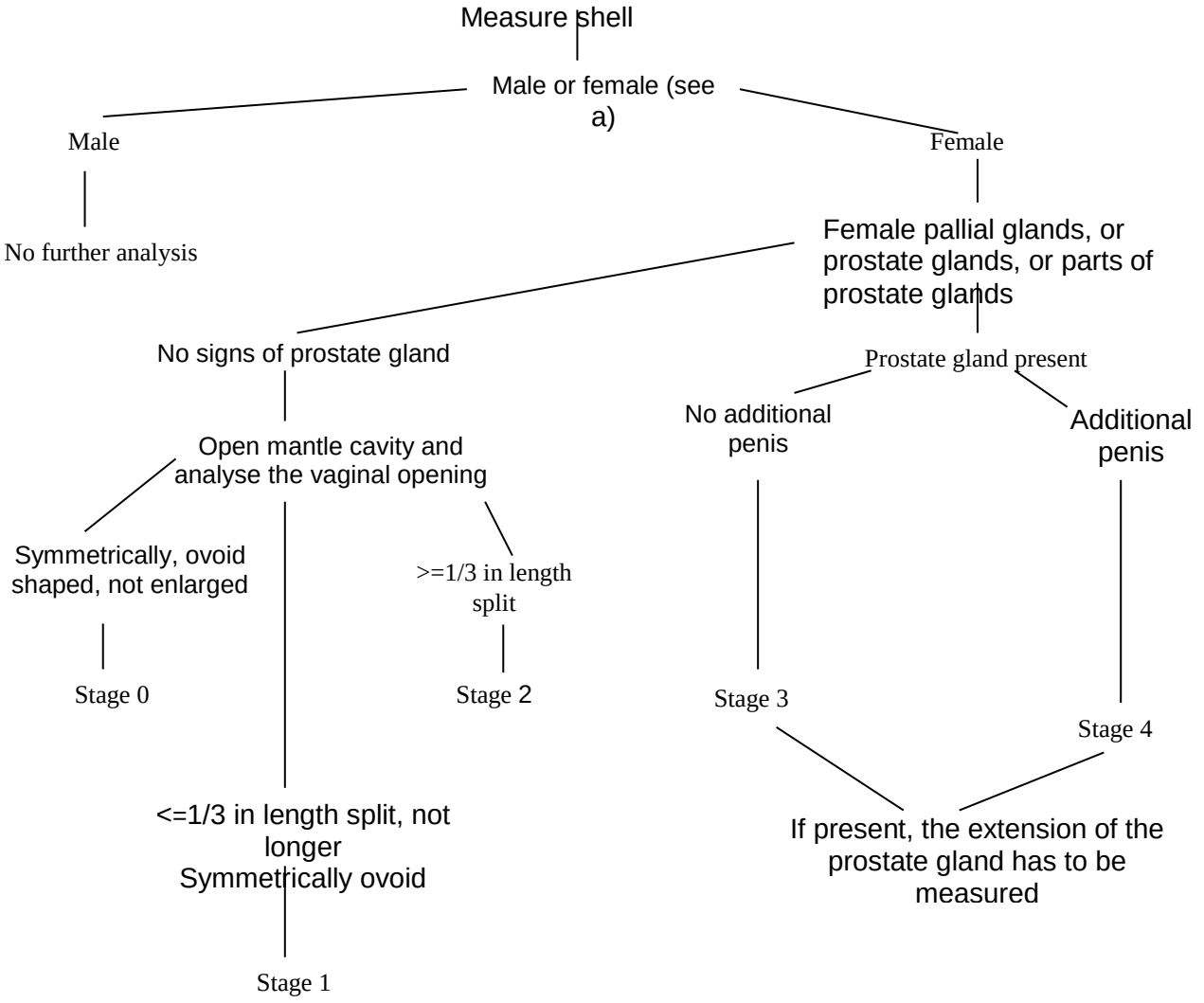


Figure 5: Identification showing how to distinguish between the sexes (a) of *Littorina littorea* and the intersex stages (b).

Technical Annex 4
Estrogen-specific biological effects monitoring

1. Introduction

1. Endocrine disruption by estrogenic substances (estrogens and xenoestrogens) and their impact on marine wildlife has become a major environmental concern in the past few years. Since fish play a pivotal role in the aquatic environment and are natural integrators of water quality conditions, they are considered to be especially suitable as sentinels for environmental endocrine disrupters (Kendall *et al.,* 1998; Vethaak *et al.*, 2006). Thus, monitoring programmes of estrogen-specific effects should primarily focus on finfish (see section 1). Methods to assess the effects on fish early life stages at breeding grounds and on shellfish and other marine invertebrates are less advanced or sensitive, and further research is required before the necessary tools become available.

2. Many substances present in the environment are able to induce changes in fish that are also caused by exposure to the naturally occurring hormone 17β-oestradiol (E2), and are thus collectively known as xenoestrogens. These substances act directly with estrogen receptors and enhance the natural effects of the female sex hormone estradiol. Apart from receptor interactions, estrogen-specific effects may also be caused by mechanism such as hormone synthesis, transport and metabolism. End points of estrogenic disruption may include effects such as reduced fertility, feminization of males, and changes in the sex ratios of progeny (e.g. CSTEE, 1999; Damstra *et al.*, 2002). Major issues of concern associated with estrogenic substances include critical time windows of exposure, additivity by two or more estrogens and low dose effects (e.g. Kendall *et al.,* 1998; CSTEE, 1999, Damstra *et al.*, 2002).

3. To date, a mounting number of studies have documented feminization of wild fish populations (featured by vitellogenin induction and intersexuality) from a variety of contaminated freshwaters (see Jobling *et al.*, 1998; Tyler *et al.*, 1998, Vos *et al.*, 2000; Vethaak *et al.*, 2006). Vitellogenin is a protein naturally produced by oviparous females as they become sexually mature, and which is incorporated into oocytes to form egg yolk. Intersex, in the case of feminization, refers to the presence of ovarian tissue in the male gonad. The widespread nature of estrogenic effects is also evident from recent studies which have demonstrated increased vitellogenin in male fish in areas away from point sources, e.g. in estuarine and marine environments (e.g.; Fossi *et al.,* 2002; Kirby *et al.,* 2004; Scott *et al.,* 2006). The causes of these phenomena are not yet known (Matthiessen, 2003). Feminization of wild fish populations (featured by vitellogenin induction and intersexuality in male fish) is well known to occur in the direct vicinity of sewage treatment plant effluents as a consequence of exposure to estrogenic steroids (estrone, estradiol and ethynylestradiol) and/or phenolic compounds (e.g. alkylphenols and their polyethoxylates, APs/APEOs). This may however be different for marine species, especially top-predator fish, which may experience estrogenic disruption due to biomagnification of weak estrogens such as organochlorines or brominated flame retardants, for example (e.g. Matthiessen, 2003; Vethaak *et al.*, 2006).

2. Sampling strategy

4. It is generally recognized that monitoring programs for estrogenic substances should be based on an integrated approach using a combination of complementary methods (Vethaak *et al.*, 1997, 2006; Routledge, 2003; Segner, 2003; OSPAR, 2005). These typically include: confirmation of exposure through determination of body burden concentrations or estradiol-equivalent potency and biological responses in the same sentinel species; joint application of chemical analyses (external and internal concentrations) and biological effects assessment techniques (using diagnostic biomarkers, mechanism-based bioassays).

5. As the biological activity of estrogen disrupters can occur at very low concentrations and because of their additive nature of their action, the presence and significance of estrogenic substances should be assessed through biological screening complementary to chemical analysis. The measurement of bile estrogenicity of male fish using reporter gene assays (CALUX or YES) is proposed as a mechanism-based screen to yield useful information about the fish's internal exposure to (xeno-) estrogens. In addition, the same assays can be used for measuring the estrogenicity of abiotic extracts and in relation to a bioassay-directed fractionation approach (e.g. Toxicity Identification and Evaluation, TIE or Effect-Directed Analysis (EDA)) to identify those chemicals responsible for the observed estrogenicity of environmental and biotic samples (e.g. Desbrow *et al.*, 1998; Houtman *et al.*, 2004).

E2 plays a major role in the sexual differentiation and maturation of female fish, acting largely through nuclear hormone receptors, but also through membrane receptors. Male fish also express oestrogen receptors (ERs) and thus show biological responses following exposure to xenoestrogens or exogenous E2. Oestrogenic exposure affects multiple mechanistic pathways, but one of the most readily determined endpoints is the production of egg-yolk precursor protein vitellogenin (Vtg) in male fish. Xenoestrogens bind to hepatic oestrogen receptors (ERs), which in turn bind to the oestrogen responsive element (ERE) on the

promoter region of oestrogen-inducible genes. This up-regulates the transcription of such genes (e.g. ERs, Vtg, choriogenin), and production of the corresponding proteins (Rotchell and Ostrander, 2003).

6. Oestrogenic exposure of finfish may be determined at the molecular-level (*in vitro* estrogenicity of deconjugated estrogens in bile fluid; hepatic mRNA expression) or the biochemical level (plasma Vtg protein concentrations). A third endpoint, gonadal intersex can be induced by oestrogenic exposure (but may have other causes, with a low-level prevalence of intersex occurring naturally in some species). The responses of these three end-points to oestrogenic exposure occur on differing time-scales, with the formation of bile estrogen conjugates responding most quickly (hours), induction of the egg yolk precursor protein vitellogenin responding quickly (days) and intersex more slowly (months). Integrated monitoring using the three endpoints may therefore give information on short-term temporal changes in exposure (Hiramatsu *et al.,* 2006).

7. The induction of Vtg in male fish exposed to multiple xenoestrogens is an additive response, based upon the relative concentrations and potencies of the individual compounds in the mixture, and individual compounds can contribute to the overall effect even when they are present at below their no effect concentrations (Silva *et al.,* 2002; Thorpe *et al.,* 2001; Brian *et al.,* 2005). Since exposure to xenoestrogens can result in million-fold increases in plasma Vtg concentrations, determination of Vtg in male fish plasma is an extremely sensitive marker of oestrogen exposure (Sumpter and Jobling, 1995; Hiramatsu *et al.,* 2005, 2006. Vtg is not a good biomarker of oestrogenic exposure in female fish as natural plasma concentrations fluctuate seasonally; sampling for environmental surveys must therefore obtain sufficient male fish at each site to provide the required statistical power.

8. Estrogens, such as E2, are eliminated by metabolic conversion to less active or inactive water-soluble metabolites that are excreted via urine and/or via bile in excreta. In fish, glucuronidation is the dominant conjugation reaction for biliary excretion of steroids as well as xeno-biotic compounds (Truscott, 1979, 1983; see also Legler *et al.*, 2002a). In rainbow trout, for example, over 90% of estrogens are excreted in bile as E2-glucuronides (Forlin and Haux, 1985). Estrogens in bile are excreted into the intestines, where they may be broken down by intestinal bacteria, forming a *de novo* source of the active parent compound (Klaassen and Watkins, 1984). Bacterial enzymes such as beta-glucuronidase, found in various microorganisms such as *E. coli*, are capable of hydrolyzing acid glucuronides back to their primary compounds (Ralovich *et al.*, 1991). Therefore the measurement of deconjugated bile forms a useful (indirect) biomarker for internal dose of xenoestrogens in male fish. Legler et al (2002a) for example have shown that measuring the estrogenic activity in deglucuronidated bile samples of bream is well associated with induction of plasma vitellogenin in the same fish, as well as with (xeno-) estrogenic activity in water from these locations.

9. Another feature of feminization of male fish in the aquatic environment associated with exposure to estrogenic substances concerns the occurrence of ovotestis (or intersexuality, a condition in which oocytes are formed in the testicular tissue). These abnormalities have been observed in rivers, coastal waters and estuaries in the United Kingdom (Jobling *et al.*, 1998; Allen *et al.*, 1999a,b; Kirby *et al.*, 2004), the Baltic Sea (Gercken and Sordyl, 2002) and freshwaters sites in other countries (Sole *et al.*, 2003; Barnhoorn *et al.*, 2004; Vethaak *et al.*, 2005; Bjerregaard *et al.*, 2006). Although intersexuality has most often been observed in waters receiving effluents of sewage treatment plants, recent reports indicate high prevalence (25%) for Mediterranean swordfish (De Metrio *et al.*, 2003). Intersex has also been reported in species such as dab from offshore waters in the OSPAR Maritime Area, although infrequently and at low prevalence (Stentiford and Feist, 2005; Scott *et al.*, 2006). Although not yet fully understood, there is evidence indicating that wild intersex fish (roach) are compromised in their reproductive capacity and that the phenomenon therefore has a potential impact on populations, at least on a local scale (Jobling *et al.*, 2002). It is known that intersex in males is a naturally occurring in low prevalence in various species of fish, including roach (Routledge *et al.*, 1998); therefore possible background levels of this condition should be taken into account (Matthiessen *et al.*, 1998).

10. High levels of Vtg in male fish sometimes coincide with the observation of intersex conditions, but both phenomena are not necessarily causally linked (Jobling *et al.*, 1998; Vethaak *et al.*, 2005). It is conceivable that ovotestis in adult male fish reflects exposure to (xeno-)estrogens during younger life stages when sexual differentiation took place while elevated levels of Vtg in the plasma of adult males is caused by more recent exposure. Allen *et al.* (1999b) also argued that ovotestis is almost certainly induced at the larval stage.

11. Given the above, monitoring programmes of estrogen-specific effects should focus on adult male fish. The following 3 endpoints of endocrine disruption in individual male fish should be measured: (1) bile estrogenic activity; (2) plasma vitellogenin induction; (3) gonadal intersex. These endpoints are collectively diagnostic of estrogenic contaminant exposure and effects and can give information on severity of exposure, its duration and the stage of development at which the organism was exposed, as well as the severity of damage or effects on fish health. The histological evaluation of gonads for signs of intersex is a laborious task and given its generally low prevalence in offshore areas (see section 3) will require large sampling sizes (about 50 male fish). Therefore, it should be undertaken only at selected sites based on historic Vtg levels. A limited number of laboratories are currently able to determine the estrogenicity of fish bile.

3. Sampling locations

12. Biological effects monitoring for estrogen-specific responses can be conducted on both a local scale, in relation to known or probable sites/sources of contamination (e.g. estuaries that are heavily populated and/or industrialised, or that receive run-off from intensive agriculture), but also more widely within the OSPAR Maritime Area, since low-level exposure may be occurring on the regional sea scale (Fossi *et al.,* 2002; Scott *et al.,* 2006). Areas away from hotspots and presumably pristine areas should be included to ensure that unexpected occurrences are not missed. TIE/EDA can be undertaken to analyse and to identify responsible substances and control estrogenic hot spots.

4. Methods to be used

13. The status of the methods to be used is described below and a summary is given at appendix 1. Detailed information on the recommended methodology is given in Section 7.

4.1 Bile estrogenicity

14. The technique needs further refinement and development of QA procedures. An ICES Times protocol, a technical annex and assessment criteria also require to be prepared. The method can be based on the method as described by Legler *et al.*, 2002a; Houtman *et al.*, 2004 and Allard *et al.* 2004. This includes the use of beta-glucuronidase for the incubation of male fish bile samples and the quantification of estrogenic activity of the deconjugated (xeno-)estrogens in the *in vitro* estrogen receptor (ER)-mediated Chemical Activated Luciferase gene eXpression (ER-CALUX) assay with stably-transfected T47D breast cancer cells (Legler *et al.*, 2002a) or the Yeast Estrogen Screen (YES) with a recombinant yeast strain containing the human ERα gene (Allard *et al.*, 2004).

4.2 Vtg induction

15. Although Vtg gene expression can also be determined by determination of hepatic mRNA, the recommended technique for determining Vtg induction in fish is by measuring Vtg protein concentrations in blood plasma (OSPAR, 2006) and for which an ICES TIMES paper already exists (Scott and Hylland, 2002). Plasma Vtg concentrations are most regularly determined by immunoassay, as described in detail in Tech Annex x.

4.3 Gonadal intersex

16. Histological examination of gonadal tissue is used to quantify both the prevalence and severity of intersex (ovotestis). Analysis can be carried out by standard histological procedures on tissues embedded in paraffin. A severity index (OSI) has been described for flounder and can be used for other OSPAR sentinel species (Bateman *et al.,* 2004; Stentiford and Feist, 2005). An ICES Times protocol and technical annex based on the flounder index require to be prepared. Integration of gonadal histology with liver histopathology will be cost-effective and should be encouraged.

5. Temporal trend monitoring

17. Long-term temporal trend monitoring will be required to assess whether oestrogenic exposure is increasing, stable, or decreasing in response to changes in the regulatory framework. It should take place on an annual basis at sites where exposure has been identified as occurring through the spatial monitoring described in Section 2.

6. Field sampling and sampling equipment

18. The recommendations of the existing JAMP Guidelines on Contaminant Monitoring in Biota and on Contaminant-specific Biological Effects Monitoring apply, with minor modification, to designing a sampling strategy to assess xenoestrogen contamination.

19. Live fish in good condition are required, therefore short fishing hauls (<30 minutes) should be used and captured fish should be maintained in running water of ambient temperature until sampled. The water temperature at the collection site should be recorded and only male fish should be sampled.

20. Flounder (*Platichthys flesus*) should be used for estuarine monitoring and cod (*Gadus morhua*) for "open sea" monitoring. Dab (*Limanda limanda*) can also be used for offshore monitoring, but Vtg concentrations are 10-20 times lower than in cod (Scott *et al.,* 2007), and therefore the latter species is preferred. The same species should be used throughout the survey. At each site, fish should be sampled at the same time of the year every year. Sampling should be undertaken outside the breeding season and, for estuarine monitoring, should take place when the flounder have been resident in the estuaries for several months and before their offshore migration begins (e.g. between January and March in the UK).

21. For bile estrogenicity and plasma Vtg it is recommended that a minimum of 15 (and preferably 20) male fish, of comparable sizes, are sampled from each site. The same individual fish should be used. The recommended size ranges are ≥20 cm (25-29 cm is preferred) for flounder, 20-25 cm for dab and 30-45 cm

for cod. However, the diet of large cod differs from that of smaller cod and they have higher plasma Vtg concentrations (Scott *et al.,* 2006). Therefore, where possible, a separate sample of large cod (>80 cm) should be taken.

22. Due to its generally low prevalence in offshore areas, the analysis of intersex requires larger sampling sizes (about 50 male fish). These male fish may include the 15 to 20 fish already taken for bile and plasma samples. However, where PAH-specific biological effect monitoring is conducted in addition to estrogen-specific effect monitoring, the size ranges of sentinel fish should be similar to those used for liver histopathology. Further, because the severity of intersex likely increases with age (Linely *et al.*, 2005), the age of each individual fish should be determined.

23. Fish should be anaesthetised and blood and bile sampled as soon as possible after capture. Because Vtg and bile deconjugates are relatively unstable, samples need to be collected into cold tubes, kept on ice, and in the case of Vtg centrifuged (at 2 - 6^{o} C) within 30 minutes to obtain plasma. Fish weight, fish length and testis weight should be recorded. Otoliths should be collected for age determination, and a portion of the testis fixed (in Bouin's solution, or buffered formal saline) for histological determination of intersex prevalence and severity.

7. Storage and pre-treatment of samples

7.1 Bile samples

24. The samples should be stored in freezers at -20^{o} C for a maximum of six months.

7.2 Plasma samples

25. Plasma samples should be aliquoted into two vials, flash frozen in liquid nitrogen and transported ultra-frozen to the laboratory. The aliquots should be stored in separate freezers at below -60 ^{o}C, and should not be re-frozen after defrosting. If an analysis needs to be repeated, the duplicate sample should be used. Samples must be pre-diluted to avoid matrix effects during the assay.

7.3 Gonadal intersex

26. Gonadal tissue sample should be stored in Bouin's solution, or buffered formal saline.

8. Analytical procedures

8.1 Bile estrogenicity

27. Oestrogenic compounds in bile samples should be deconjugated using b-glucuronidase–arylsulfatase according to Legler *et al.* (2002a) and the estrogenic activity measured using the ER-CALUX assay or the YES assay. Both *in vitro* reporter gene assays are recommended methods for the CEMP. There are, however, important differences between the two assays. First, the ER-CALUX assay appears to be more sensitive (about a factor 20) than the YES assay (Legler *et al*., 2002b). Second, differences are observed in the sensitivity between the two assays to both (xeno-)estrogens and anti-estrogens (Legler *et al.*, 2002b; Murk *et al*., 2002). The YES assay detects estrogenic agonists, but does not consistently detect anti-estrogenic activity and sometimes identifies it as agonistic (Legler *et al.,* 2002b). Third, the mammalian-based ER CALUX assay has the major drawback, compared to the yeast-based YES assay, that mammalian cells are more difficult and expensive to cultivate, and are more susceptible to cytotoxic effects. A main problem in the utilization of *in vitro* assays to analyse aquatic environmental samples is the presence of inhibitory/cytotoxic compounds. The YES assay may perform better for monitoring of environmental samples, as these samples are frequently contaminated with substances other than (xeno)estrogens that may interfere with the growth and viability of animal cells, but not with yeast cells (Kinnberg, 2003). Since a low detection limit is important for the marine environment, the preferred reporter gene assay is the ER-CALUX, although this is more costly and more difficult to use. Further work is needed to conclude on this issue.

8.2 Plasma Vtg

28. Plasma Vtg is normally determined through immunoassay; and usually this is an enzyme-linked immunosorbent assay (ELISA), although radioimmunoassay (RIA) may be used. The general procedure required to establish an ELISA for Vtg determination has been described (Specker and Anderson, 1994; Scott and Hylland, 2002). However, there are at least three different types of Vtg in some species of fish, and each responds differently to xenoestrogenic exposure (Hiramatsu *et al.,* 2006). To ensure comparability of results between laboratories, it is therefore recommended that a common procedure and a common source of Vtg antibody and antigen are used by all laboratories involved in surveying a particular species. Detailed descriptions of ELISAs are available for the determination of Vtg in flounder (Lahr *et al.,* 2006; Madsen *et al.,* 2003; Kirby *et al.,* 2004), dab (Scott *et al.,* 2007) and cod (Meier *et al.,* 2007). However, cod Vtg appears to be particularly unstable, readily breaking down into its constituent proteins, lipovitellin (Lv) and phosvitin (Pv), and an ELISA using Lv as a standard might be more appropriate for investigating xenoestrogenic exposure in this species (Scott *et al.,* 2006).

8.3 Gonadal intersex

29. Testis samples should be assessed for the prevalence and severity of intersex (ovotestis), as described for flounder (Bateman *et al.,* 2004).

9. Quality assurance

30. A future ICES TIMES paper for bioassays of specific matrices (currently being prepared by the NL, UK and Norway) will describe the deconjugation methodology for fish bile. The YES reporter gene assay has been recommended for biological effects monitoring, and the CALUX assay for oestrogen receptor-active compounds is a promising technique that has been widely used (e.g. OSPAR, 2005; Legler *et al.*, 2002; Vethaak *et al.*, 2006). However, an ICES TIMES document, OSPAR technical annex and QA procedures require development for both of these assays before they can be implemented within the CEMP.

31. An international laboratory performance trial of cod Vtg determination was conducted (2004/05) with five participating laboratories (NIVA, 2005). This demonstrated that, although the laboratories were not using common methodologies and reagents, such as is recommended above, all were able to distinguish between samples from control fish and those exposed to E2. The laboratories also showed acceptable agreement in the Vtg concentrations reported for the induced samples, although inter-laboratory variability was much greater for the control samples. Further performance tests are required that also include plasma from flounder, and plasma samples with a wider range of Vtg concentrations than in the cod Vtg trial.

32. BEQUALM also runs liver histopathology (fish disease) ring tests, distributing images of histological slides of liver tissue. They do not currently conduct similar ring tests to allow intercomparison of laboratories determining gonadal intersex, although such an external QA scheme would be required for proficiency testing of laboratories involved in assessing oestrogenic exposure through intersex determinations.

33. All the samples from a spatial survey should be analysed at the same time to allow randomisation of samples and the use of a single batch of standard and reagents to reduce the inter-assay variation.

10. Assessment criteria

34. Reference criteria are required to allow an assessment of whether determined concentrations or effects are of environmental significance. Based upon the extensive UK dataset, a background concentration (BC) has been proposed for Vtg in flounder (0.13 µg/ml; OSPAR, 2007). The same document also proposed a BC for cod Vtg (0.22 µg/ml), although this was based upon a more limited dataset, and it may be possible to revaluate this in the light of more recent data (Meier et al., 2007). However, significant differences in the methodologies employed to measure cod Vtg concentrations by the UK and in the latter study may not allow this.

35. It may be possible to develop BC/BACs for the estrogenicity of fish bile (as E2-equivalent concentrations, e.g. pMol EEq/ml), and equivalent assessment criteria for the prevalence/severity of intersex in male gonads, although this has not been established.

11. Reporting requirements

36. Data reporting should be in accordance with the current requirements for National Comments and meet the latest ICES reporting formats. The following data are required (or *recommended):

11.1 Contaminant concentrations

37. The additive oestrogenic response following exposure to low environmental concentrations of natural and synthetic (xeno-)estrogens makes biological effects measurements desirable. Whilst chemical analysis methodologies are available for some (xeno-)estrogens, suitable External Quality Assurance (EQA) programmes do not yet exist; this may change in the near future as many of the contaminants of interest are required for monitoring under the WFD. Until the EQA requirements for monitoring chemical monitoring are fulfilled, the following compounds are recommended determinants, but not required:

- Alkylphenols and alkylphenol ethoxylates in sediment and biota*
- Natural (E1, E2) and synthetic hormones (EE2) in water only when suspected*
- Phthalates in sediment and biota*

11.2 Biological effects measurements

38. See earlier comments on the availability of assessment criteria and EQA schemes for biological effects measurements.

- *Bile estrogenicity (pMol EEQ/ml)
- Plasma Vtg concentrations (µg/ml)

- *Intersex prevalence (presence / absence)
- *Ovotestis Severity Index (according to Bateman *et al.,* 2004)

11.3 Supporting parameters

- Fish identifier
- Site code and name
- Latitude and longitude
- Taxonomic identification
- Temperature and salinity at each collection site
- Date of sample collection
- Gender (males only should be sampled)
- Location and description of grossly visible anomalies, lesions and parasites
- Fish weight (g) and length (mm)
- *Fish age
- Gonad somatic index
- Stage of sexual maturation

12. References

Allard, A.-S., Gunnarsson, M., Svenson, A. (2002) Estrogenicity in bile of juvenile rainbow trout as measure of exposure and potential effects of endocrine disruptors. Environmental Toxicology and Chemistry, 23: 1187-1193.

Allen, Y., Scott, A. P., Matthiessen, P., Haworth, S., Thain, J. E., & Feist, S. (1999a) 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(8): 1791-1800.

Allen, Y., Matthiessen, P., Scott, A. P., Haworth, S., Feist, S., & Thain, J. E. (1999b) The extent of oestrogenic contamination in the UK estuarine and marine environments - further surveys of flounder. Science of the Total Environment, 233(1-3): 5-20.

Barnhoorn, I.E.J., Bornman, M.S., Pieterse, G.M., Van Vuren, J.H.J. (2004) Histological evidence of intersex in feral sharptooth catfish (*Clarias gariepinus*) from an estrogen-polluted water source in Gauteng, South Africa. Environmental Toxicology, 19: 603-608.

Bateman, K. S., Stentiford, G. D., and Feist, S. W. (2004) A ranking system for the evaluation of intersex condition in European flounder (*Platichthys flesus*). Environmental Toxicology and Chemistry, 23: 2831-6.

Bjerregaard, L.B. Madsen, A.H. Korsgaard, B., Bjerregaard, P. (2006) Gonad histology and vitellogenin concentrations in brown trout (*Salmo trutta*) from Danish streams impacted by sewage effluent. Ecotoxicology, 15: 315-327.

Brian, J. V., Harris, C. A., Scholze, M., Backhaus, T., Booy, P., Lamoree, M., Pojana, G., Jonkers, N., Runnalls, T., Bonfa, A., Marcomini, A., and Sumpter, J. P. (2005) Accurate prediction of the response of freshwater fish to a mixture of estrogenic chemicals. Environmental Health Perspectives, 113:721-8.

CSTEE. (1999) Human and Wildlife Health Effects of Endocrine Disrupting Chemicals, with emphasis on Wildlife and on Ecotoxicology test methods. Opinion of the European Commission Scientific Committee for Toxicity, Ecotoxicity and the Environment. European Commission, DG XXIV Consumer Policy and Consumer Health Protection, Brussels, Belgium.

Damstra T, Barlow S, Bergman A, Kavlock R, Van der Kraak G. (2002) Global Assessment of the State-of-the-Science of Endocrine Disruptors. WHO/PCS/EDC/02.2. World Health Organization, Geneva, Switzerland.

De Metrio G., A. Corriero , S. Desantis , D. Zubani , F. Cirillo , M. Deflorio, C.R. Bridges , J. Eicker; J.M. De La Serna , P. Megalofonou , D. Kime (2003) Evidence of a high percentage of intersex in the Mediterranean swordfish (*Xiphias gladius* L.). Marine Pollution Bulletin, 46:358-361.

Desbrow C, Routledge EJ, Brighty GC and Waldock M. (1998) Identification of estrogenic chemicals in STW effluent. 1: Chemical fractionation and *in vitro* biological screening. Environmental Science and Technology, 32:1549-1558.

Forlin L, Haux C. (1985) Increased excretion in the bile of 17b- [3H]estradiol-derived radioactivity in rainbow trout treated with b-naphthoflavone. Aquatic Toxicology, 6:197–208.

Fossi, M. C., Casini, S., Marsili, L., Neri, G., Mori, G., Ancora, S., Moscatelli, A., Ausili, A., and Notarbartolo-di-Sciara, G. (2002) Biomarkers for endocrine disruptors in three species of Mediterranean large pelagic fish. Marine Environmental Research, 54: 667-671.

Gercken J.1; Sordyl H. (2002) Intersex in feral marine and freshwater fish from northeastern Germany. Marine Environmental Research, 54: 651-655.

Hiramatsu, N., Cheek, A. O., Sullivan, C. V., Matsubara, T., and Hara, A. (2005) Vitellogenesis and endocrine disruption. In: Biochemistry and Molecular Biology of Fishes. (Eds. T. P. Mommsen and T. W. Moon.) pp. 431-71. Elsevier, Amsterdam.

Hiramatsu, N., Matsubara, T., Fujita, T., Sullivan, C. V., and Hara, A. (2006) Multiple piscine vitellogenins: biomarkers of fish exposure to estrogenic endocrine disruptors in aquatic environments. Marine Biology, 149:35-47.

Houtman CJ, van Oostveen AM, Brouwer A, Lamoree MH, and Legler J. (2004) Identification of Estrogenic Compounds in Fish Bile Using Bioassay-Directed Fractionation. Environmental Science and Technology, 38:6415 -6423.

Jobling S, Nolan M, Tyler CR, Brighty G, Sumpter JP. (1998) Widespread sexual disruption in wild fish. Environmental Science and Technology, 32:2498-2506.

Jobling S, Beresford N, Nolan M, Rodgers-Gray TP, Brighty G, Sumpter JP, Tyler CR. (2002) Altered sexual maturation and gamete production in wild roach (*Rutilus rutilus*) living in rivers that receive treated sewage effluents. Biology of Reproduction, 66:272-281.

Kendall RJ, Dickerson RL, Suk WA, Giesy JP. (1998) Principles and Processes for Evaluating Endocrine Modulation in Wildlife. SETAC, Pensacola, FL.

Kinnberg, K. (2003) Evaluation of in vitro assays for determination of estrogenic activity in the environment. Danish Environmental Protection Agency, Danish Ministry of the Environment Working Report 2003 No. 43.

Kirby, M. F., Allen, Y. T., Dyer, R. A., Feist, S. W., Katsiadaki, I., Matthiessen, P., Scott, A. P., Smith, A., Stentiford, G. D., Thain, J. E., Thomas, K. V., Tolhurst, L., and Waldock, M. J. (2004) Surveys of plasma vitellogenin and intersex in male flounder (*Platichthys flesus*) as measures of endocrine disruption by estrogenic contamination in United Kingdom estuaries: temporal trends, 1996 to 2001. Environmental Toxicology and Chemistry, 23: 748-58.

Klaassen CD, Watkins JB. 1984. Mechanisms of bile formation, hepatic uptake and biliary excretion. Pharmacology Reviews, 36:1–67.

Lahr, J., Kuiper, R.V., van Mullem, A., Verboom, B.L., Jol, J., Schout, P., Grinwis, G.C.M., Rouhani Rankouhi, T., Pieters, J.P.F., Gerritsen, A.A.M., Giesy, J.P., Vethaak, A.D. (2006) A field survey of estrogenic effects in freshwater and marine fish in The Netherlands Chapter 7. In: Estrogens and Xeno-Estrogens in the Aquatic Environment: An Integrated Approach for Field Monitoring and Effect Assessment. Vethaak, A.D., Schrap, S.M., de Voogt, P. (eds.), Society of Environmental Toxicology and Chemistry (SETAC). SETAC Press, Pensacola, FL, USA, pp 151-178. ISBN 1-880611-85-6.

Legler J, Jonas A, Lahr J, Vethaak AD, Brouwer A, Murk AJ. (2002a) Biological measurement of estrogenic activity in urine and bile conjugates using the in vitro ER-CALUX reporter gene assay. Environmental Toxicology and Chemistry, 21:473-479.

Legler J, Dennekamp M, Vethaak AD, Brouwer A, Koemann JH, van der Burg B, Murk AJ (2002b) Detection of estrogenic activity in sediment associated compounds using in vitro reporter gene assays. The Science of the Total Environment, 293: 69-83.

Liney Katherine E., Susan Jobling, Jan A. Shears, Peter Simpson, and Charles R. Tyler (2005) Assessing the Sensitivity of Different Life Stages for Sexual Disruption in Roach (*Rutilus rutilus*) Exposed to Effluents from Wastewater Treatment Works. Environmental Health Perspectives, 113(10): 1299–1307.

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

Matthiessen P. (2003) Endocrine disruption in marine fish. Pure and Applied Chemistry, 75:2249–2261.

Matthiessen, P., Allen, Y. T., Allchin, C. R., Feist, S. W., Kirby, M. F., Law, R. J., Scott, A. P., Thain, J. E. and Thomas K. V. (1998) Oestrogenic endocrine disruption in flounder (*Platichthys flesus* L.) from United Kingdom estuarine and marine waters. Science Series, Technical Report No. 107. Centre for Environment, Fisheries and Aquaculture Science, Burnham-on-Crouch, UK. 48pp.

Meier, S., Andersen, T. E., Norberg, B., Thorsen, A., Taranger, G. L., Kjesbu, O. S., Dale, R., Morton, H. C., Klungsoyr, J., and Svardal, A. 2007. Effects of alkylphenols on the reproductive system of Atlantic cod (*Gadus morhua*). Aquatic Toxicology, 81: 207-18.

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. Environmental Toxicology and Chemistry, 21:16-23.

NIVA (2005) BEQUALM vitellogenin intercalibration exercise – draft report, June 2005. Available from: Kjetil Hylland, Norwegian Institute for Water Research (NIVA), Brekkeveien 19, P.O. Box 173, Kjelsaas, N-0411 Oslo, Norway.

OSPAR (2005) Possible routes towards the assessment of endocrine disruption (ED) in the marine environment, including an overview of recent UK Government-funded marine ED research. Presented by the UK. OSPAR SIME 05/02 16-E.

OSPAR (2007) Background document for biological effects monitoring techniques: fish vitellogenin as a biomarker of exposure to xenoestrogens. Presented by the UK. OSPAR SIME 07/03/09-E.

Ralovich B, Ibrahim GAM, Fabian A, Herpay M. (1991) Beta-Dglucuronidase (BDG) activity of gram-negative bacteria. Acta Microbiologica Hungarica, 38:283–291.

Rotchell, J. M. and Ostrander, G. K. (2003) Molecular markers of endocrine disruption in aquatic organisms. Journal of Toxicology and Environmental Health B Critical Reviews, 6:453-496.

Routledge, E. J., Sheahan, D., Desbrow, C., Brighty, G. C., Waldock, M., and Sumpter, J. P. (1998). Identification of estrogenic chemicals in STW effluent. 2. *In vivo* responses in trout and roach. Environmental Science and Technology, 32:1559-1565.

Routledge EJ. (2003) Identifying the causative agents: The use of combined chemical and biological strategies in monitoring programs. Pure and Applied Chemistry, 75:2461–2466.

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

Scott, A. P., Katsiadaki, I., Witthames, P. R., Hylland, K., Davies, I. M., McIntosh, A. D., and Thain, J. (2006) Vitellogenin in the blood plasma of male cod (*Gadus morhua*): a sign of oestrogenic endocrine disruption in the open sea? Marine Environmental Research, 61: 149-70.

Scott, A. P., Sanders, M., Stentiford, G. D., Reese, R. A., and Katsiadaki, I. (2007). Evidence for estrogenic endocrine disruption in an offshore flatfish, the dab (*Limanda limanda* L.). Marine Environmental Research, 64: 128-148.

Segner H. (2003) The need for establishing integrated monitoring programs. Pure and Applied Chemistry, 75:2435-2444.

Silva, E., Rajapakse, N., and Kortenkamp, A. (2002) Something from "nothing"--eight weak estrogenic chemicals combined at concentrations below NOECs produce significant mixture effects. Environmental Science and Technology, 36: 1751-6.

Solé, M., Raldua, D., Barceló, D., Porte, C. (2003) Long-term exposure effects in vitellogenin, sex hormones, and biotransformation enzymes in female carp in relation to a sewage treatment works. Ecotoxicology and Environmental Safety, 56 (3):373-380.

Specker, J. L. and Anderson, T. R. (1994) Developing an ELISA for a model protein -- vitellogenin. In: Biochemistry and Molecular Biology of Fishes. (Eds. P. Hochacka and R. Mommsen.) pp. 567-78. Elsevier, Amsterdam.

Stentiford, G. D. and Feist, S. W. (2005) First reported cases of intersex (ovotestis) in the flatfish species dab *Limanda limanda*: Dogger Bank, North Sea. Marine Ecology Progress Series, 301:307-310.

Sumpter, J. P. and Jobling, S. (1995) Vitellogenesis as a biomarker for estrogenic contamination of the aquatic environment. Environmental Health Perspectives, 103 Suppl 7: 173-178.

Thorpe, K. L., Hutchinson, T. H., Hetheridge, M. J., Scholze, M., Sumpter, J. P., and Tyler, C. R. (2001) Assessing the biological potency of binary mixtures of environmental estrogens using vitellogenin induction in juvenile rainbow trout (*Oncorhynchus mykiss*). Environmental Science and Technology, 35: 2476-81.

Truscott B. (1979) Steroid metabolism in fish. Identification of steroid moeities of hydrolyzable conjugates of cortisol in the bile of trout *Salmo gairdnerii*. General and Comparative Endocrinology, 38:196–206.

Truscott B. (1983) Steroid metabolism in fish. II. Testosterone metabolites in the bile of the marine winter flounder *Pseudopleuronectes americanus* and the freshwater Atlantic salmon *Salmo salar*. General and Comparative Endocrinology, 51:460–470.

Tyler CR, Jobling S, Sumpter JP. (1998) Endocrine disruption in wildlife: a critical review of the evidence. Critical Reviews in Toxicology, 28:319-61.

Vethaak AD, Jobling S, Waldock M, Bjerregärd P, Dickerson R, Giesy J, Grothe D, Karbe L, Munkittrick K, Schlumpf M, Sumpter JP. (1997) Approaches for the conduct of field surveys and toxicity identification and evaluation in identifying the hazards of endocrine modulating chemicals to wildlife. In Tattersfield L, Matthiessen P, Campbell P, Grandy N, Länge R, eds. Society of Environmental Toxicology and Chemistry (SETAC) SETAC-Europe/OECD/EC Expert workshop on endocrine modulators and wildlife: assessment and testing; 1997 April 10-13; Veldhoven, the Netherlands. SETAC-Europe, Brussels, Belgium.

Vethaak AD, Lahr J, Schrap SM, Belfroid AC, Rijs GBJ, Gerritsen A, de Boer J, Bulder AS, Grinwis GCN, Kuiper RV, Legler J, Murk TAJ, Peijnenburg W, Verhaar HJM, de Voogt P. (2005) An integrated assessment of estrogenic contamination and biological effects in the aquatic environment of the Netherlands. Chemosphere, 59:511-524.

Vethaak, D, M. Schrap and P. de Voogt (eds) (2006) Estrogens and Xenoestrogens in the Aquatic Environment: An Integrated Approach for Field Monitoring and Effect Assessment. Society of Environmental Toxicology and Chemistry (SETAC), Pensacola, FL. ISBN1-880611-85-6. 481 pp.

Vos JG, Dybing E, Greim HA, Ladefoged O, Lambré C, Tarazona JV, Brandt I, Vethaak AD. (2000) Health effects of endocrine-disrupting chemicals on wildlife, with special reference to the European situation. Critical Reviews in Toxicology, 30:71-133.

Appendix 1

Status of methods for estrogen-specific effects monitoring (as at March 2008)

Method	in JAMP	QA	Recommended by ICES	ICES TIMES document	Background document	Assessment Criteria
In vitro bile estrogenicity*	no		yes	no	In prep by NL	Is possible but not established yet
Specific matrix bile extraction-deconjugation				draft by NL, Norway and UK		
in vitro bioassays YES or		intercalibrated	yes	not available	yes	see above
in vitro bioassay ERCALUX		no	yes	not available	yes	see above
Vitellogenin	no	Bequalm	yes	yes no 31	yes	yes
Gonadal intersex	no	Bequalm	yes	in prep by UK	in prep by UK	Not established yet