**Monitoring Bathing Waters - A Practical Guide to the Design and Implementation of Assessments and Monitoring Programmes**
Edited by Jamie Bartram and Gareth Rees

ISBN 0-419-24390-1

# Chapter 3*: RESOURCING AND IMPLEMENTATION

* *This chapter was prepared by G. Rees, J. Bartram, E.B. Pike and W. Robertson*

Monitoring programmes, by necessity, must be commensurate with the socio-economic and technical and scientific development of the country where they are implemented. For example the extent of development of national legislation and co-ordinating or oversight programmes will affect activities undertaken. Similarly, more complex analytical variables require highly trained technicians and costly laboratory facilities. In general terms, it is possible to distinguish three levels into which monitoring programmes can be classified (Table 3.1). All elements of assessment, from objective setting to data interpretation, are related to these three levels. The aim is to progressively develop monitoring operations from the basic level to the more comprehensive levels. Each level is associated with increasing demands on staff (expertise and numbers), inspection and fieldwork (complexity and frequency), laboratory facilities (range of analysis, throughput) and data management and reporting capacity.

**Table 3.1** Levels of monitoring competence in relation to resource requirements

| Level | Basic information/visit rate | Accident hazards[1] | Microbiological parameters[2] | Cyanobacteria and algae[3] | Other |
|---|---|---|---|---|---|
| Local (no national organisation yet) | Local action comparable to basic level in some locations only | Local action comparable to basic level in some locations only | Local action comparable to basic level in some locations only | Local action comparable to basic level in some locations only | Local action comparable to basic level in some locations only |
| Basic (no access to equipment or staff resources at national level; limited local resources) | At least one pre-season visit; creation of a catalogue of basic characteristics; all beaches registered, but more used and higher risk beaches inspected and monitored | Annual inspection for identification of any hazards and interventions (e.g. signs, warning systems) | Inspection for faecal pollution or sewage odour; delimitation of high risk areas; initial screening of faecal streptococci (marine or freshwaters), *E. coli* or faecal coliforms | Inspection for scum, type, and transparency | Register of local special problems |

| | | | | | |
|---|---|---|---|---|---|
| | | | (freshwater) for primary classification; internal quality control at laboratories; at least one sample a month in season once the beach is classified | | |
| Intermediate (limited access to resources both local and national level) | Comprehensive cataloguing and timetabling of visits; additional visits during peak seasons (e.g. monthly); greater proportion of beaches monitored | Periodic verification of interventions during bathing season; central capacity for incident investigation | Identification and cataloguing of potential sources of contamination; all beaches at primary classification; monthly sampling; resampling and investigation of unexpected peak values; reclassification scheme initiated; investigation of rain effects and design of preventative measures; internal quality control at laboratories; occasional inter-laboratory comparison studies | Phosphate analysis (freshwater) Chlorophyll a (freshwater) | Check on local information availability; active warning and management response |
| Full (no significant resource limitations) | Additional visits during peak seasons (e.g. fortnightly or weekly); complete cataloguing, including updating for each recreational area; all beaches with significant use monitored | Central register of recorded incidents; decentralised capacity and procedure for incident investigation | Additional microbiological parameters if necessary; possible reclassification investigated where indicated; internal and external quality controls regularly operated; convergence amongst participating laboratories | Toxicity detection and toxin analysis capacity if necessary (not routine); remote sensing methods where relevant | Chemical monitoring (for necessary parameters) |

Each level also demands inclusion of the requirements specified at lower levels

[1] See Chapter 7
[2] See Chapters 8 and 9
[3] See Chapter 10

## 3.1 Staffing and training

The personnel in charge of sample collection, field handling and field measurements must be trained for these activities (Table 3.2). The choice of personnel for sampling depends on a number of factors, including the geographic features of the region and the systems for transportation. For example, in a small country with good transport infrastructure, sampling may be carried out by laboratory personnel going to the field to take samples, conduct field analyses and transport samples back to the laboratory. In countries of a larger size that possess a more developed monitoring system, specially-trained field personnel often conduct the sampling and inspection. In large countries that have a poor transportation system, relatively more personnel are required. In this situation specialists from decentralised facilities, such as health centres or hydrometereological and hydrological stations, may be involved in sampling, inspection and testing. Such personnel may not always possess all appropriate training.

**Table 3.2** Principal tasks undertaken by staff type as an indication of skills and training requirements for differing levels of monitoring competence

| **Level** | **Field staff**[1] | **Laboratory staff** |
|---|---|---|
| Local | Basic inventory for beach registration<br><br>Collection of water samples for microbiological and cyanobacteria analysis<br><br>Transparency measurement (Secchi disc)<br><br>Cyanobacterial scum recognition<br><br>Basic sanitary survey | Data analysis and management<br><br>Analysis of faecal indicator bacteria (according to indicator and method available) |
| Basic | More complex sanitary survey<br><br>Selection of sampling sites<br><br>Intensive microbiological sampling for primary classification<br><br>Observational verification of effectiveness of interventions | Phosphate analysis<br><br>Chlorophyll a analysis<br><br>Organisation and implementation of necessary quality assurance |
| Intermediate | Local follow-up of unexpected peak microbiological values (including liaison with local authorities) | Participation in occasional, informal interlaboratory comparisons (probably “round-robin”) |
| Full | Participation in accident investigation | Cyanobacterial toxicity detection and |

| | | |
|---|---|---|
| | | toxin analysis, where relevant<br><br>Participation in regular, formal inter-laboratory comparisons |

For a description of the different levels of monitoring competence, see Table 3.1

[1] The balance between field and laboratory staff is determined by local factors

This guidebook provides a major part of the information needed for carrying out successful and adequate fieldwork and sampling. The quality of information produced by a monitoring programme depends on the quality of the work undertaken by field and laboratory staff. The importance of appropriate training cannot be over-stressed.

Separate training packages should be developed for field staff, laboratory staff and others. It should be emphasised that training is not a "once-only" activity, but should be continuous. Supervision, as a form of training, is especially relevant to laboratory and field staff. It is vital that the training function is flexible, responding to experience and feedback and taking account of specific needs. Training is especially important when programmes for monitoring are implemented in several countries or independently managed areas or regions, from which the results will be compared and used outside the country or region where the monitoring has taken place.

Assuming a good general education or relevant previous experience, the training period for staff responsible for fieldwork and sampling is about one week, with approximately one additional week of further training as monitoring develops through basic, intermediate to full levels. In order to maintain motivation, it is recommended, that short follow-up events are provided. If the staff are less experienced more training will be necessary.

Although not exhaustive, training for field staff should include the objectives of the water quality monitoring programme and its local, national and international significance. The training should stress the importance of samples being of good quality and representative of the water body from which they are taken and it should give guidance on how to ensure that the samples meet those requirements. The training should also include planning of field sampling and map reading, as well as how to make field notes describing the sampling site and station and how to undertake on-site inspections. Safety aspects of field sampling are an important component of the training programme.

Staffing requirements for servicing a monitoring programme may vary widely and it is not possible to make general statements about the number of staff needed for fieldwork. Estimating staffing requirements includes allowing for travel time or distance between beaches and laboratory, the choice of laboratory infrastructure (e.g. centralised or decentralised) and co-ordination between participating institutions.

The head of the laboratory and/or the programme manager are generally responsible for:

- Laboratory management.

- Determining and procuring the equipment and supplies that will be required.

- Ensuring that Standard Operating Procedures (SOPs) are being followed (Chapter 4).

- Ensuring that adequate quality control procedures are being followed.

- Enforcing safety procedures.

Laboratory technicians must have suitable training. They will generally be responsible for:

- Maintenance of the laboratory including, cleanliness and safe storage of all equipment, glassware and other reusables.

- Storage and preparation of reagents and media.

- Checking the accuracy of field equipment.

- Training of junior staff.

- Performing the tests and recording the results of field analyses.

## 3.2 Laboratory and analytical facilities

The choice of laboratories to be involved in the monitoring programme can have a major influence on the time required for collecting the samples. There may be reasons such as ensuring staff training, equipment repair and analytical quality control, that suggests using a central laboratory is more appropriate. Nevertheless, when total sample numbers are high and samples are transported over long distances, it may be preferable to use more than one laboratory. Under these circumstances it is normal practice for one of the laboratories to act as the central or co-ordinating laboratory. Some countries have developed mobile laboratories (typically in small vans or minibuses) as an alternative solution to the problems of sample transport. Where long distances are encountered this has often been found to be an effective approach. The need for co-ordination between institutions, for example when sampling is undertaken by one agency and analysis by another, may reduce the necessary workload but may lead to inefficiency if the co-ordination is not effective.

In many countries, monitoring laboratories are organised on two tiers: regional laboratories (lower level) to conduct basic determinations not requiring very complex equipment, and central laboratories (higher level) to conduct more complex analyses requiring elaborate equipment and well trained personnel. In addition, the central laboratories often provide the regional laboratories with methodologies and analytical data quality control.

During the initial stages of development of a recreational water monitoring system, it is reasonable to focus on the basic variables that, as a rule, do not require expensive and sophisticated equipment. Gradually, the number of variables measured can be increased in relation to the financial resources of the monitoring agency. Even in fully-developed monitoring programmes, the elaborate equipment and technical skills necessary for the measurement of complex variables are not needed in every laboratory.

Laboratories must be selected or set up to meet the objectives of each assessment programme. Attention should be paid to the choice of analytical methods. The range of concentrations measured by the chosen methods must correspond to the concentrations of the variable in a water body and to the concentrations set by any applicable water quality standards. Ideally, a laboratory should consist of four sections:

- Reception and registration of samples.
- Production of analytical media.
- Sample analysis.
- Washing up and autoclaving equipment.

However, it is possible for a laboratory to function properly in fewer sections, depending on the number of analyses to be undertaken.

## 3.3 Transport and scheduling

Resource elements in fieldwork include transportation, personnel requirements and equipment. Problems of transport are largely related to distance and accessibility of sampling sites. Where access to sites is known to be problematic, reliance has been placed traditionally on four-wheel drive vehicles, but these are often expensive to purchase and to operate. It may generally be assumed that the main part of the beaches included in the routine monitoring programme can be reached using ordinary vehicles because beaches difficult to access have most probably been excluded when selecting the participating beaches. Some beaches may be located such that using motorcycles or public transport is a possibility; however, most frequently, ordinary cars equipped for transportation of the samples will be required. Sampling transport equipment can consist of an insulated box with melting ice, or a refrigerator installed in the vehicles, to ensure that the samples are kept cool.

Travel time to and from sampling sites is a major constraint for staff undertaking fieldwork. Realistic estimates of travel time to each sampling site should be made as early in the programme as possible. An inventory of sampling stations should be developed, including actual travelling time, in order to facilitate programme planning. When planning the route for visiting the sites, the demand of having the samples analysed at the laboratory within the appropriate time frame must also be taken into account (see Chapter 8 for microbiological analyses and Chapter 10 for cyanobacteria).

## 3.4 Inspection forms and programmes

Public health authorities should have, at least, a basic inspection programme in place for all recreational-water sites within their jurisdiction. The primary purpose of the inspection programme is to minimise the risk of illness or accident to bathers. The programme should be based on an SOP (refer to Chapter 4). This will ensure that the on-site inspection, laboratory analyses and interpretation of data are carried out in an objective and uniform manner. An on-site survey form should be prepared as a guide for inspectors to make certain that all aspects of the site receive adequate review and evaluation.

In most programmes at least two forms will be used. The first “basic registration” form collects the minimum background information to construct a register. The beach register is a high priority during the basic level of monitoring. The principal uses of the beach register are:

- To determine which beaches should be considered by the programme, based upon criteria such as extent of use, degree of development (or plans to develop) and already-recognised hazards.

- To provide data (such as transport options and travel times) to assist in programme planning.

Details of the information to be included in a beach registration form are provided in Chapter 2.

Once a beach enters into a monitoring programme it will be subject to periodic, usually annual, inspections. These inspections are principally orientated towards the identification of hazards that might lead to physical injury or contribute to drowning; towards the adequacy of measures (signs, lifeguards, communications) in place to reduce these risks; and towards sources of microbiological pollution such as sewage outfalls, combined sewer overflows, rivers and storm drains. At freshwater sites, the inspection may also be concerned with the likelihood of cyanobacterial blooms (see Chapter 10). An example of a sanitary survey form is included in Chapter 8 (see Box 8.1) and the components of an on-site inspection are discussed further in Chapter 7 (physical hazards), Chapters 8 and 9 (microbiological aspects) and Chapter 12 (aesthetic aspects). Other inspections may be required sporadically depending on local circumstances, for example to assess bather load (Chapter 8), cyanobacterial hazards (Chapter 10) or to verify the effectiveness of interventions to control microbiological quality (Chapter 9).

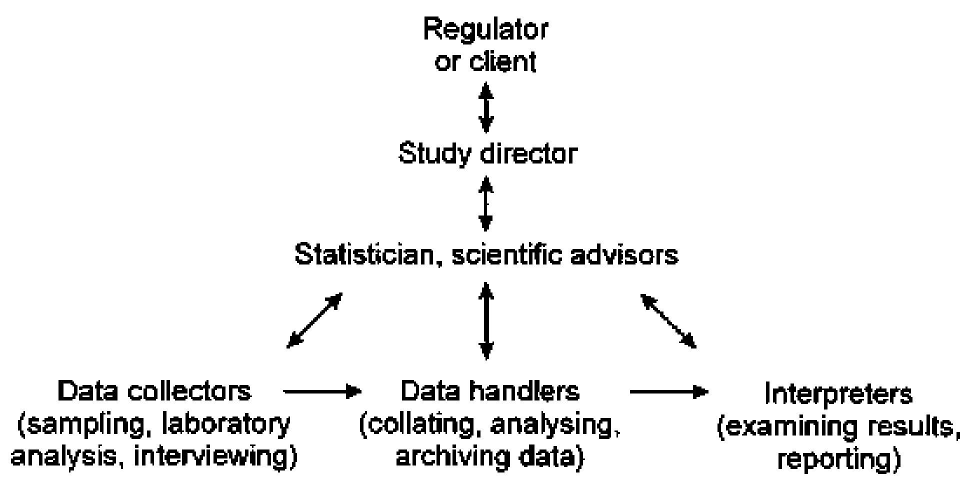
A key requirement of any inspection programme is the availability of qualified personnel. Ideally, inspectors should have a basic knowledge of public health microbiology, environmental chemistry, limnology, oceanography, estuaries and meteorology. Training workshops, based on SOPs and actual case studies, should be held during programme development and implementation. Additional workshops should be held periodically to review inspection procedures.

The co-operation of all involved or interested individuals or organisations is essential if assessments and monitoring programmes are to result in improved water quality and reduced health risks. Representatives from user groups, tourist associations, beach and resort owners, industries and sewage treatment facilities and public health laboratories should be aware of, and invited to participate in, the programme. The complexity and completeness of surveys are likely to increase as programmes develop from local through basic and intermediate to full levels (see Table 3.1). At the most advanced level, for each site, three to five days may be required to prepare for the survey, conduct an inspection and prepare the report, although subsequent re-verification visits may require less time. From these findings public health authorities will be able to classify the beaches within their jurisdiction with ratings ranging from excellent to very poor and to relate these ratings to requirements for monitoring and management. Chapter 9 discusses in depth the development of classification schemes for microbiological quality. In this way, resources can be directed to those beaches that present the greatest risks to public health.

## 3.5 Data processing and interpretation

If data have been collected in a careful manner in a properly designed study, they will be representative and unbiased and suitable for analysis, interpretation and reporting. Processing, management and storage of data can be collectively described by the term "data handling". This activity has a central position in any study as shown in Figure 3.1. The data handlers receive data from samplers and interviewers in the field, as well as the results of analyses from laboratories. They then assemble and archive the data in central files and analyse the data as instructed, so that they can be interpreted by those responsible for reporting the results. It is most desirable to prevent bias, from the introduction of personal or political perceptions, by arranging for separate groups of people to carry out the collection, handling and interpretation of data. Data handlers must be appropriately trained and qualified for their duties. It is also desirable that the data handlers do not have any particular interest in the results of the study, apart from carrying out their work with dedication and accuracy. The whole process of the study should be under the control of a study director, who will receive instructions on the objectives and on the programme to be carried out as well as statistical and other scientific advice from appropriate experts. In this way, the requirements for effective data management (Box 3.1) will be fulfilled. The period over which the data may be needed and the purpose for which they could be used should be considered at the outset, particularly for large or continuing studies. Computer hardware and software systems inevitably change over time and future compatibility could be a problem. For example, it is doubtful whether data stored on punched cards or paper tape some 30 years ago could now be retrieved. Moreover, changes in methods of analysis will reduce the compatibility of data from older studies with those of recent studies, thus rendering the raw data (but not necessarily the conclusions) less useful. Consideration should be given to storing data in several places, and in more than one format, in order to forestall the risk of loss and damage.

**Figure 3.1 The central relationship of data handling in the conduct of monitoring programmes and scientific studies**



**Box 3.1 Requirements for effective data management**

The requirements of the monitoring programme must be defined. This is the responsibility of the study director, instructed by the regulator or client and as advised by the statistician and scientific experts. In the case of a scientific survey or epidemiological study, the effects under study are framed as conceptual models, requiring the testing of null hypotheses by statistical methods. The data required are determined by the monitoring schedule imposed by the regulatory authority or, in the case of a survey, by the desired statistical power.

Appropriate forms or other recording instruments are produced for collection of the raw data required to meet the above requirement. These must be approved by the collectors and handlers of the data, so that all suppliers of data use a standard, approved format. It is essential that the raw data are in a form that can be transcribed easily and with minimal risk of error to a permanent file, suitable for processing and archiving.

The data are analysed by appropriate methods to produce the desired output. The appropriateness of the methods must be decided earlier in fulfilment of the first requirements. Suitable analytical quality control procedures must be used for input and analysis of data.

For the long term, it is vital to ensure that published reports carry adequate details of the methods of surveys and that the key data and basic statistics, such as means and standard deviations of measurements, are recorded as appendices to the final reports. The basic test of the adequacy of survey recording is whether an independent investigator could understand what was done sufficiently to be able to repeat the work and to be able to test the results and conclusions from the data presented.

The greatest cost of any study is the data collection. These data will be unique and there will usually be no opportunity of repeating a missing or erroneously recorded observation. The raw data sheets should be examined critically by the person in charge of collecting the data as soon as possible after recording, so that discrepancies can be detected and corrected while the events can still be remembered. It may be possible to check suspect meteorological and tidal records with those obtained at nearby, official recording stations. Different sets of data (e.g. from different sites, sampling runs or days) must not be pooled, unless it has been shown that there is no significant statistical difference between the data sets.

It must be remembered that data are collected for a specific purpose, and thus they may not be useful for meeting subsequent objectives. Most data from routine, regulatory monitoring fall into the “data-rich, information-poor” category because other information, necessary for explaining the circumstances or trends, was not recorded at the time. The use of statistical methods to test *a priori* null hypotheses is an obligatory part of scientific method. However, the temptation to carry out further analysis of the complete data set at a later date, in order to detect statistically significant associations or correlations between variables, is strictly invalid unless used solely as a method of suggesting hypotheses for further study. Such “data dredging” is analogous to fitting targets to holes (Jolley, 1993).

The effects studied in behavioural and epidemiological research of water recreation are typically small and are difficult to separate from confounding factors (i.e. factors that affect the responses of the exposed and control groups differently). Potentially confounding factors must be identified at the planning stages of a study and suitable measures should be introduced into the design of the study (e.g. by the matching of exposed and non-exposed subjects, or by random assignment) and in the analysis of the data to detect or nullify them. Confounding can be detected or corrected for in analysis by various methods, such as stratification of the exposed and non-exposed groups (e.g. by common potential confounders such as age, sex or socio-economic class) or by including potential confounders as variables in multivariate analyses, such as logistic regression analyses (for a description, see McCullagh and Nelder, 1989). It is not possible to eliminate the effects of confounding in the data, if it has not been considered, measured and examined in the design of the study (see Chapter 13) (Datta, 1993; Leon, 1993; WHO, 1998).

### 3.5.1 Database construction

The construction of the database will depend on:

• The level of monitoring as defined in section 3.2 and the intended use of the data, e.g. for compliance monitoring, risk assessment, baseline data, acquisition, new scheme design, epidemiological investigations and/or post-audit evaluations or remediation efficacy.

• The technology available.

• The requirement for data transfer to regulators or to other agencies.

• The requirements for “data audit” and “chain of evidence” procedures.

There are many data storage systems available that have been developed by laboratories and software engineers. Perhaps the most stringent and sophisticated database would be characterised by a data storage system able to accommodate the information suitable for a full “chain of evidence” assessment. This type of database is being developed in drinking water surveillance monitoring. The reason for the high level of laboratory audit stems from the potential litigation that could derive from contaminated waters and associated disease outbreaks. This type of database is emerging in drinking water assessment in the UK following a consultation document produced by the Drinking Water Inspectorate (DWI, 1998). The requirements of such a data storage programme are that the data could be used in a court of law to achieve a criminal prosecution. To achieve this the prosecution is generally required to demonstrate that the data are accurate beyond reasonable doubt, i.e. the analyst can prove that there are systems in place to prevent tendentious or accidental misreporting through laboratory analytical error or database construction mistakes (such as recording a result from the wrong sample).

Laboratory audit control is now routine in laboratories undertaking compliance assessment programmes for recreational and drinking waters in some countries. Here, the recording laboratory must be in a position to prove that the recorded value is correct and that laboratory procedures have been followed. The level of internal control is less than for “chain of evidence” assessments, i.e. the laboratory would have to prove that all reasonable measures had been taken but it would not have to demonstrate that there was no error “beyond reasonable doubt”.

Any environmental sampling programme should implement appropriate intralaboratory and interlaboratory analytical quality control (AQC) procedures. These would normally involve the collection of field and laboratory blank (e.g. sterile) samples to investigate the integrity of aseptic techniques and field duplicate samples to investigate reproducibility. Generally, it is preferable to use split samples from the same bottle in the case of duplicate enumerations. In inter-laboratory trials duplicate samples are split, or spiked samples are prepared and delivered to participating laboratories for analysis. Analytical quality control is described in full in Chapter 4.

There are many commercial data storage software systems available. The developing international standard is a spreadsheet that facilitates storage, rudimentary statistical analyses, graphical representation and data export for external communication and reporting. Spreadsheet packages are commonly used for the storage and recording of recreational water quality data worldwide. Such systems are appropriate for the user familiar with computation methods and standard package use. They may require adaptation for the lay user or to make them effective tools for clerical staff with a low level of information technology (IT) skills. In such circumstances, the spreadsheet can be modified with a Graphical User Interface (GUI). This facilitates rigid, but user friendly, data access and export, appropriate and tightly controlled analysis and, where appropriate, data cleaning and security. Perhaps the most widely used programming language in the production of a GUI is Visual Basic. This language can be used to design forms and screen menus to control data input, data analysis and reporting, as well as to provide security through the use of passwords. Visual Basic programming requires a competent programmer but the advantages in its use for data security and quality are very significant.

The data storage system should be appropriate to the intended purpose of the storage exercise. If the objective is to demonstrate chain of evidence or a future audit of the data for compliance, a bespoke system will probably need to be constructed to facilitate clerical input, data checking by non-technical staff, data cleaning after extreme values have been flagged, and clearly defined data reporting as defined by legislation. Here, linking a GUI with a spreadsheet through Visual Basic is probably the most appropriate route. Where the data are simply being stored for scientific and/or baseline definition purposes, and no immediate audit beyond normal AQC is required, the spreadsheet alone can be a suitable vehicle for data storage.

### 3.5.2 Preliminary examination of data

It is important, in the interests of consistency, to agree at the outset the procedure to be used for dealing with missing, indeterminate and outlying values, and to adhere to these procedures consistently. Sets of data frequently contain missing values. These values can often be estimated from trends in the data set (where the subsequent analysis requires a complete set) but it must be realised that such “patching” reduces the number of degrees of freedom on which to base statistical decisions. Missing data cannot be estimated or reported when monitoring to assess compliance with a standard for water quality.

Indeterminate values are frequently found, implying that the volume of sample examined, or the concentration of the determinand, was too small or too large to be within the limits of the method. The analysts should be instructed to record the facts, i.e. “less than” or “greater than” the analytical limit, together with the volume examined. This, at least, enables a rank order of values to be established. Procedures for estimating a mean from data with indeterminate values are described below. The particular difficulties encountered in microbiological analyses are explained in Chapter 8.

Detailed examination of data often reveals values that lie outside the normal range of values or trends in the data and which therefore seem improbable (known as “outliers” or doubtful values). Some computer programmes are able to identify outliers, according to defined criteria. Such doubtful values must be investigated by going back to the original records and, if at all possible, to all laboratory personnel and samplers who were responsible for obtaining the value recorded. Only if there are strong technical reasons (e.g. contamination of a batch of culture medium, or a fault in a recording instrument) should such values be deleted from the data set. There are no valid statistical reasons for excluding outliers from sets of data. Their occurrence provides a strong case for investigation of sampling and laboratory procedures. They also indicate the value of carrying out simple checks for the consistency of data in the laboratory at the time when they are first recorded. If no technical problem is found, such values must be accepted. These values could be the first indication of a change in water quality or they may be a random, infrequent event.

Due regard must be taken of the underlying nature of the probability distribution within the data collected, because this will determine the most appropriate way of expressing the central tendency and dispersion of the data. This should be taken into account when standards are derived, because standards will invariably specify their requirements and limiting values in terms of statistics, such as the average, median or geometric mean, or an upper percentile value.

**Table 3.3** Examples of the different probability distributions which may be encountered in surveys of recreational water use areas

| Distribution | Properties and examples |
|---|---|
| *Discrete (whole-number) random variables* | |
| Binomial | Results of a sequence of independent trials, specified in advance and with two possible outcomes (“yes”, “no”) and constant probability of success from trial to trial. |
| Poisson | Describes occurrence of random events in a continuum (e.g. annual deaths by drowning in a region or counts of randomly distributed particles in independent, identically sized samples); variance and mean are identical. |
| *Continuous random variables* | |
| Normal | Conforms to the normal probability density function, distributed symmetrically about the arithmetical mean (average); mean and median are identical. Distribution is described by a mean and the standard deviation. Heights and weights of individuals, errors of analysis and data for many physical and chemical measurements usually approximate to normality. |
| Log-normal | The logarithms of the values are normally distributed. Distribution is described by the geometric mean, which is equivalent to the median, and by the standard deviation of log values (the log standard deviation). Generated by random variation |

| | |
|---|---|
| | of the rate constant in natural processes subject to exponential decay or growth; sets of microbiological counts or chemicals diffusing in water, or particle-size analyses of sediments are therefore typically log-normally distributed. |
| *Ordered, categorical variables* | |
| | Data can be arranged and ranked in order of size or categorised, but do not have discrete or continuous values. For example, water users can be arbitrarily ordered according to their observed degree of contact with water (i.e. none, wading and paddling or head immersion). In other cases, the categories do not have a size or implied order (e.g. water-contact recreation may be swimming, surfing, rafting or diving). Data can be tested by appropriate non-parametric methods. |

Parametric tests for statistical significance assume that the data conform to a particular model of distribution, such as the normal distribution, and thus are only valid when applied to data that are known to conform approximately to that distribution, or can be transformed appropriately so that they do so. This is a problem where statistical advice must be sought before attempting analysis of data. However, many non-parametric tests of significance, such as those using ranked values instead of actual numbers, are inherently distribution-free, but they lack the statistical power of the parametric tests.

Table 3.3 lists examples of frequency distributions that are commonly encountered in the data collected in surveys of recreational water-use areas, together with their properties. Invariably, there will be data sets that do not conform to the frequency distributions listed in Table 3.3 and, at best, only an approximate fit to the frequency distributions will be possible. A full treatment of probability distributions and their properties can be obtained from textbooks of probability and statistics (e.g. Devore, 1991).

The use of the mean and the standard deviation to describe central tendency and variability in data is appropriate if the data are distributed normally or approximately normally. This is usually the case with physical and chemical data. Microbiological, virological and biological data are almost invariably found to be skewed and distributed log-normally or approximately log-normally. This is thought to occur because of growth, decay and dispersion processes in natural waters, which tend to follow exponential (first-order) reactions, in which the rate constants are subjected to random environmental fluctuations. Skewness caused by log-normality can be detected in several ways. It should be suspected if the data contain many relatively small values and relatively few very large values, or if the average (arithmetical mean) is much larger than the median.

Tests for skewness are given in many statistical textbooks. Log-normally distributed data can be made to conform to the normal distribution before analysis if they are first transformed to logarithms. The geometric mean (i.e. the antilog of the average of the log values) and the standard deviation of the logarithms are the appropriate statistics for such data. The geometric mean cannot be calculated if any of the values are zero (e.g. “undetectable in the volume examined” or “less than the limit of detection”). In such cases, the median can be recorded as an equivalent to the geometric mean, with an explanatory note, or a log ($x + 1$) transformation can be used, by adding 1 to all the values of $x$ before taking log values. The reverse transformation is used in expressing the geometric mean.

Other transformations are sometimes used. Whole number counts of random events, such as deaths by drowning, usually conform to the Poisson distribution. This should be suspected if the variance (the square of the standard deviation) of the data is similar numerically to the mean. The appropriate transformation is to take square roots of the values, if they lie in the range 10-100, or $\sqrt{(x + 0.5)}$ if the values are less than 10. If most of the numbers are greater than 100, transformation is not needed. In the case of values which are rates or ratios, such as velocities ($m\ s^{-1}$), reciprocals should be used and the correct mean of the *n* values of *x* is the harmonic mean:

$$\frac{1}{\sum(1/x)/n}$$

Skewness and the effects of transformation can be detected by constructing frequency distributions. This is most readily done by graphical analysis on the computer, but may also be done with small data sets by constructing cumulative frequency plots on normal probability graph paper. Such plots will approximate to straight lines, with slopes proportional to the standard deviation if the data are normally distributed, or if the transformation is appropriate.

### 3.5.3 Internal data check mechanisms

Mechanisms for checking the quality of data require links between sampling and AQC (which is achieved through rigorous programme design) and appropriate data checking mechanisms. For purposely built systems, the links can be built into the GUI or Spreadsheet or Graphics system. For example, the implementation of field and laboratory blank (sterile) and duplicate samples, together with participation in inter-laboratory AQC programmes, should ensure the numerical integrity of the data acquired, provided appropriate dilutions are employed. Thus, data reporting unexpected low or high values should provide a true representation of indicator bacterial concentration or physicochemical parameters. However, the entry of the data to the database offers a further opportunity for automated data checking, principally in the form identifying outliers in the data set. The definition of an outlier requires a historical data set. Identification of a specific data item as an outlier value is based on a knowledge of the statistical distribution of the environmental determinand. In the case of microbiological data, it is generally accepted that environmental data measured at recreational waters follow a $\log_{10}$-normal probability density function. Thus an outlier could be defined as a data item where the $\log_{10}$ value was greater than, for example, two standard deviations from the historical $\log_{10}$ mean value.

This above approach is perhaps the most scientifically valid, although, it presupposes historical data and has a significant problem which derives from the choice of a cut-off at which the data item is considered an outlier. It is certainly the case that microbiological concentrations in recreational waters can commonly increase by several orders of magnitude (i.e. 3-4) following rainfall events. Thus, the definition of a numerically low cut-point tied to an automatic data cleaning system designed into a GUI and operated by clerical staff with little scientific insight into the acquired data, might result in very significant data loss from precisely the most high-risk periods against which the system was seeking to provide protection. For this reason, it is essential that data cleaning is closely supervised by competent scientific staff and that any automated systems simply define the items on which a scientific decision is required, rather than being allowed to carry out any automated change to the raw data matrix by deletion of apparent outlier values.

### 3.5.4 Determining compliance with a standard

Many monitoring programmes are designed to ascertain compliance with a standard or other objective for water quality. The standard should be carefully checked to determine the method that will be needed to assess compliance. Any deviation from the specified method will lead to doubts about comparability of the conclusions between parallel programmes in different regions, nationally or internationally. The compliance method will have three basic components: the design of the sampling scheme, (including the number of samples, their frequency and the period of sampling); the description of water quality (such as the chemical, physical or microbiological variable), the units of measurement and the descriptive statistics used to describe the level attained in the set of observations (such as the range of values, the average, median, or geometric mean, the standard deviation, or a given percentile value); and the criterion for judging passing or failing of the standard, or for classifying the quality at the water recreation area.

Water quality at a particular site is, essentially, a continuous population of measurements of quality, from which a sampling programme can only provide a limited number of discrete measurements. Because of the errors of measurement that occur when a small number of samples are taken, the sampling programme only gives an approximate estimate of the conditions existing over the whole period of the programme. These errors may, or may not, be regarded as important, although they will have most significance when assessing compliance of waters which are of borderline quality in terms of the standard. Once this problem is recognised, it is important to consider the burden of proof required to assess compliance. It is most usual to take results at their face value, without taking account of sampling error. This is justifiable, but is an empirical decision which will always carry a risk of misclassifying waters, particularly if small numbers of samples are taken. Alternatively, allowance may be made for sampling errors by adopting a “benefit of the doubt” or a fail-safe approach, so that when water quality is borderline the risk of incorrectly failing or passing the standard, respectively, is acceptably small and defined in statistical terms. A general description of these problems in the design of water quality monitoring programmes has been given by Ellis (1989).

Most standards specify a limiting value, which may be either a measure of central tendency, such as a mean or median value obtained over a specified period, or an upper limit which must not be exceeded. An upper limit may be absolute, or it may allow for the natural variability that occurs with time and which is caused by such factors as weather and tidal conditions. Such allowance can be made by defining an upper confidence limit, defined statistically by the observed variability or by the percentage of samples (percentile) which must not exceed the upper limit.

### 3.5.5 Data presentation

In all except the smallest surveys, the study director and data handlers will wish to use electronic computation for speed and accuracy. They should, however, use reliable and proven statistical packages and assure themselves that the calculation routines give the correct output. Manual calculation may be the only method available to small teams, handling small data sets and lacking computation facilities, or for field use. It can also be invaluable as a means of checking the data as they are produced, or for checking the output from computation. Desk calculators that provide a printed output of the operations and calculations are recommended because the printout provides a means of checking errors of input.

Graphical methods can be used manually when the data set is small. Linear probability graph paper, which has ordinates ruled in equal divisions and abscissae ruled proportionally to the percentage points of the standard normal density function, can be used to check that the data (transformed if necessary) conform approximately to the normal distribution. If so, the graph paper can then be used to estimate the median, other percentile points and the standard deviation of a set of data, with two-figure accuracy.

Much data processing is associated with assessments of microbiological quality and this is described in Chapter 8. Water quality data points may be plotted sequentially on a chart to indicate changes of quality with date of sampling. The occurrence of high values may be used to initiate investigation. In these circumstances the plot becomes a control chart. The choice of a limit value and appropriate action, for when the limit is exceeded, will be chosen to suit the need. The limit value can be set to coincide with the value given in a standard. More conventionally, two upper limit values are set on a control chart: at the mean plus twice and at the mean plus three times the sample standard deviation. These represent values that would be expected to be exceeded only once in 20 or 100 samples and which indicate, respectively, a warning and the need for remedial action.

Use of a computer is obligatory for analysing large sets of data. Even when it is possible to use a hand-held calculator or even to use graphical methods, the computer is able to produce results free from error and, in many cases, to reduce the total time expended in data analysis. There are many statistical packages available that can take raw data from a spreadsheet and subject them to statistical analyses and in many cases use them to produce excellent graphics. Such packages are available to suit different needs. For general use, typical packages that are commercially available are MINITAB Statistical Software, SAS (Statistical Analysis System), SPSS (Statistical Package for the Social Sciences) and BMD (Bio-Medical Data programs). Specialised computer programs include GENSTAT (a comprehensive collection of statistical programs available from Numerical Algorithms Group Ltd) and EpiInfo (public domain software for epidemiological investigations). Each of these has graphical capabilities adequate for most users, although they can also produce output appropriate for more sophisticated and internationally available graphics and publishing software.

### 3.5.6 Data interpretation and communication

Interpretation of data and communication of results are the final two steps in an assessment programme. Correctly interpreted data will not be of much use if they are not disseminated to relevant authorities, to scientists and the public, in a form that is readily understandable by, and acceptable to, the target audience. The form and level of data presentation is, therefore, crucial. Often, it is advisable to produce two types of reports: a comprehensive, detailed report containing all relevant data and associated interpretation and an executive summary (in an illustrated and simple form) which highlights the major findings. Usually, the interpretation of data is undertaken by specialised professionals, such as the relevant scientists, (e.g. microbiologists, chemists or epidemiologists) the data treatment team and professionals from other organisations such as environmental protection agencies, health authorities, national resource agencies. As a courtesy, the results and recommendations should be discussed with all interested groups and individuals before reports are formally released. A contingency plan should also be developed with the assistance of all those with a vested interest to investigate and respond to cases of adverse health effects or to any unforeseen event or conditions that could lead to a deterioration in water quality and possibly increase the risk of illness to bathers.

A very important part of any sampling exercise is to review the extent to which the desired objectives have been achieved. There are many reasons for periodic adjustments to any assessment or monitoring programme for recreational water quality. The initial objectives may have been achieved and the programme may need reorientating from a baseline study to routine monitoring, with the establishment of new objectives and possibly the addition of supplementary monitoring variables or the substitution of existing activities with new activities. Once the samples have been taken, contemporary information is available on distributions and variability. If the required precision has not been achieved, these new data may be used to establish how many extra samples are required, and how the sampling strategy may be further optimised. If necessary, any extra sampling may then be carried out immediately.

## 3.6 Elements of good practice

The logistical planning of any monitoring programme or study should take account of socio-economic, technical or scientific and institutional capacities, staffing, equipment availability, consumable demands, travel and safety requirements and sample numbers, without compromising achievement of the objectives or scientific validity of the programme or study.

- The hierarchy of authority, responsibility and actions within a programme or study should be defined. All persons taking part in the programme or study should be aware of their roles and interrelationships.

- Staff should be trained adequately and be appropriately qualified, including in respect of health and safety aspects.

- Collection of data and information should use the most effective combination of methods of investigation, including observation, water quality sampling and analysis, interview of appropriate persons and review of published and unpublished literature.

- Frequency and timing of sampling and selection of sampling sites should reflect beach types, use types and density of use, as well as temporal and spatial variations in the recreational water-use area that may arise from seasonality, tidal cycles, rainfall, discharge and abstraction patterns, beach types and usage.

- Sampling should provide a data set amenable to statistical analysis.

- Data handling and interpretation of results should be done objectively without personal or political interference.

- The need for transformation of raw data, before analysis, to meet the conditions for statistical analysis should be agreed with a statistical expert before commencing analysis.

- Data handlers and collectors should agree on a common format for recording results of analyses and surveys and should be aware of the ultimate size of the data matrix. Forms and survey instruments should be compatible with this format. Likewise, data handlers should agree on a format for the output of results with those responsible for interpreting and presenting the data.

- Procedures for dealing with inconsistencies, such as omissions in records, indeterminate results (e.g. indecipherable characters, results outside the limits of the analytical methods) and obvious errors should be agreed in advance of data collection. On receipt from the data collectors, record forms should be examined and the agreed procedure followed. Discrepancies should be referred immediately to the data collector for correction or amendment. Where re-sampling is impossible, estimates are preferable to leaving gaps in the data record, (estimates should always be recorded as such) although they will reduce the statistical degrees of freedom.

- Ideally, arrangements should be made to store data in more than one location and format, to avoid the hazards of loss and obsolescence. Data should be transcribed accurately, handled appropriately and analysed to prevent errors and bias in the reporting.

- The statistical routine should be selected by a statistical expert.

- Data should be handled and stored in such a way to ensure that the results are available in the future for further study and for assessing temporal trends.

- Data should be interpreted and assessed by experts with relevant recommendations for management actions prior to submission to decision makers. Interpretations should always refer to the objectives and should also propose improvements, including simplifications, in the monitoring activities - stressing the needs for future research and guidelines for environmental planning.

- Interpretation of results should take account of all available sources of information, including those derived from inventory, catalogue of basic characteristics, sanitary and hazard inspection, water quality sampling and analysis, and interview, including any historical records available.

## 3.7 References

Datta, M. 1993 You cannot exclude the explanation you have not considered. *The Lancet,* **342**, 345-347.

Devore, J.L. 1991 *Probability and Statistics for Engineering and the Sciences,* 3rd Edition. Duxbury Press, Belmont, California.

DWI 1998 *Standard Operating Protocol for the Sampling of* Cryptosporidium *Oocysts in Treated Drinking Water.* Drinking Water Inspectorate, London.

Ellis, J.C. 1989 *Handbook on the Design and Interpretation of Monitoring Programmes.* WRc Publication NS 29, WRc plc, Medmenham, Marlow, Bucks, UK.

Jolley, D. 1993 The glitter of the table. *The Lancet,* **342**, 27-29.

Leon, D.A. 1993 Failed or misleading adjustment for confounding. *The Lancet,* **342**, 479-481.

McCullagh, P. and Nelder, J.A. 1989 *Generalised Linear Models,* 2nd Edition, Chapman & Hall, London.

WHO 1998 *Guidelines for Safe Recreational Water Environments.* Draft for consultation. World Health Organization, Geneva.