Journal of Aquatic Ecosystem Health **1**: 295–308, 1992.
M. Munawar (ed.), Assessing Aquatic Ecosystem Health: Rationale, Challenges, and Strategies.

An overview of the assessment of aquatic ecosystem health using benthic invertebrates

Trefor B. Reynoldson & Janice L. Metcalfe-Smith
Environment Canada, National Water Research Institute, P.O. Box 5050, 867 Lakeshore Road, Burlington, Ontario L7R 4A6, Canada

Keywords: benthos, biomonitoring, community structure, indices

Abstract. Community structure or species composition of benthic invertebrates has frequently been used in environmental monitoring and assessment of aquatic systems. Three general approaches have been taken: the 'saprobic' approach, which requires detailed knowledge of taxonomy and is most effective in measuring impacts from sewage effluents; diversity indices, which do not require detailed knowledge of species requirements but ignore information provided by important species and tend to lose information; and biotic indices, which combine both approaches. In the past few years considerable advances have been made by applying multivariate statistical techniques to large data matrices and relating benthic community structure to key environmental variables. Using these techniques it is possible to establish reference communities for a set of environmental conditions, to predict the benthic community that should occur at new sites and thus measure deviation from an expected community type. This suggests that environmental criteria and objectives can be established based on biological variables as opposed to the more traditional chemical approach.

Measurement of ecosystem health using functional attributes of benthic invertebrates is generally in the development stage. In the future, functional measures of ecosystem health, such as chronic measures of toxicity or stress, should be incorporated into any assessment process.

1. Introduction

A major concern of environmental agencies is the remediation of areas that have been adversely affected by human activity. The successful accomplishment of this task requires detailed assessment of the degraded system to determine both the extent to which degradation has occurred and the factors responsible. In addition, remediation targets require precise definition to establish when sufficient action has been taken and to maximize the improvement, either per dollar or per unit time. This is particularly important given the fact that cost per unit improvement relationships are usually asymptotic (Fig. 1).

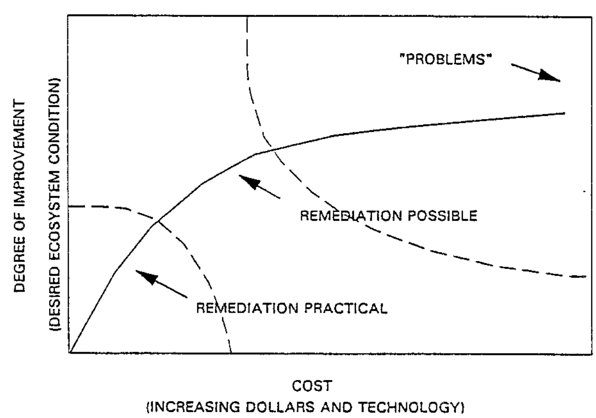

Fig. 1. Usual relationship between cost and level of remediation.

At present government agencies which set regulations, environmental quality criteria, guidelines, and objectives have exclusively employed a chemical concentration approach. This is somewhat ironic as it was landmark publications such as Rachel Carson's 'Silent Spring' and Hynes (1960) 'Biology and Polluted Waters' that played a major part in developing public environmental awareness of the biological implications of pollution. Far-sighted individuals (Patrick, 1949, 1951; Hynes, 1960) suggested several decades ago that biological systems were an essential part of ecosystem health assessment and could provide a basis for environmental management. This broader view is particularly well expressed at the policy level in the 1987 revision of the Great Lakes

Water Quality Agreement between the United States and Canada, which explicitly endorses an ecosystem approach. We suggest further that biological systems should be the standard for monitoring, assessment, and target formulation (i.e. criteria and objectives), and that the role of chemistry and physics is most important in the identification of factors causing impairment and the selection of appropriate remedial actions.

This paper discusses the tools and technology available for determining stress within aquatic ecosystems using benthic invertebrates. The merits of other groups of organisms have not been considered in this paper, but the reader should be aware that other biological groups can also be used for assessment purposes and are discussed in other papers in this volume.

2. Utility of benthos: past and present

There are valid reasons for using benthic invertebrates as part of any aquatic health assessment programme. Benthic invertebrates are present in most habitats and are relatively easy to sample quantitatively compared to other groups such as phytoplankton, zooplankton, and fish; they have recognized community responses to water quality changes (Hynes, 1960; Reynoldson, 1984; Sheldon, 1984; Hilsenhoff, 1987, 1988) and their taxonomy is well established. Furthermore, populations are relatively stable in time compared to microbial or planktonic communities, thus requiring less frequent sampling. Benthic invertebrates are largely non-mobile and thus representative of the location being sampled. The life histories of many species are known, as are their responses to certain pollutants. Laboratory culture is comparatively easy and their size renders them capable of study at the level of the individual. The main disadvantage of the group is that taxonomy to the species level in some groups requires verification by specialists, and spatial heterogeneity requires appropriate sample replication. Sample processing can be time-consuming, precluding rapid enumeration in quantitative studies, although rapid assessment techniques have been proposed that provide qualitative information. The time required for sample processing in quantitative studies can also result in a relatively high labour cost. Some of the disadvantages relating to cost and time required for analysis relate to the fact that routine analyses and standard operating procedures have not been established in environmental laboratories (Reynoldson, 1984; Sheldon, 1984). Lastly, the responses of taxa to toxic materials are not as well documented as their responses to sewage pollution.

A fundamental distinction can be made between structural and functional approaches to bioassessment. Much of the earliest work was of the structural or taxonomic type, where changes in community structure and indicator species were related to anthropogenic and naturally induced stresses (Hynes, 1960; Warren, 1971). Community structure information, which provides evidence of impact at one or more trophic levels within the ecosystem, can be easily and rather inexpensively obtained (Reynoldson, 1984). Benthic communities are particularly useful for identifying sediment-related stress (IJC, 1987, 1988). Although structural information at the community level is necessary, desirable and readily attained, its specificity is often insufficient to initiate remedial action or produce regulatory change, since cause and effect cannot be easily inferred from structural data alone. Functional tests, which measure physiological or ecological processes, can provide this necessary ancillary information because they are more specific and are capable of discriminating factors responsible for structural changes, as well as quantifying dose-response relationships. Functional tests also tend to be more sensitive than community-based assessment as they largely operate at the level of the individual as compared to the population. They can therefore provide an early warning system.

In the context of this paper the term 'functional biological monitoring' includes various measures of physiological and ecological processes. At the level of the individual these functional processes are frequently measured by toxicity tests. The more widely used tests include acute and chronic measures of toxicity, indicators of biochemical-physiological change, gentoxicity, mutagenicity, carcinogenicity, teratogenicity, reproductive impairment, and simple mortality tests. At the community or population level, such measurements include rates of species colonization or emigration, the rate of re-establishment of equilibrium densities following perturbation (Cairns *et al.*, 1979) or the measurement of species' function (e.g., feeding mechanisms) within a community. Many of these tests can be conducted *in situ*, but the need for strict control over exposure conditions means that most tests are conducted in the laboratory. In addition, most of these tests are performed using single species which may not be indigenous to the area under investigation. Therefore, although functional tests can provide quantifiable relationships between contaminants and organisms and can isolate contributing factors, their lack of environmental realism and failure to incorporate ecosystem complexity limits their investigative utility and possibly their ability to establish actual impairment (Monk, 1983). The obvious solution to this dilemma is to use both structural and functional methods to examine stress.

3. Structural approaches

In the modern era the first critical use of the benthos in ecological assessment was the 'Saprobien system' developed by Kolkwitz & Marsson (1909), and in one way or another the majority of score systems subsequently used are modifications or refinements of this fundamental concept. The Saprobien system assigns scores to individual taxa based on their tolerance to organic pollution. The main variations in most of the subsequent score systems were responses to local conditions and adjustments to the actual scores or their summary values. Only two Saprobien systems are currently in use, in the Netherlands and in W. Germany (Metcalfe, 1989). As originally conceived, it incorporated all trophic levels; however, much of the taxonomy, particularly for the micro-organisms, was, and still remains, unavailable. Furthermore, the pollution sensitivity of many species is unknown and the system suffers from a lack of community integration. The original system was geographically specific and was only appropriate for sewage pollution (Chutter, 1972; Persoone & De Pauw, 1979; Slooff, 1983).

The second major innovation in utilizing benthic invertebrates in ecological assessment was the advent of diversity indices. Wilhm & Dorris (1968) initiated the widespread adoption of the diversity approach, particularly in North America. Diversity indices use three components of community structure, abundance (total numbers), richness (number of species), and evenness (distribution of individuals among species) to produce a single number. The underlying premise of this system is that undisturbed environments will be characterized by high taxa richness and an even distribution of individuals among the species. Stressed environments may respond to sewage pollution by a decrease in richness as sensitive species are lost and tolerant species increase, or a decrease in both richness and abundance in response to toxic wastes. This approach was seen to have several advantages over the saprobic approach. Diversity indices appear to be independent of sample number (Pinder *et al.*, 1987), thus reducing the need for replication. They do not require information on, nor do they make assumptions about, the sensitivities of species. Lastly, the fact that they generate a single index value was seen by some to be advantageous in quickly transferring information to non-biologists (Wilhm & Dorris, 1968; Cook, 1976; Green, 1979; Mason *et al.*, 1985; Pinder *et al.*, 1987). However, the recent decline in the frequency of their use suggests that the inherent disadvantages of using diversity indices outweigh their benefits. These disadvantages are: the reduction of information on individual species to single numbers disregards known data on pollution tolerances of various species and results in a loss of information on the composition of the community; the response of a community to stress is not normally linear, therefore strictly linear comparisons cannot be made; the standards set for the interpretation of indices (Wilhm & Dorris, 1968; U.S. EPA, 1990) consider neither non-linearity or naturally low diversity; and the level of taxonomic detail and other conditions can markedly affect the index value.

The biotic approach combines the benefits of the saprobic system with those of diversity indices into a single index or score that can be statistically analyzed. Numerous biotic indices and scores have been developed for use in Europe (review by Metcalfe, 1989). Most biotic indices trace their origins to the Trent Biotic Index, which was developed for the Trent River in England (Persoone & De Pauw, 1979). The most advanced systems to date are the Biological Monitoring Working Party (BMWP) score system, applied mainly in the U.K., and the Belgian Biotic Index (BBI).

The BMWP score system is based on the sensitivities of key families of invertebrates to pollution (Armitage *et al.*, 1983). All organisms from a site to be scored are identified to the family level and each family is assigned a value according to its pollution tolerance. Values for all families present are summed to arrive at a site score. The BBI (De Pauw & Vanhooren, 1983) also ranks key indicator groups of taxa according to their pollution tolerance. However, the score for a given site corresponds to that for the most sensitive group and considers the number of taxa that are represented. Thus, unlike the BMWP score system, this index does not involve tallying scores for all families.

In the United States, the Invertebrate Community Index (ICI) (Ohio EPA, 1987) and Hilsenhoff's Biotic Index (Hilsenhoff, 1987) are frequently used. The ICI is derived from the Index of Biotic Integrity (IBI), which is based on attributes of fish communities and is described in detail in Karr (1991). The ICI consists of ten structural and functional community metrics, each with four scoring categories (6, 4, 2, and 0 points) that correspond to exceptional, good, fair, and poor community condition. Individual metric scores are summed to generate a site ICI score. The ten metrics used are: total number of taxa, total number of mayfly taxa, total number of caddisfly taxa, total number of dipteran taxa, percent mayfly composition, percent caddisfly composition, percent Tanytarsini midge composition, percent other dipteran and non-insect composition, percent tolerant organisms (oligochaetes and various dipterans and molluscs), and richness of Ephemeroptera, Plecoptera, and Trichoptera (EPT) taxa.

Hilsenhoff's Biotic Index was developed for assessing the response of riffle-dwelling arthropod communities in Wisconsin streams to organic pollution. Tolerance values have been assigned to some 400 species or genera. The index is calculated for a given site as follows: the number of individuals of each taxon is multiplied by the tolerance value for that taxon; the products for all taxa are than summed and divided by the total number of organisms to derive an index value for the site. (Table 1 of Hilsenhoff, 1987 is a guide to water quality based on this index).

Community similarity indices, which basically measure the similarity of the structure of two communities, have mostly been developed for use in terrestrial ecology and have not been extensively applied to aquatic ecosystems (Washington, 1984). Similarity indices can only be used where a clean water station is available for comparison, and have therefore mainly been used for upstream-downstream contrasts of the response to point source pollution. There is some evidence that similarity indices may be more sensitive to subtle changes in community structure than diversity indices (Perkins, 1983; Washington, 1984).

A wide range of similarity indices are currently available, all of which have somewhat different mathematical and ecological properties. Three indices that appear to have been most successfully applied are the Percent Similarity Index (PSC), Pinkham and Pearson's B (Washington, 1984), and the Bray-Curtis index (Hruby, 1987). Formulae for a large selection of indices, including these, are presented in Perkins (1983) and Washington (1984). Because of their differing properties, various indices generate somewhat different results when applied to the same data set. Most require both validation under controlled conditions and careful consideration of their intrinsic properties to determine which index is most suitable to the objectives of a particular study and the type of data available. Because PSC is based on relative abundance whereas Pinkham and Pearson's B is based on actual abundance, the two indices respond differently to certain changes in community structure. Brock (1977) found that sites with very similar proportions of taxa but very different total abundances registered a high degree of similarity according to PSC and less similarity according to B. The removal of a few rare species from a data set for two stations hardly affected the value for PSC, whereas B indicated a definite change. Conversely, B was insensitive to the addition of a substantial number of individuals to a dominant taxon. Brock (1977) concluded that if B overemphasizes shifts in rare taxa and de-emphasizes changes among dominant species, then it is too sensitive to normal sampling errors. It could, however, have an important application where the loss of rare or endangered species is of interest. Hruby (1987) identified the following properties of an index as most desireable: species abundances should be included; the index should use transformed data, which increases the importance of rare species; the index values should not be clumped (i.e. the range of values should not be too narrow); and there should be a linear response to changes in species numbers and abundances. The Bray-Curtis index satisfied most of these conditions.

Most indices and assessment techniques are based on the response of the entire macroinvertebrate benthic community to pollution. It has been demonstrated, however, that certain components or assemblages are more suitable for assessing lakes versus rivers or water versus sediment quality, or for investigating specific types of pollution. Focusing on a reduced assemblage has several advantages. It can simplify the collection, sorting, and identification of benthic samples, thereby reducing time and effort. The use of selective sampling techniques can provide more precise estimates of the diversity and abundance of the organisms under consideration. The resources saved by limiting the investigation to one group of organisms can be redirected into more intensive or extensive studies and species-level taxonomic identification. The latter is an extremely important consideration. Indices that use species-level identification have better potential for site discrimination than those using higher taxonomic levels (Furse *et al.*, 1984; Hilsenhoff, 1988). This is not surprising, considering the findings of Slooff (1983) who subjected various invertebrate species to a variety of chemicals in laboratory toxicity tests and observed that species belonging to the same group often showed as much variability in susceptibility as species from different groups. Both Slooff (1983) and Chapman *et al.* (1982) found that species that were intolerant to organic pollution in general were often very tolerant to specific toxicants. Working at the species level also allows the identification of indicator species for specific types of pollution. Observational data can then be verified by laboratory toxicity tests on these species (e.g. Chapman *et al.*, 1982) in order to establish cause/effect links between specific types of pollution, environmental perturbations, or even specific chemicals, and the presence or absence of key indicator species or species assemblages.

There have been several approaches to focusing on a reduced assemblage, including the identification of characteristic combinations of species that are indicative of certain types of pollution, the determination of ratios of pollution-tolerant to pollution-sensitive organisms, and the identification of key indicator species. The groups that have received the most attention are the oligochaetes and the insects, especially the chironomids and caddisflies.

Chironomid and oligochaete communities have been used for both trophic classification and pollution assessment in large lakes. Saether (1979) identified 15 distinct chironomid communities that were characteristic of a range of trophic conditions in Nearctic and Palearctic lakes. In his opinion, oligochaete communities could not provide as distinct a classification system as chironomids because their environmental requirements are less restricted. However, he noted that an increasing ratio of oligochaetes to chironomids was a good indicator of eutrophication. A trophic classification system based on oligochaete taxa was developed by Mozley & Howmiller [1977; in Lauritsen, *et al.* (1985)] using data from Green Bay, Michigan. Other authors (Lang & Lang-Dobler, 1979; Chapman *et al.*, 1982) have since demonstrated that oligochaete species have a broad range of tolerances to organic pollution and specific chemicals. These authors have shown that oligochaetes have considerable potential for bioassessment.

Rosenberg *et al.* (1986) promoted the use of aquatic insect communities for environmental impact assessment because of their great diversity, importance as food for fish, representation in various functional roles in the ecosystem (predators, herbivores, detritivores, etc.) and because they often respond to environmental perturbations in characteristic ways. He stressed the need for developing taxonomic keys for species-level identifications so that indicator species could be identified and used. At present, it may be possible to describe representative insect communities for broad categories of pollution such as sewage, agricultural run-off, and heavy metals. Dance & Hynes (1980), for example, observed characteristic changes in a variety of insect orders due to the combined stresses of agricultural land use. Once again, the Chironomidae and Trichoptera have received the most attention. For example, Rae (1989) found that several groups of common chironomid genera were indicative of certain chemical conditions in the Scioto River basin in Ohio. He strongly recommended focusing on indicator taxa rather than studying the entire benthic community because little information is gained by examining the distributions of facultative organisms. Winner *et al.* (1980) assessed the response of aquatic insect communities to heavy metal pollution (Cu, Cr, and Zn) in two Ohio streams. Where metal stress was severe and overwhelmed other environmental features, chironomid populations were consistently dominant. Winner *et al.* (1980) therefore proposed that the ratio of chironomids to total insects was a useful index of heavy metal pollution. There appears to be a general consensus that chironomids as a group are excellent indicators of many types of pollution in both lentic and lotic environments.

Trichoptera are also a diverse group occupying a wide variety of habitats and trophic levels. Basaguren & Orive (1990) identified a succession of species from the headwaters to the lowland reaches in a river system in Spain in relation to selected physico-chemical features. In organically polluted river sections, species substitutions occurred that deviated from those in unpolluted sections; Trichoptera were particularly susceptible to low oxygen. The Hydropsychidae have been used to classify water types; deviations from the expected patterns for given areas were attributed to pollution (Higler & Tolkamp, 1983).

The computerization of taxa and environmental data in Great Britain in the early 1980s allowed Armitage *et al.* (1983) to examine the possibility of predicting the expected communities from physical and chemical data. Multiple regression relationships showed strong correlations between Biological Monitoring Working Party (BMWP) score as the dependent variable and various physical and chemical parameters as the independent variables. Wright *et al.* (1984) and other authors have used multivariate techniques to classify unpolluted sites and to predict community type from environmental data (Table 1). In the United Kingdom sites were classified into groups using clustering and ordination algorithms (Wright *et al.*, 1984). Sixteen site groups were selected, based on the taxa lists, from combined spring, summer, and fall sampling at 268 sites on 41 rivers. Correlation coefficients were determined between the ordination scores and environmental variables, thus determining the validity of the selected variables. The highest correlations were observed using substratum characteristics, water chemistry, and variables that related to distance from the source. Using multiple discriminant analysis (MDA) the site groupings were well correlated with many of the 28 environmental

Table 1. The relative success of various authors in predicting benthic invertebrate community structure

Habitat	Sites no.	Communities no.	Predictor variables	Communities predicted (%)	Author(s)
Rivers	286	16	28	76.1	Wright *et al.*, 1984
Streams	79	5	26	70.9	Corkum & Currie, 1987
Rivers	45	6	5	68.9	Ormerod & Edwards, 1987
Rivers	54	5	9	79.6	King, 1981
Lakes	68	7	11	90.0	Johnson & Wiederholm, 1989
Rivers	35	5	6	75.4	Reynoldson & Zarull, 1989

variables. The effects of both season and level of identification on predictive capability of the classification were investigated by Furse *et al.* (1984) to determine whether effort in either area could be reduced. They found that qualitative species-level data led to more reliable classifications and predictions than either quantitative or qualitative family level data, because of the greater number of taxa and the narrower environmental requirements of species. The quality of the prediction was also improved by combining results from three seasons as this ensured that species missing from one sampling event would be captured in another. The least information loss was achieved by maintaining the taxonomic effort and reducing the sampling effort (Table 2). Furse *et al.* (1984) recommended species level identification on spring samples, because animals are larger in spring and thus easier to identify. Moss *et al.* (1987) tested the prediction of benthic invertebrate communities from environmental variables. Almost 90 percent of the species predicted to occur with more than 75 percent probability were found, and almost 80 percent of the $P > 0.50$ predicted species were found. This suggests that the occurrence of the majority of expected species can be predicted from variables that are readily obtained from maps or meteorological stations. The most recent development in the U.K. study (Wright *et al.*, 1988, 1989) has been the adaptation of the reference communities, now from a 370-site data base, and the predictive equations for group occurrence and probability of species occurrence, into software that can be run on a personal computer. The software, called RIVPACS, uses environmental data collected from a site to predict the probability of the site being in a group and the probability of the occurrence of individual taxa for that site based on the environmental data options selected. There are four input options using various environment variables (Moss *et al.*, 1987) and output options can be set to either species or family level. This approach provides a method for the classification of unpolluted aquatic habitats, a technique for establishing target communities and, by comparing predicated with observed communities, for determining the extent of impact.

The approach has also been used in smaller geographic or taxonomic studies (Table 1). Ormerod & Edwards (1987) examined the benthic assemblages within a single drainage basin in South Wales and, using five environmental variables, were able to predict site occurrence with almost 70 percent accuracy. Other authors (King, 1981; Corkum & Currie, 1987; Johnson & Wiederholm, 1989) have used this approach in a

Table 2. Effects of taxonomic and sampling effort on the prediction of benthic invertebrate community structure (from Moss *et al.*, 1987)

Taxonomic effort COMMUNITIES	8 % Predicted	8 % Loss	16 % Predicted	16 % Loss
Species (Qualitative)	79.1		76.1	
Family (Quantitative)	57.5	21.6	53.7	22.4
Family (Qualitative)	69.0	10.1	60.1	16.1
Sampling effort COMMUNITIES	8 % Predicted	8 % Loss	16 % Predicted	16 % Loss
Spring	71.3	7.8	61.9	14.2
Summer	69.4	9.7	63.4	12.7
Fall	75.4	3.7	69.0	7.1
Combined	79.1		76.1	

range of habitats, using different taxonomic groups. Using classification and discriminant analysis, such studies have been able to predict community assemblages, usually with a prediction accuracy of more than 70 percent. Furthermore, from these studies there has been considerable concurrence in the environmental variables that are the most appropriate predictors in running water systems. Similarly, basic physical and chemical variables were good discriminators of the benthic community in lakes in Sweden (Johnson & Wiederholm, 1989). The results of a study on the Detroit River (Reynoldson & Zarull, 1989) provide the only data, using this approach, for a highly polluted habitat. This study showed that contaminants such as hexachlorobenzene and chromium, together with physical characteristics, were important discriminating variables.

4. Functional approaches

Measurements of biological effect from structural information provides only circumstantial evidence of causation even when correlations between biological and environmental data are very strong (Chapman & Long, 1983). The provision of direct proof of toxicity requires the exposure of animals to contaminated waters or sediments (i.e. use of bioassays). The use of benthic invertebrates as bioassay test organisms has been strongly recommended as part of the sediment assessment protocol developed by the International Joint Commission (IJC, 1988). This report identified six major groups of benthic invertebrates for which tests have been developed (Table 3), and suggested that the following factors be considered when selecting a benthic invertebrate for bioassay testing: the origin of the test population, the endpoints being considered, the exposure period, whether or not the species are indigenous and relevant, and the background information available on the test. Giesy & Hoke (1989) reviewed the various organisms and tests available for measuring sediment toxicity in freshwater systems and proposed a number of characteristics for selecting an appropriate test. Of these characteristics we view simplicity, replicability, ecological relevance, and good relationship to field effects to be of primary importance. We have ranked these four characteristics (*sensu* Giesy & Hoke, 1989) and found benthic invertebrates to compare

Table 3. Benthic invertebrate species used as bioassay organisms (from IJC, 1988).

SPECIES	SOURCE	ENDPOINTS	DURATION
AMPHIPODA			
Pontoporeia affinis (Lind.)	Field	Growth	
Hyalella azteca (Sauss.)	Culture	Reproduction	28 d
Gammarus lacustris (Sars)	Field	Mortality	10—28 d
ISOPODA			
Asellus communis Say	Field	Mortality	4 d
INSECTA			
Chironomus tentans (Fabr.)	Culture	Growth	48 h—25 d
C. riparius Meig.	Culture	Growth	24 h
Hexagenia limbata (Serv.)	Field	Growth	5 d
H. rigida (McDunn.)	Field	Growth	6 h—10 wk
ANNELIDA			
Nais communis (Piguet)	Field	Mortality	96 h
Ilyodrilus frantzi (Brink.)	Field	Mortality	96 h
Lumbriculus variegatus (Müll)	Culture	Mortality	10—28 d
Tubifex tubifex (Müll.)	Culture	Reproduction	28 d
NEMATODA			
Panagrellus redivivus	Culture	Growth	96 h
MOLLUSCA			
Lymnaea acuminata	Field	Reproduction	72—96 h

favourably with more frequently used species (Table 4).

Most bioassay tests utilize a laboratory approach with cultured or field-collected organisms. Cairns (1981) and Chapman & Long (1983) have made a plea for *in situ* tests. A recent development has been the technique known as 'scope for growth'. This was developed for use with marine organisms but has been adapted by Naylor *et al.* (1989) for use with *Gammarus pulex* in fresh waters. The technique involves monitoring the energy budget of individuals and testing how the various components are affected by stress, where:

$C - F - R = \text{SfG}$.
C = energy consumed as food;
F = energy lost as faeces;
R = energy metabolized (measured as respiration).

Scope for Growth (SfG) can be positive (energy is available for growth and reproduction), zero (energy input = output) or negative (animals are using reserves for essential metabolism). Although growth and fecundity could be measured directly, the advantage of SfG is that it can be measured rapidly (hours or days) and shows which component of the energy budget is affected by stress. Naylor *et al.* (1989) have shown the test to be an order of magnitude more sensitive than acute toxicity tests for determining toxic effects of zinc and pH.

A considerable amount of work has been made on the occurrence of deformities in benthic invertebrates. The majority of published material is that of Warwick (1980, 1985, 1990) on deformities in chironomid antennae and mouthparts. Some information is also available on the occurrence of deformities in oligochaete setae (Milbrink, 1983). This work is still very much in the research phase, and cannot be widely used at this time.

Most other physiological assays are still in the developmental stage. These include measuring various processes such as ATP incorporation, MFO induction, and thymidine incorporation as indicators of environmental stress. Although these techniques may well be more sensitive than whole organism tests, their ecological relevance is unknown at present.

According to the River Continuum Concept (Vannote *et al.*, 1980), stream morphology, current velocity, substrate composition, temperature, and allochthonous vs. autochthonous food sources all interact to influence food availability to inverte-

Table 4. Relative ranking of some sediment bioassays (adapted from Giesy & Hoke, 1989) 1 = poor, 2 = fair, 3 = good, 4 = excellent. A total score of 13—16 indicates the bioassay has high utility.

	Simplicity	Replicability	Ecological relevance	Relationship with field	Total
Polytox	4	2	1	1	8
Microtox	4	4	1	1	10
Protozoan	2	2	4	2	10
Algal flask	3	3	2	2	10
Algal Microplate	3	3	2	2	10
Algal flow	2	3	2	2	9
Algal ETS	4	2	2	2	10
Oligochaete	3	2	4	4	13
Tubifex	4	3	4	4	15
Mollusc	2	2	4	4	12
Amphipod	2	2	4	4	12
Hexagenia	1	2	4	4	11
Chironomus	2	3	4	4	13
Daphnia	3	4	3	3	13
Cerodaphnia	2	3	3	3	11
Pimephales	2	3	3	3	11

brates. These interactions vary systematically with stream order, thereby regulating distribution patterns of invertebrate functional feeding groups (Hawkins & Sedell, 1981). The feeding groups, which are classified on the basis of morphological-behavioral mechanisms of food acquisition rather than taxonomic or trophic designations, are referred to as the scrapers, collector-filterers, collector-gatherers, predators, and shredders (Cummins, 1988). Under unperturbed conditions, headwater areas are normally dominated by shredders and collectors due to the mainly allochthonous energy source, mid-portions of streams will be autotrophically driven and dominated by scrapers and collector-filterers, and communities downstream in large rivers will be mainly composed of collector-gatherers due to the accumulation of fine particles from upstream (Rabeni *et al.*, 1985).

It has been suggested that a water quality classification system based on functional feeding groups of aquatic invertebrates may be superior to a system based on community structure because it reflects 'more ecologically significant attributes of streams and rivers' (Rabeni *et al.*, 1985). Such a system would also reduce the variability associated with taxonomic complexity (Hawkins & Sedell, 1981), allowing easier recognition of trends and understanding of ecosystem function. We would suggest that this approach is also more universally applicable because local taxonomic differences do not seriously affect a system based on functional feeding groups. In contrast, biotic indices require considerable modification before they can be used in different geographical areas. The main drawback to the functional feeding group approach is that it is based on nutrient dynamics and can therefore only be used to assess the effects of organic enrichment, not toxic chemicals. Also, it does not eliminate the need for detailed taxonomy because organisms must be correctly identified before they can be assigned to their appropriate group.

In a study of the Penobscot River in Maine, Rabeni *et al.* (1985) demonstrated that functional feeding groups responded to organic pollution in a manner predicted by the River Continuum Concept, despite the fact that there was no significant longitudinal gradient. This suggests that pollution can 'reset' the normal sequence of shifts in feeding groups. Once pollution had been reduced, functional analyses showed that the system had begun to shift back from a heterotrophic to its normal autotrophic condition. Olive *et al.* (1988) assessed the water quality of the upper Cuyahoga River in Ohio by means of a composite index of community diversity, biotic and functional group-based measures. The functional index was the ratio of scrapers to detritivores. In the presence of organic pollution, they expected scrapers to be reduced and detritivores to be encouraged. However, they found that the different indices did not always agree in the classification of certain sites because of the influence of habitat. These findings caution us against using indices of community response interchangeably and against relying on only one type of index for assessing the effects of pollution.

The last broad use of benthic invertebrates has been the assessment of contaminant bioavailability through bioaccumulation studies. Although bioaccumulation is a functional response of aquatic organisms to the presence of toxic chemicals in their environment it is not *per se* a health assessment parameter. Organisms can often accumulate high concentrations of certain contaminants without any adverse health effects. Nevertheless, the chemical analysis of aquatic organisms to determine concentrations of pollutants in their tissues, which is defined as 'bioanalysis' by Butler *et al.* (1985), can be used in several ways to provide information on both the environment occurrence and hazard of toxic chemicals.

Recently, many agencies responsible for monitoring water quality have incorporated bioanalysis components into their routine water and sediment quality monitoring programs for two major purposes: to improve on their ability to determine the presence, levels, and distribution of contaminants in the environment, and to provide information on the bioavailability, persistence, and potential for food chain transport of toxic chemicals. The latter properties are important aspects of risk assessment. Benthic invertebrates are particularly suitable for both purposes because they are sedentary (thereby reflecting local conditions), they live in the sediment (often a major source of contaminants), and they are important fish food items and, therefore, are a major vector for transfer of contaminants up the food chain.

In order to understand the interactions between benthic invertebrates and toxic chemicals in their environment, several processes must be examined: exposure (levels of contaminants in water, sediment, and food), bioaccumulation (contaminant residues in the organisms themselves) and toxicity (both instream assessment of the response of benthic communities to contaminants and bioassays on effluents or contaminated sediments). To date, the links between exposure and bioaccumulation and between exposure and toxic effects have been explored. However, few studies have addressed the relationship between bioaccumulation and toxic effects. At present, then, bioaccumulation data can only be used as indirect evidence of the impact of contaminants on biota. Ultimately, the most powerful biomonitoring and bioassessment tools will be those which combine a measure of contamination in the environment and in the organisms with a measure of effects, thus establishing cause/effect or dose/response relationships. Bioanalysis will become a more meaningful activity when a 'critical' threshold concentration of a contaminant in an organism can be associated with a specific toxic response.

5. An integrated approach

Cairns & van der Schalie (1980), in their extensive review of biological monitoring criticized the fact that assessment of aquatic effects was based almost exclusively on direct observation. Such observations require little basis in theoretical constructs of how systems operate and do not allow extrapolation to other ecosystems.

One of the first proposals combining structural and functional methods in aquatic ecosystem assessment was the sediment quality triad (Chapman & Long, 1983; Long & Chapman, 1985). This approach included coincident measurements of toxicity, community structure, and physical and chemical variables. More recently, this approach has been modified and formally presented for use in sediment assessment in the Great Lakes (IJC, 1987, 1988; Reynoldson & Zarull, 1989). The IJC approach to assessment relies heavily on benthic invertebrates because its primary concern is the determination of sediment quality. The functional tests recommended by Chapman & Long (1983) and the IJC (1987, 1988) rely on single species assays; few formally accepted protocols are available for bioassays using benthic invertebrates. Only the *Chironomus tentans* (Fabr.) and *Hexagenia limbata* (Serv.) assays were recommended for standard use (IJC, 1988) but routine tests using *Chironomus riparius* (Meig.), *Hyalella azteca* (Sauss.), *Tubifex tubifex* (Müll.), and *Lumbriculus variegatus* (Müll.) may soon be available. However, the relative sensitivity of these various tests is presently unknown (Calow, 1989) and the results, relative to *in situ* effects, have not been examined as far as we are aware.

One of the most serious objections to incorporating benthic community assessment data into water management policy has been the definition of reference communities. This definition could provide community based criteria against which observational data could be matched to determine the extent of impact. There has been a general recognition of the need for the identification and characterization of reference sites from unimpaired locations to define attainable water quality objectives. Because of the great natural variation among aquatic ecosystems, it is unlikely that community based criteria could be universally applied, yet, site-specific criteria are too expensive and cumbersome to use. In the United States an 'ecoregion' approach is being developed. These ecoregions are areas defined by similarities in geography, hydrology, land use, or other ecologically relevant variables (U.S. EPA, 1990). Sites within a given ecoregion should have similar ecosystems, and 'attainable quality is then based on assessment of conditions in minimally impacted reference sites that characterize the region' (Hughes & Larsen, 1988). Biological criteria based on structural and functional attributes or resident fish and benthic invertebrate communities are currently being developed by the U.S. EPA and integrated with chemical-specific criteria and whole-effluent toxicity evaluations to achieve a more holistic approach to water management. The focus is presently on streams and rivers, but lakes, wetlands, and coastal areas have not been excluded.

In the United States, individual states are required to develop narrative biological criteria during the next three years, to be followed by the development of numerical biocriteria. Biocriteria will be used to improve water quality standards, identify impairment of beneficial uses, assist in setting program priorities, and detect problems which might otherwise be missed or underestimated. Biocriteria are especially suited for the detection of non-point source, cumulative and episodic pollution, and physical changes such as habitat deterioration, none of which would be detected using traditional chemical assessment methods. Biocriteria are also generally more sensitive than chemical criteria. In Ohio, for example, an evaluation of instream biota indicated that 36 percent of impaired sites were not identified using chemical-based criteria (Ohio EPA, 1987).

Twenty states now use some form of biological assessment; some of these (e.g. Ohio) have already instituted numerical biocriteria for classifying sites and assessing water quality. Various states apply biocriteria somewhat differently. For example, Florida has a legal criterion based on macroinvertebrate species diversity whereby diversity may not fall below 75 percent of established background values. Maine uses macroinvertebrate community data to assess attainment or non-attainment of standards for designated water uses (e.g. drinking water supply; fish and wildlife habitat; recreational, agricultural, and industrial use). Oklahoma uses biotic information to assess long-term trends in water quality.

The recent developments in using multivariate techniques show promise in establishing biocriteria. There appears to be general agreement within the European community (Metcalfe, 1989) on the need to define reference communities based on chemical, physical, geological, and geographical features unrelated to pollution as a first step toward identification of the best achievable communities for each type of water. These communities could then be used as water quality objectives, expressed either in terms of whole communities or important discriminating species. Although the European efforts are primarily directed toward running water systems there is no reason why this strategy should not be applicable to lakes. The task is expensive because it requires the collection of biological, chemical, and physical data from large numbers of unpolluted sites. Nevertheless, such an effort has begun on the Canadian Great Lakes as part of a project to develop biological sediment criteria based on benthic communities and toxicity responses in selected benthic invertebrates. The project proposes to use both toxicity response categories and community composition as grouping variables, and establish relationships between these and physical and chemical variables for a set of reference communities. The expected toxicity test responses and community structure will be established in a range of habitats and used to compare predicted to observed responses. This combination of a structural and functional approach should expand the utility of benthic invertebrates in aquatic ecosystem health assessment.

6. Acknowledgements

We wish to thank Dr P. Ross and particularly Dr D. M. Rosenberg for their constructive comments on this paper.

References

Armitage, P. D., D. Moss, J. F. Wright & M. T. Furse, 1983. The performance of a new biological water quality score system based on macroinvertebrates over a wide range of unpolluted running-water sites. Water Res. 17: 333—347.

Basaguren, A. & E. Orive, 1990. The relationship between water quality and caddisfly assemblage structure in fast-running rivers. The River Cadagua basin. Environ. Monit. Assess. 15: 35—48.

Brock, D. A., 1977. Comparison of community similarity indexes. J. Water Pollut. Control Fed. 49: 2488—2494.

Butler, G. C., B. G. Bennett, D. R. Miller, J. K. Piotrowski & A. Sors, 1985. Methods for estimating exposure to chemicals. In: V.B. Vouk, G.C. Butler, D.G. Hoel & D.B. Peakall (eds), *Methods for Estimating Risk of Chemical Injury: Human and Non-human Biota and Ecosystems*. pp. 7—28. John Wiley & Sons, New York.

Cairns, J. Jr, 1981. Biological monitoring. Part VI — Future needs. Water Res. 15: 941—952.

Cairns, J. Jr & W. H. van der Schalie, 1980. Biological monitoring. Part I — Early warning systems. Water Res. 14: 1179—1196.

Cairns, J. Jr, D. L. Kuhn & J. L. Plafkin, 1979. Protozoan Colonization of Artificial Substrates. In: R.L. Wetzel (ed.), *Methods and Measurements of Attached Microcommunities: A Review*. Amer. Soc. Test. Materials, Philadelphia, PA.

Calow, P., 1989. The choice and implementation of environmental bioassays. Hydrobiologia 188/189: 61—64.

Chapman, P. M. & E. R. Long, 1983. The use of bioassays as part of a comprehensive approach to marine pollution assessment. Mar. Pollut. Bull. 14: 81—84.

Chapman, P. M., M. A. Farrell & R. O. Brinkhurst, 1982.

Relative tolerances of selected aquatic oligochaetes to individual pollutants and environmental factors. Aquat. Toxicol. 2: 47—67.
Chutter, F. M., 1972. An empirical biotic index of the quality of water in South African streams and rivers. Water Res. 6: 19—30.
Cook, S. E. K., 1976. Quest for an index of community structure sensitive to water pollution. Environ. Pollut. 11: 269—288.
Corkum, L. D. & D. C. Currie, 1987. Distributional patterns of immature Simuliidae (Diptera) in northwestern North America. Freshw. Biol. 17: 201—221.
Cummins, K. W., 1988. The study of stream ecosystems: A functional view. In: L. R. Pomeroy & J. J. Alberts (eds), *Concepts of Ecosystem Ecology*. pp. 247—262. Springer-Verlag, New York.
Dance, K. W. & H. B. N. Hynes, 1980. Some effects of agricultural land use on stream insect communities. Environ. Pollut. 22: 19—28.
De Pauw, N. & G. Vanhooren, 1983. Method for biological quality assessment of watercourses in Belgium. Hydrobiologia 100: 153—168.
Furse, M. T., D. Moss, J. F. Wright & P. D. Armitage, 1984. The influence of seasonal and taxonomic factors on the ordination and classification of running-water sites in Great Britain and on the prediction of their macroinvertebrate communities. Freshw. Biol. 14: 257—280.
Giesy, J. P. & R. A. Hoke, 1989. Freshwater sediment toxicity bioassessment: rationale for species selection and test design. J. Great Lakes Res. 15: 539—569.
Green, R. H., 1979. *Sampling Design and Statistical Methods for Environmental Biologists*. John Wiley and Sons Ltd., New York, 257 pp.
Hawkins, C. P. & J. R. Sedell, 1981. Longitudinal and seasonal changes in functional organization of macroinvertebrate communities in four Oregon streams. Ecology 62: 387—397.
Higler, L. W. G. & H. H. Tolkamp, 1983. Hydropsychidae as bio-indicators. Environ. Monit. Assess. 3: 331—341.
Hilsenhoff, W. L., 1987. An improved biotic index of organic stream pollution. Great Lakes Entomol. 20: 31—39.
Hilsenhoff, W. L., 1988. Rapid field assessment of organic pollution with a family-level biotic index. J. North Amer. Bentholog. Soc. 7: 65—68.
Hruby, T., 1987. Using similarity measures in benthic impact assessments. Environ. Monit. Assess. 8: 163—180.
Hughes, R. M. & D. P. Larsen, 1988. Ecoregions: An approach to surface water protection. J. Water Pollut. Control Fed. 60: 486—493.
Hynes, H.B. N., 1960. *The Biology of Polluted Waters*. Liverpool Univ. Press, Liverpool, 202 pp.
IJC (International Joint Commission), 1987. Guidance on Characterization of Toxic Substances Problems in Areas of Concern in the Great Lakes Basin. Rep. Surveillance Work Group, Windsor, Ontario, 177 pp.
IJC, 1988. Procedures for the Assessment of Contaminated Sediment Problems in the Great Lakes. Rep. Sediment Sub-Committee, Windsor, Ontario, 140 pp.
Johnson, R. K. & T. Wiederholm, 1989. Classification and ordination of profundal macroinvertebrate communities in nutrient poor, oligo-mesohumic lakes in relation to environmental data. Freshw. Biol. 21: 375—386.
Karr, J. R., 1991. Biological integrity: A long-neglected aspect of water resource management. Ecolog. Applic. 1: 66—84.
King, J. M., 1981. The distribution of invertebrate communities in a small South African river. Hydrobiologia 83: 43—65.
Kolkwitz, R. & M. Marsson, 1909. Ökologie der tierische Saprobien. Beiträge zur Lehre von der biologische Gewässerbeurteilung. Internat. Rev. Hydrobiol. 2: 126—152.
Lang, C. & B. Lang-Dobler, 1979. The chemical environment of tubificid and lumbriculid worms according to the pollution level of the sediment. Hydrobiologia 65: 273—282.
Lauritsen, D. D., S. C. Mozley & D. S. White, 1985. Distribution of oligochaetes in Lake Michigan and comments on their use as indices of pollution. J. Great Lakes Res. 11: 67—76.
Long, E. R. & P. M Chapman, 1985. A sediment quality triad: Measures of sediment contamination, toxicity and infaunal community composition in Puget Sound. Mar. Pollut. Bull. 16: 405—415.
Mason, W. T. Jr, P. A. Lewis & C. I. Weber, 1985. An evaluation of benthic macroinvertebrate biomass methodology. Environ. Monit. Assess. 5: 399—422.
Metcalfe, J. L., 1989. Biological water quality assessment of running waters based on macroinvertebrate communities: History and present status in Europe. Environ. Pollut. 60: 101—139.
Milbrink, G., 1983. Characteristic deformities in tubificid oligochaetes inhabiting polluted bays of Lake Vanern, southern Sweden. Hydrobiologia 106: 169—184.
Monk, D. C., 1983. The uses and abuses of ecotoxicology. Mar. Pollut. Bull. 14: 284—288.
Moss, D., M. T. Furse, J. F. Wright & P. D. Armitage, 1987. The prediction of the macro-invertebrate fauna of unpolluted running-water sites in Great Britain using environmental data. Freshw. Biol. 17: 41—52.
Naylor, C., L. Maltby & P. Calow, 1989. Scope for growth in *Gammarus pulex*, a freshwater benthic detritivore. Hydrobiologia 188/189: 517—523.
Ohio EPA (Ohio Environmental Protection Agency), 1987. Biological criteria for the protection of aquatic life: Volume II. Users manual for biological field assessment of Ohio surface waters. Division of Water Quality Monitoring and Assessment, Surface Water Section, Columbus, Ohio, 44 pp.
Olive, J. H., J. L. Jackson, J. Bass, L. Holland & T. Savisky, 1988. Benthic macroinvertebrates as indexes of water quality in the upper Cuyahoga River. Ohio J. Sci. 88: 91—98.
Ormerod, S. J. & R. W. Edwards, 1987. The ordination and classification of macroinvertebrate assemblages in the catchment of the River Wye in relation to environmental factors. Freshw. Biol. 17: 533—546.
Patrick, R., 1949. A proposed biological measure of stream conditions, based on a survey of the Conestoga Basin, Lancaster County, Pennsylvania. Proc. Acad. Natur. Sci. Philadelphia 101: 277—342.

Patrick, R., 1951. A proposed biological measure of stream conditions. Verh. Internat. Verein. Limnol. 11: 299—307.
Perkins, J. L., 1983. Bioassay evaluation of diversity and community comparison indexes. J. Water Pollut. Control Fed. 55: 522—530.
Persoone, G. & N. De Pauw, 1979. Systems of biological indicators for water quality assessment. In: O. Ravera (ed.), *Biological Aspects of Freshwater Pollution*. pp. 39—75. Pergamon Press, Oxford.
Pinder, L. C. V., M. Ladle, T. Gledhill, J. A. B. Bass & A. M. Matthews, 1987. Biological surveillance of water quality — 1. A comparison of macro-invertebrate surveillance methods in relation to assessment of water quality, in a chalk stream. Arch. Hydrobiol. 109: 207—226.
Rabeni, C. F., S. P. Davies & K. E. Gibbs, 1985. Benthic invertebrate response to pollution abatement: Structural changes and functional implications. Water Resources Bull. 21: 489—497.
Rae, J. G., 1989. Chironomid midges as indicators of organic pollution in the Scioto River Basin, Ohio. Ohio J. Sci. 89: 5—9.
Reynoldson, T .B., 1984. The utility of benthic invertebrates in water quality monitoring. Water Qual. Bull. 10: 21—28.
Reynoldson, T. B. & M. A. Zarull. 1989. The biological assessment of contaminated sediments — the Detroit River example. Hydrobiologia 188/189: 463—476.
Rosenberg, D. M., H. V. Danks & D. M. Lehmkuhl, 1986. Importance of insects in environmental impact assessment. Environ. Managem. 10: 773—783.
Saether, O. A., 1979. Chironomid communities as water quality indicators. Holarct. Ecol. 2: 65—74.
Sheldon, A. L., 1984. Cost and precision in a stream sampling program. Hydrobiologia 3: 147—152.
Slooff, W., 1983. Benthic macroinvertebrates and water quality assessment: Some toxicological considerations. Aquat. Toxicol. 4: 73—82.
U.S. EPA (United States Environmental Protection Agency), 1990. Biological Criteria: National Program Guidance for Surface Waters. EPA-440/5-90-004, 57 pp.
Vannote, R. L., G. W. Minshall, K. W. Cummins, J. R. Sedell & C. E. Cushing, 1980. The River Continuum Concept. Can. J. Fish. Aquat. Sci. 37: 130—137.
Warren, C. E., 1971. *Biology and Water Pollution Control.* W.B. Saunders Co., Philadelphia, 434 pp.
Warwick, W. F., 1980. Palaeolimnology of the Bay of Quinte, Lake Ontario: 2800 years of cultural influence. Can. Bull. Fish. Aquat. Sci. 206.
Warwick, W. F., 1985. Morphological abnormalities in Chironomidae (Diptera) larvae as measures of toxic stress in freshwater ecosystems: indexing antennal deformities in *Chironomus* Meigen. Can. J. Fish. Aquat. Sci. 42: 1881—1914.
Warwick, W. F., 1990. Morphological deformities in Chironomidae (Diptera) larvae from the Lac St Louis and Laprairie basins of the St. Lawrence River. J. Great Lakes Res. 16: 185—208.
Washington, H. G., 1984. Diversity, biotic and similarity indices. A review with special relevance to aquatic ecosystems. Water Res. 18: 653—694.
Wilhm, J. L. & T. C. Dorris, 1968. Biological parameters for water quality criteria. BioScience 18: 477—481.
Winner, R. W., M. W. Boesel & M. P. Farrell, 1980. Insect community structure as an index of heavy-metal pollution in lotic ecosystems. Can. J. Fish. Aquat. Sci. 37: 647—655.
Wright, J. F., D. Moss, P. D. Armitage & M. T. Furse, 1984. A preliminary classification of running-water sites in Great Britain based on macro-invertebrate species and the prediction of community type using environmental data. Freshw. Biol. 14: 221—256.
Wright, J. F., P. D. Armitage, M. T. Furse & D. Moss, 1988. A new approach to the biological surveillance of river quality using macroinvertebrates. Verh. Internat. Verein. Limnol. 23: 1548—1552.
Wright, J. F., P. D. Armitage, M. T. Furse & D. Moss, 1989. Prediction of invertebrate communities using stream measurements. Regulated Rivers: Research & Management 4: 147—155.