Chapter 10: Biocriteria Implementation

10.1 Characteristics of Effective Biocriteria

Development of narrative or numeric biocriteria depends on the premise that biota provide a sensitive screening tool for measuring the condition of a water resource. Properly defined biocriteria can be used to protect the biological integrity of waterbodies and establish aquatic life use classifications.

Following the development of biocriteria, sites are evaluated to determine how well they meet the biocriteria or whether they have been significantly degraded. This determination is made by comparing the aquatic biota at potentially disturbed sites to the biocriteria, which are in turn based on minimally impaired reference conditions. The greater the anthropogenic impact in a watershed, the greater the impairment of the water resource. A corollary is that drainage basins not subject to anthropogenic impacts contain natural communities of aquatic organisms that reflect unimpaired conditions. These assumptions provide the scientific basis for formulating hypotheses about impairments - departures from the natural condition that result from human disturbances.

The establishment of formal biocriteria warrants careful consideration of planning, management, and regulatory goals. Effective biocriteria function to:

- Provide for scientifically sound evaluations.
- Protect the most sensitive biological value.
- Support and strive for protection of chemical, physical, and biological integrity.

Generally, optimal biocriteria share several common characteristics:

- They include specific assemblage characteristics required for attainment of designated use.
- They are clearly written and easily understood.
- They adhere to the philosophy and policy of antidegradation of water resource quality.
- They are defensible in a court of law.

In addition, biocriteria should be written to consider the best attainable condition at a site. Overly stringent criteria that are unlikely to be achieved serve little purpose. Similarly, biocriteria that support a degraded biological condition defeat the intent of the Clean Water Act. Well-designed biocriteria are set at levels sensitive to anthropogenic impacts; they are not set so high that sites that have reached their full potential are considered in nonattainment or so low that unacceptably impaired sites are scored as meeting the criteria. It will be difficult to determine the full potential of a given lake. Balanced biocriteria will allow multiple uses to be considered so that any conflicting uses are evaluated at the outset. The best balance is achieved by developing biocriteria that closely represent the natural biota, protect against further degradation, and stimulate restoration of degraded sites.

Several kinds of biocriteria are possible, and both narrative and numeric biocriteria have been effectively implemented. Narrative biocriteria consist of statements such as “aquatic life as it naturally occurs” or “changes in species composition may occur, but structure and function of the aquatic community must be maintained.” Numeric values, such as measurements of community structure and function, can also serve as biocriteria as such or as quantitative refinements of narrative biocriteria. To account for a measure’s natural variability in a healthy environment, the numeric criterion should be a defined range rather than a single number. Numeric criteria may also combine several such values in an index. Regardless of which kind is chosen, biocriteria should be both quantitatively based and supported by effective implementation guidelines and adequate capabilities including people, resources, methods, historical data, and management support. Additional general guidance regarding the writing of biocriteria is provided in EPA 440/5-90-004 (1990a) and EPA 822-B-92-002 (1992e).

10.2 Steps To Implementation

The first phase in a biocriteria program is the development of narrative biological criteria (USEPA 1992e). These criteria are essentially statements incorporated into water laws and regulations to formally consider the fate and status of aquatic biological communities. These statements of intent should include the following objectives:

1. Support the goals of the Clean Water Act to provide for the protection and propagation of fish, shellfish, and wildlife, and to restore and maintain the chemical, physical, and biological integrity of the Nation’s waters.
2. Protect the most natural biological community possible by emphasizing the protection of its most sensitive components.
3. Refer to specific community characteristics that must be present for the waterbody to meet a particular designated use; for example, natural diverse systems with their respective communities or taxa indicated.
4. Include measures of community characteristics, based on sound scientific principles, that are quantifiable and written to protect or enhance the designated use.
5. In no case should impacts degrading existing uses or the biological integrity of the waters be authorized.

The use of multiple measures, or metrics, to develop biocriteria is a systematic process involving discrete steps. The process includes site classification (Chapter 4), a biological survey, evaluation of metrics with aggregation into indices (where indicated), formulation of biocriteria, and monitoring and assessment. The conceptual model for processing biological data into a biocriteria framework is adapted from EPA 822-B-96-001(USEPA 1996a) and summarized in Table 10-1. The process is as follows:

Table 10-1. Sequential progression of the biocriteria process.

Step 1	Preliminary Classification to Determine Reference Conditions and Regional Ecological Expectations - Resource classification - Determination of best representative sites (reference sites representative of class categories)
Step 2	Characterization of Reference Condition - Historical data - Survey of reference sites and selected test sites. - Applicable models - Expert consensus
Step 3	Final Classification - Test preliminary classification - Revise if necessary
Step 4	Metric Evaluation and Index Development - Data analysis (data summaries) - Testing and validation of metrics by resource class - Evaluation of metrics for effectiveness in detecting impairment - Aggregation of metrics into index - Selection of biological endpoints
Step 5	Biocriteria Development - Adjustment by physical and chemical covariates - Adjustment by designated aquatic life use
Step 6	Implementation of Monitoring and Assessment Program - Determination of temporal variability of reference sites - Identification of problems
Step 7	Protective or Remedial Management Action Initiate - Programs to preserve exceptional waters - Implement management practices to restore the biota of degraded waters and to identify and address the causes of this degradation
Step 8	Continual Monitoring and Periodic Review of References and Criteria - Biological surveys continue to assess efficiency of management efforts - Evaluate potential changes in reference condition and adjust biocriteria as management is accomplished

Step 1: Preliminary Classification of the Resource - The first decision that a resource agency must make is to determine the resource classes to which biocriteria will apply. Successful classification will result in less variation within a class, leading to more refined characterization of the reference condition and, therefore, to criteria with better resolution in detecting impairment. The preliminary classification should be based on lake characteristics that are not subject to pollution or disturbance, such as size, depth, morphology, or characteristics of the lake watershed.

Multijurisdictional collaboration is encouraged so that common methods and metrics can be established among states or other monitoring entities, and common reference conditions for multijurisdictional ecoregions can be characterized.

A set of reference sites are selected for each resource class; the reference sites are those least impacted by human influence, and they are characteristic of the resource class.

Step 2: Biological Survey - To determine the discriminatory power of the metrics within a lake class, the best-quality sites available, as well as those known to be impaired, are surveyed for biota and physical habitat. The use of standardized field collection methods allows a better interpretation of the raw data than does the use of a conglomeration of techniques.

Step 3: Final Classification - The preliminary classification is tested with biological data to determine whether it is reflected in the biota. If necessary, the classification is revised.

Any characterization of a reference condition should allow for the variability in biological data by using measures of central tendency and variability. Statewide or broader characterization of reference condition can be expected to exhibit high variance. The goal of classification is to minimize variability within classes by allowing the variability to be attributed to differences among classes.

Step 4: Metric Evaluation and Index Development - Potential metrics that have ecological relevance are identified in this step. Metrics are then evaluated for the ability to differentiate between impaired and nonimpaired sites. Values from various scales of measurement are transformed to scores, which are normally incorporated into an index, such as an Index of Biological Integrity (IBI) or an invertebrate index, which in turn becomes part of the final assessment. Metrics may also be used individually as indicators of biological condition in the overall assessment.

Step 5: Biocriteria Development - Biocriteria may be based on an aggregated index, or established for several biological metrics and adjusted by aquatic life uses. The component information and data should always be retained so future indexes or improvements in initial indexes can be calibrated with the data, and continuity of information preserved over time.

For example, a biocriterion for “Class A” lakes might be “a biotic index greater than the 25th percentile of least-impacted reference conditions.” A “Class A” lake would be rated impaired if its biotic index fell below the 25th percentile of reference condition.

Outline of Evaluation Criteria for Bioassessment Programs.

1. *Development of quality assurance and quality control bioassessment program plans.*
2. *Careful preparation of data quality objectives (DQOs) and design of field and laboratory studies to ensure the collection of representative data that will enable the biologists to achieve the objective of their program.*
3. *Preparation of standard operating procedures (SOPs) for field and laboratory methods.*
4. *Staff with adequate training and experience; division of labor within the program that permits specialization.*
5. *Use of approved methodology, use of technically defensible methodology if approved methodology is not available.*
    1. *Sample collection*
    2. *Sample processing*
    3. *Organism identification*
    4. *Counting*
    5. *Biomass measurements*
    6. *Data analysis and interpretation*
6. *Adequate space and physical facilities.*
7. *Adequate state-of-the-art field equipment, laboratory*

instrumentation, and supplies.

8. *Adequate safety procedures.*
9. *Use of replication in sample collection and analysis to determine the precision.*
10. *Frequent calibration of field and laboratory instruments; log book documentation.*
11. *Chain-of-custody procedures for proper sample identification, handling, and logging to prevent misidentification and intermixing of samples.*
12. *Development and use of a taxonomic reference library for identifying specimens to the lowest possible taxonomic level.*
13. *Development and use of a reference specimen collection and use of outside experts to solve difficult problems in specimen identification.*
14. *Careful editing of data before they are placed in a computer file or used in reports.*
15. *Use of appropriate statistical analyses and other methods of data evaluation and interpretation.*

Step 6: Implementation of the Monitoring and Assessment Program - Use of biocriteria requires an operational monitoring and assessment program for two primary reasons: assessment of potentially impaired test sites and continued monitoring of selected reference sites to determine seasonal and annual variability and trends. A biocriteria program is the basis for a representative sampling program to determine statewide status and trends of the resource. The resources required to initiate a monitoring and assessment program are presented in the text box entitled “Outline of Evaluation Criteria for Bioassessment Programs.”

Step 7: Protective and Remedial Management Action - The purpose of the entire process is to improve the water resource quality. Where problems have been identified through this effort, land use changes, discharges, abatements, and in-lake use adjustments are part of the management response. This may be done to improve degraded lakes or reservoirs or to protect exceptionally good ones from future damage. It should be recognized that implementing management action is potentially a multi-year process.

Step 8: Continual Monitoring and Periodic Reviews - The biocriteria-biomonitoring effort is designed to be a continuing process. Progress is expected but failures must be documented so monitoring and management efforts can be improved. The process progressively improves water resources by cycling back through the sequence.

10.3 Technical Considerations

The technical design of a biocriteria program affects the program’s total cost. The sampling and analysis effort and data storage are two major cost elements of a biocriteria program. An optimal design balances the information needs of the monitoring agency with the cost of obtaining the information.

10.3.1 Taxonomic Level

Assemblages in Tiers 2A and 2B are identified to the lowest practical taxonomic level. Species level identification can be time consuming, especially for phytoplankton and benthic macroinvertebrates, and identification to family or genus might be more cost-effective.

10.3.2 Subsampling

Consistency of sampling methods and effort is critical in bioassessment. A sample is usually subsampled, in a random manner, to obtain a reasonable number of organisms for identification and enumeration (typically 100 to 500). Using fewer than 100 organisms might yield unreliable results, whereas using more than 500 is not cost-effective.

Taxonomic richness metrics, such as total taxa, diversity indices, and number of orders are sensitive to sample size. These values increase asymptotically with subsample size up to 500 organisms. Percent composition metrics (e.g., feeding groups, higher tax metrics) are less sensitive to subsample size; that is, the precision of an estimate for percent composition does not improve with subsamples greater than 100 organisms.

In order to control for the effects of sample size, it is critical that the methods are consistent in the number of organisms identified. For example, if the target subsample is 100 organisms, then subsamples smaller than 80 organisms should be rejected and subsamples larger than 120 organisms should be reduced mathematically by rarefaction (Hurlbert 1971) to make them comparable.

10.3.3 Spatial Variability and Replication

Replicating field samples by repeated measurements at a site is integral to biological surveys. These analyses have typically tested for significant differences between upstream and downstream pairs of sites. Significant differences were inferred to be due to discharges. However, Hurlbert (1984) pointed out that treatment of multiple measurements as replicates to infer cause is incorrect use of statistical inference. He pointed out that the site is the sampling unit, and repeated measurement of a sampling unit is not replication. True replication is achieved by replicating independent sampling units.

Repeated measurements, however, do have benefits, which must be weighed against the cost. Repeated measurements are used to estimate measurement error, which is variability among measurements at the same site. Measurement error is due to spatial and temporal variability within a lake as well as actual errors made in sampling and analysis. It may be necessary to determine whether the measurement methodology adequately characterizes the site, and to determine the precision of metrics and indices (Fore et al. 1994). If measurement error is too large, it may be reduced by repeated measurements at a site or by a change of methodology (sample more microhabitats for a larger composite sample; increase subsample size). If the measurement error is acceptable, it is necessary only to take repeated measurements (replicates) for quality assurance at randomly selected sites (typically 10 percent of all sites). The QA replicates are then used to estimate measurement error.

Repeated measurements are used to sample more microhabitats at each site because the spatial distribution of organisms at a site can be patchy and a single measurement might not represent the composition of the assemblage. This usually results in a better estimate of the assemblage at the site. Since the site is the sampling unit, the measurement methodology can be altered to reduce measurement error. This is usually done by sampling multiple locations or habitat types, with several deployments of the specified sampling gear, and combining the hauls into a single composite sample. With a composite sample, a single measurement is taken, but the measurement is thought to be more representative of the site than a single, non-composited sample. Composite sampling that is representative of the sites is usually the most cost-effective sampling methodology. It avoids the costs of multiple measurements, allowing more sites to be sampled and increasing statistical sample size.

10.3.4 Temporal Variability

All aquatic assemblages go through annual cycles of composition and abundance changes. In addition, short-lived species also exhibit short-term temporal variability. Index period sampling, in which measurements are made during the same period each year (e.g., mid-summer), is intended to control short-term variability. Index period sampling is effective if assemblage composition and abundance are relatively stable and predictable among years. If the assemblage is not stable within the index period, it might be necessary to make repeated measurements during a season or year to obtain growing season or annual average estimates of the metrics. Repeated measurements over a season or year are more expensive and reliable than index-period sampling. For cost-effectiveness, assemblages that can be adequately characterized using index period sampling are therefore preferable to those which require repeated sampling, unless the information from the repeated sampling is more valuable.

In view of major seasonal changes in lakes, it is possible to have more than one index period. Warm temperate and subtropical lakes, in particular, might require two or more index periods because biological activity remains high year-round. Multiple index periods must be analyzed separately. Therefore, there will also be separate reference expectations and biocriteria for each season represented by an index period. Two index periods require double the sampling effort of a single index period but provide greater information on biological variability throughout the year.

10.3.5 Classification

Each lake class requires a separate reference characterization (hence, separate reference sites) and separate biocriteria. For better statistical validity, each class should have a minimum of 5 or 10 reference sites (preferably up to 30 sites). Excessive proliferation of lake classes results in an unwieldy and expensive biocriteria program.

10.3.6 Status and Trends

Estimating status and trends of lakes as a resource requires a different sampling design from that proposed here. Unbiased estimation of status requires random selection of sampling units (lakes) within sampling strata (lake classes). One approach is to assign all lakes to the classes and then randomly select a sample of lakes from each class (list-frame sampling). An alternative approach is to use a grid and sample lakes nearest the grid points, as is being done in EMAP (USEPA 1991e).

Trends can be assessed in single lakes or in a region. Several years of sampling are required for trend assessment. EPA 841-R-93-003 (USEPA 1993d) outlines trend analysis methods for lakes.

10.4 Program Resources

> *A successful bioassessment and biocriteria program depends on (1) a clear definition of goals, (2) the active use of biomonitoring data in decision making, and (3) the allocation of adequate resources to ensure a high-quality program.*

The implementation of a bioassessment and biocriteria program requires proper management and the appropriate combination of resources and expertise. Agencies already having well-developed programs usually have experienced and well-trained biologists, appropriately equipped facilities, and properly maintained sampling gear. Areas just beginning a bioassessment and/or biocriteria program need to evaluate their existing biological expertise, facilities, and equipment and expand accordingly. A cost-effective way to accomplish this is to coordinate efforts and share data with adjacent states or tribes, especially when lake or reservoir systems cross political boundaries.

10.4.1 Program Elements

Monitoring agencies can and should enhance their programs through cooperation with other agencies. For example, they should seek coordination with staff from state fishery, land management, geology, agriculture, and natural resource agencies. If federally employed aquatic biologists are stationed in a state or if the state has substantial federal lands, cooperative bioassessments and biocriteria development programs could be initiated. Scientists at universities should also be included in the planning and monitoring phases of the program - their students make excellent field assistants and future ecologists and natural resource managers. The selected team of specialists from this above pool of talent can also provide the “expert consensus” referred to earlier in defining reference conditions and developing biocriteria.

A cost-effective way to develop a bioassessment and biocriteria program is to coordinate efforts and share date with adjacent states or tribes, especially when lake or reservoir systems cross political boundaries.

10.4.2 Personnel and Resources

Several trained and experienced biologists and natural resource specialists should be available to provide thorough evaluations, support various activities, and manage quality. They should have training and experience commensurate with the needs of the program. At least one staff member should be familiar with establishing a quality assurance framework.

A biocriteria and biomonitoring program has several required elements, as well as optional elements, that determine the costs and resources of the program. Program elements include:

- *Quality assurance and quality control (e.g., standard operating procedures, training).*
- *Delineated reference conditions with annual monitoring of selected sites.*
- *Multiple assemblage biosurvey.*
- *Habitat assessment.*
- *Status and trends monitoring of a representative sample of lakes (optional).*
- *Computer hardware and software (database management, data analysis) and staff training.*
- *Documentation of program and study plans, periodic updates of analyses, and periodic review of reference conditions and biocriteria.*

Laboratory and field facilities and services should be in place and operationally consistent with the designed purposes of the program so that high-quality environmental data can be generated and processed in an efficient and cost-effective manner (USEPA 1992b). Adequate taxonomic references and scientific literature should support data processing and interpretation.

Quality management is an important planning aspect that focuses attention on establishing and improving quality in all aspects of the biocriteria development process. Quality management requires that all personnel involved in a biocriteria project (from senior management to field and laboratory technicians) be aware of and responsive to data needs and expectations.