CWS GUIDELINES TO PRACTICAL QUALITY
ASSURANCE FOR CONTRACTED CHEMICAL ANALYSIS

Richard Turle
R.J. Norstrom

TECHNICAL REPORT SERIES NO. 21
Headquarters 1987
Canadian Wildlife Service

This series may be cited as
CWS Guidelines to Practical
Quality Assurance for Contracted
Chemical Analysis. Technical Report
Series No. 21, Canadian Wildlife
Service Headquarters.
Abstract

The concern over the quality of the environment has led to many chemical analyses. In order to identify problems, accurate and precise analyses are demanded. The reasons for a Quality Assurance Program for an analytical contract are outlined. Various procedures such as control charts, use of reference materials and audits are discussed. Recommendations are made for the reporting of detection limits and low level data. Criteria are presented for the acceptance or rejection of results from a contract analytical laboratory.

Proposed contract clauses and suggestions for additional reading are included.
CWS Guidelines to Practical Quality

Assurance for Contracted Chemical Analysis

Introduction

In the last 15 or so years the public's concern over the causes of pollution and the quality of the environment has led governments worldwide to embark on monitoring and research programs. The majority of these require analysis of various chemicals both naturally occurring and anthropogenic. Over the last few years, it has become evident to scientists that many of the analyses were subject to criticism because the results differed from those produced by other groups. This difference was often just different people working with different methods in different laboratories. Particularly evident were the results in such diverse areas of analysis as pH and PCBs. To maintain credibility in the eyes of a very sceptical public, these differences have had to be resolved. Many of the causes of variability between two projects investigating similar problems in the environment are beyond the scope and influence of the analytical chemist (1). Unfortunately, the analytical chemist is the one producing the numbers and he is invariably the one at which the finger is pointed. Thus, it is essential that the numbers produced by analytical chemists are described in terms of quality, i.e., are they acceptable (in terms of accuracy and precision). Much has been written (Appendix I) to guide managers responsible for "good" numbers from their analytical laboratory. Some of these documents are useful for the manager of a contract for chemical analysis (2, 3, 4, 5).
The view that a manager of a contract can do little to assure high quality in a contract lab is now less prevalent than in former years. It is the view of the authors that a manager can demand consistent quality within a contract, providing a clear method of determining quality is outlined at the proposal stage. These Guidelines include model clauses (Appendix II) to assist in the drawing of suitable QA/QC clauses in contracts.

Approach and Applicability

The CWS approach to quality awareness has been based on the use of Reference Materials (RM) (6, 7). This approach is only suitable for stable analytes such as metals and organochlorine residues. It has the distinct advantage that a comparison to the RM over the years allows for accurate assessment of analytical quality during long-term trend monitoring. These Guidelines are thus primarily intended for contracts in which stable analytes are determined. For pesticides and other labile compounds similar approaches based on spikes and replicates are very useful in assessing precision, but not bias or accuracy. In such cases, inter-laboratory comparisons can be used to measure bias, providing the analyte stability for the length of the experiment can be assured.

For contracts where there are only a very limited number of samples, such as often occurs with pesticides projects, then only a rough estimate of precision may be made. The approach as circulated to Federal Interdepartmental Committee on Pesticides (FICP) Check Sample Program Coordinators is recommended for this situation (8).
This approach is acceptable since in assessing exposure of wildlife to pesticides, the sampling variability often exceeds the analytical variability.

These Guidelines are, in essence, a special application of the approach outlined for the Toxic Chemicals Program (4). They do not cover field or sampling preparation but address the special quality assurance protocols needed for contracting of analyses. They reflect the main thrust of the CWS component of the Toxic Chemicals Program, of long-term trend monitoring. They will prove useful to laboratories tendering bids for CWS analytical contracts. Users of CWS analytical data will find them useful in assessing the quality of data obtained under contract.

Sources of Uncertainty

Total uncertainty can be defined as the sum of uncertainties from the following sources (9, 10):

1. uncertainty associated with the natural non-uniformity of the sample population;
2. total uncertainty associated with collection strategies (sample size, location, frequency, storage);
3. total uncertainty associated with analytical processing (preparation, aliquoting);
4. total uncertainty associated with the final determination step (i.e., the measurement, data processing).
Uncertainty arising from sources of error 1 and 2 may often be dominant. At times, the uncertainty in these two categories may be an inherent part of the experimental design, that is, not subject to minimization. In this case, some measure of the uncertainty is still valuable because it may influence the precision which must be achieved in the final determination step. The size, frequency and location of sampling required in a field study will be quite different to that of a laboratory study. Storage of samples, particularly with difficult substrates such as water or air, may also present many difficulties.

Therefore, although all sources of uncertainty must be considered carefully in a good experimental design, general guidelines for the first two categories are difficult to set down because each field or experimental situation has its own unique problems. This document will therefore primarily address uncertainty arising from sources 3 and 4 above, which together constitute the determination steps.

**Quality Assurance**

Definitions of quality assurance seem to abound in the literature. One document (4) defines Quality Assurance (QA) as referring "to a total program or activity designed to assure the reliability of data", while Quality Control (QC) "refers to those activities undertaken by the field, laboratory and data management personnel for the attainment of prescribed performance standards". The objective of a QA program should be to identify and measure sources of error in techniques and minimize them in order to achieve the best practical
accuracy and precision of the data (9). It is important to include the qualification "practical" because accuracy and precision are terms which must be put into the context of the data to be obtained. The QC activities would include laboratory control charts, interlaboratory check samples, standard operating procedures, etc. Much of the framework for laboratory QC has been outlined for Long Range Transport of Airborne Pollutants projects (3). Although this document refers to high volume inorganic parameters, the principles are valid for all environmental analyses.

Some aspects are entirely within the control of the contract laboratory management. Thus, in the contract situation, the QA covers such activities as the writing of QA clauses into the contract. An audit of the laboratory's QC procedures, inspection of the data to ensure acceptability and feedback to the contracting laboratory to correct any deficiencies. The QC activities would include laboratory control charts, interlaboratory check samples, GC column performance checks, etc.

"The Principles of Environmental Analysis", as developed by the American Chemical Society (5), contain many of the features necessary to a laboratory QA/QC program. The key factors relevant to analysis done by contract are:

1) the contracting laboratory should have a QC program in which control charts are the basic method for control;

2) the attainment of statistical control must be met before assessment of accuracy can be made;

3) audits should be a feature of quality assurance programs;
4) the acceptability of analytical measurement depends upon rigorous completion of all the requirements stipulated in a properly documented method;

5) interlaboratory check samples are essential for
   a) method validation; and
   b) measuring bias between participating laboratories;

6) reports should contain sufficient information for each analysis including standards, Standard Reference Material (SRM), blanks and replicates to indicate if an analyte was present and, if so, was it above or below the level of quantitation.

Control Charts

The use of control charts is a well established feature of laboratory QC (12, 13) allowing a quick decision as to whether or not a method is in control. Also, control charts provide a pictorial representation of day-to-day variability.

For multi-step methods such as contaminant analysis, two types of charts are useful. These are method controls and instrument controls. The method control is ideally an SRM of like material to that which is being analyzed. This material is analyzed along with samples so that at least one SRM analysis is conducted per batch of samples. Ideally, two SRMs would be employed: one of low level concentration, the other towards the top end of the analytical range. Even if an SRM is not available, a homogenous material can be used. The main constraint is that no analyte or sample degradation should
occur over a long period. The instrument control sample is normally a solution that is prepared directly for injection into a GC or for aspiration into an AA. It is always prepared totally independently of calibration standards and will normally fall between two of the calibration standards.

It is essential that it be understood what the purpose is of each of these control analyses. The method control analysis provides information as to the variability of the total method (extraction, partitioning, derivitization, etc.). When the control value falls beyond the control limits, it is essential to reanalyze those samples. The instrument control analysis provides information as to the variability of instrument (analysis) steps. The sample is analyzed immediately after the instrument has been calibrated and every 20 samples or so in a long sample run. This allows immediate correction of instrument faults, e.g., blocked syringe, wrong calibrating solutions, etc. Extracts can then be reanalyzed from the point when the instrument goes out of control, avoiding complicated correction of results from poor calibration. Further, the instrument control sample is easily prepared and, if stable, can be used over a long period. It is also possible to construct control charts based on differences between replicate analyses (11). Sometimes, this is the only way control can be asserted if no suitable stable SRM can be found. If duplicates are analyzed on a regular basis, it is worthwhile to use the available data to construct the charts as they complement those based on SRM.
Audits

The concept of an audit is one that has been borrowed from accountants. Just as the income tax department does not accept the books of a company without an audit, there is no reason for an agency to accept blindly the results of a contractor. The various steps of a laboratory quality control program have been outlined (12). The major point of an audit is that only documented written procedures can be subject to audit, otherwise the procedure cannot be verified. The audit allows the contracting agency to verify the actual work has been done and performed to written procedures. It verifies that control charts have been prepared and used. It also allows a consultative procedure whereby the contracting agency can inform the contract laboratory of deficiencies or of new procedures so that, after a period, the whole quality of work improves. The audit should never be viewed as a procedure with which to find fault with a contractor's laboratory. Indeed, it is possible for an audit to be conducted by an independent auditor (an external audit) which would audit both the contracting agency's part as well as the contract laboratory. Ideally, such an external audit would cover the whole project from design through to final report.

Detection Limits

This is perhaps the thorniest issue in the environmental laboratory. To the authors' knowledge, there is no reliable and easy method to determine the detection limit for the situation where
samples are limited, the cost of analysis is high and for which there is no reliable blank. The literature assumes either you have infinite time and resources (12) or a true blank (13). Although it is probably not too difficult to determine the detection limit for a simple extraction and analysis method, for a multi-step extraction, derivitization and clean-up followed by analysis, the task can be quite daunting.

It should be observed that most of the detection limits reported in the literature are not statistically derived from actual measurements but are really Minimum Reporting Values (1). These, if applied conservatively, at least give the user some confidence in the data for which a number is reported but inevitably censor out low level data. For environmental contaminants for which trends are desired, this is not acceptable.

The problem of detection levels has been discussed (1, 16), however most approaches place a great burden on the laboratory in determining the detection limit.

We propose the application of the method outlined by the American Society for Testing and Materials in Standard Practice D4210-83 (15). Whilst this procedure is intended for use within water quality laboratories using large data sets, it is applicable to all environmental measurements providing it is recognized that limitations are imposed by having only very limited data points (degrees of freedom). Thus, for contract analyses, it is possible to calculate the Criterion of Detection (CD) and the Limit of Detection (LD) from either
an SRM run several times (at least five) or from duplicate analyses. Since duplicate analyses of samples are usually more readily performed, it is expected that this will be the usual approach. For analytes for which it is difficult to obtain samples in the low range (1 to 10 standard deviations above the CD), an estimate of the CD of an analyte can be made by calculating the ratios of the relative response factors (GC) or sensitivities (AA) to analytes for which CD have been determined. This can only be done where the analyte is carried through the same extraction and analysis procedure. It should be recognized that matrix effects can be dissimilar, especially in the case of heavy metals.

It should be stated that detection limits are established not just to compare sensitivity of methods but for the qualification of low level results.

**Reporting of Low Level Data**

This has been a controversial item among analytical chemists because no chemist wishes to put his name on results for which he has doubts as to the accuracy or precision of the numbers. However, it must be borne in mind that, for monitoring purposes (as compared to regulatory purposes), censoring of data at very low levels can occur both by reporting "less thans" or "ND" as well as by reporting numbers with insufficient significant figures. We propose that, for contract analyses, all results will be reported in a manner similar to that of the ASTM (15) by means of letter codes, W and T, as defined below.
The T code has the following meaning: "Value reported is less than the CD". The use of this code warns the data user that the individual datum with which it is associated does not, in the judgement of the laboratory that did the analysis, differ significantly from zero.

The W code has the following meaning: "Value observed is less than the lowest value reportable under T code". This code is used when a positive value is not observed or calculated for a result. In these cases the lowest reportable value, which is the lowest positive value which is observable, is reported with the W.

Thus, for contract analysis on pesticide residues, a W code will be used wherever an identifiable peak did not appear in the chromatogram. The T code would be associated with positive results with identified peaks in the chromatogram, but only those below the Criterion of Detection.

Since instrument conditions can change from day to day and interference levels are always variable, we suggest that the W and T codes be applied generously. Further, we suggest for the reporting of GC data that the code letter I be used for reporting Interferences at levels above the CD where the analyst knows or suspects interfering substances. Finally, we recommend use of a further code letter, A, whenever a positive peak is reported, but the analyst has some reason to doubt its accuracy, e.g., abnormal matrix or untested method change.
The advantage of codes is that a description of the quality of the data can be maintained in reports and in computer storage. If codes are applied consistently, then data users will become aware of them and will be cautious in interpreting low level or questionable data. It is generally recognized that this practice suggested by ASTM standard does not follow that suggested by the American Chemical Society but we have chosen it as it is more likely to produce statistically useful numbers at lower levels.

The use of significant figures is often not properly understood. Particularly in an age of calculators and computers which produce far too many significant figures, it is essential that proper practices are followed. The basic principle is that more, rather than less, significant figures should be stated. Often the degree of uncertainty of an individual measurement is not known (or can only be crudely approximated). Thus, if sufficient significant figures are provided, any statistical treatment of the results will be valid. If, as often happens, too few significant figures are recorded, then the statistics are based on a censored pool of data. Sufficient significant figures should be provided so that the last digit is probably random. Thus, if results are expressed in ng/g for a method with a Criterion of Detection of 10 ng/g (i.e., a T of 6 ng/g), then results reported to the nearest 10 ng/g will be censored, but those reported to the nearest ng/g will probably have some randomness in the last figure. Generally, four significant figures are sufficient and three normally insufficient.
Negative results can occasionally be encountered. This can happen in flame AA due to either negative interferences or from the results of the method of standard addition. It is less likely to occur with GC but can occur if a blank correction is applied. As the ASTM (15) suggests negative results should always be reported, normally with the T code.

The Acceptability of Results

In any contract it is only fair to the contracting laboratory and to the final user of the data that well thought out criteria are used to decide if a batch of results is acceptable or not. These criteria cannot be based on single samples but only on clearly defined statistics. Thus, it is essential that, in assessing the performance of a contract laboratory for a given batch of samples, a set of historical data is used for comparison. This data should be comparable as to method used, analyte concentration and matrix. Normally, good laboratories can achieve consistent precision by use of control charts, the problem comes in assessing bias (or accuracy) if no suitable SRM is available. Obviously, in such cases it may be more difficult to compare laboratories.

The Canadian Wildlife Service (CWS) has tested several criteria for acceptability of contract results, mainly based on RMs. The RMs have either been developed in CWS for the express purpose of contractor QA or are purchased from suppliers of RMs, e.g., the National Bureau of Standards. In the case of CWS RMs, the reference values and s.d. are based on at least 30 separate determinations.
The criteria used are systematic error (SE), total error (TE), acceptable ranges based on multiples of s.d., and maximum coefficient of variation (CV max).

The SE criteria (or percent relative bias) is defined as

\[
SE = \frac{\bar{x} - x_{\text{ref}}}{x_{\text{ref}}} \times 100
\]

where \( \bar{x} \) = mean of replicate SRM analyses

\[ x_{\text{ref}} \] = true or reference value

Normally, the SE must be less than 50%. The criterion of 50% is based on the observation by Elgar (16) that, in interlaboratory studies with organochlorine pesticides at low levels, rarely is there agreement below a 25% SE. Thus, a limit of 50% provides a suitable tolerance at very low levels when the absolute error (or bias) is acceptable but the percentage bias can be quite high.

The total error, TE, provides a value of the effects of both SE and precision

\[
TE = SE + 2 \times SD \times 100
\]

This concept proposed by McFarren (14) sets an arbitrary value of 50% as the maximum acceptable total error. The TE criterion would consider acceptable a set of analyses of an RM with a CV 15% and SE 20%. As applied to this situation with a single laboratory being judged, both the CV and SE would be quite good.
The third criterion of acceptable ranges based on multiples of standard deviation is very useful when there is a history of analyses. We have chosen ±2 s.d. for a set of at least five RM analyses in a batch of samples. Experience has shown that this can be obtained for residue analysis by a single laboratory.

The CV\textsubscript{max} criterion is based on the work of Frehse and Timme (18). Briefly, it has been observed that, as analyte concentrations decline, the CV increases. From this, they developed a concept of "first category measured value curves".

This can be extended to the general use

\[
CV_{\text{max}} = CV_0 \left| \frac{x_0}{\bar{x}} \right| \log f
\]

where \(CV_{\text{max}}\) = maximum permissible CV

\[
CV_0 = \text{CV at the detection limit}
\]

\[
x_0 = \text{detection limit}
\]

\[
\bar{x} = \text{mean of replicate analysis of a blind sample}
\]

\[
f = \text{the factor by which } CV_{\text{max}} \text{ diminishes per order of magnitude}
\]

For example, if \(f = 2\) and \(CV_0 = 100\), the equation is:
For pesticide residue analysis, the form of the equation

\[ CV_{\text{max}} = \left| \frac{x_0}{\bar{x}} \right|^{0.3} \times 100 \]

0.3

based on \( x_0 = 0.01 \text{ ppm}, CV_0 = 100\%, \text{ and } CV_{\text{max}} = 15\% \text{ at } 1 \text{ ppm}. \) This is equivalent to \( f = 2.7. \)

These equations allow \( CV_{\text{max}} \) to increase as concentration declines in a systematic manner. A practical lower limit of \( CV_{\text{max}} \) is 15%. Normal practice has been to determine the value of \( CV_{\text{max}} \) in our own laboratories by use of SRMs or by replicate analyses.

These criteria have been grouped so that the data manipulation can be performed by computer and the data sets failing to meet the criteria can be identified. Thus, acceptability can be based on:

a) \( TE < 50\% \)

b) \( M_{\text{ref}} < M \pm 2SD \) where \( M_{\text{ref}} \) is contract mean of reference material and \( M \) is the established mean of reference material

c) \( SE < 50\%, CV < CV_{\text{max}} \)

An acceptable result would be positive in all cases. Negative results are compared to previous data sets and judged to see if there is consistent bias, and if previous contracts had similar problems in the past, etc.
References


APPENDIX I

Suggestions for Additional Reading

1) Principles of Quality Assurance of Chemical Measurement.
John K. Taylor, National Bureau of Standards, Center for
Analytical Chemistry, Gaithersburg, MD 10899 (NBSIR 85-3105).
1985.

John K. Taylor, National Bureau of Standards. NBS Special

States Department of Agriculture, Food Safety and Quality

4) Manual of Analytical Quality Control for Pesticides and Related
Compounds in Humans and Environmental Samples by J. Sherlock.
Edited by Randall R. Watts, U.S. Environmental Protection Agency,

Chemistry in Britain, p. 1019, November 1985. Correspondence,


Proposed Contract Clauses for QA/QC

These clauses are intended for contracts in which the analytes are metals or organochlorine compounds. Suitable clauses in lieu of clauses 7, 8 and 9 will be negotiated for other analytes.

1. The contractor is expected to participate in appropriate intra-laboratory check sample programs in a timely manner and report his results to the Scientific Authority.

2. The contractor should participate in appropriate external quality assessments continuously to establish their credibility.

3. The contractor shall maintain records of quality assurance activities and make these available to the Scientific Authority and project manager on request.

4. The proposed handling, storage, preservation procedures and analytical methodologies shall be approved by the Scientific Authority and project manager before work is initiated.

5. All laboratory work, including spikes, blanks, replicates, controls, sample preparation and data reduction, will be subject to on-site inspection and audit.

6. For each batch of samples, every tenth sample will be analyzed and reported in duplicate at no extra cost.

7. For each batch of samples, a standard reference material will be analyzed five times, or every fifteenth sample, whichever is the greater, and reported at no extra cost, provided the total batch
size is greater than 75. For batches less than 75, a surcharge will be added as per fee schedule.

8. The Reference Material will be supplied by the Scientific Authority if no suitable commercial RM is available. Commercial RMs are to be supplied by the contractor.

9. For each batch of samples, a minimum of RMs (as characterized by CWS or NBS) will be included in the batch and the whole batch blind numbered either by CWS or by the contractor's quality awareness officer. For each residue or metal for which a reference value has been determined, the following conditions for the mean of the replicate analyses of the blind samples will be met:

a) the systematic error of the mean is less than 50%

b) the coefficient of variation, CV of the mean, is less than a limit described by the following:

1) for pesticide residues

\[ CV = 13.7 \times \frac{X_r}{x} \]

or 15%, whichever is the greater; where \( X_r \) is the reference material residue level, in mg/kg wet weight for each residue, as determined by CWS.

2) for trace metals

\[ CV \left| \frac{x_o}{x} \right| 0.3 x 100 \]

or 15%, whichever is the greater; where \( x_o \) is the detection limit of the particular metal, in that matrix.
10. For each batch of samples, calibration curves or a table of response factors for each reported residue covering the whole range of reported values for that residue must be included in the report.

11. For each batch of samples, example chromatograms, if appropriate, of standards and contractor's internal reference material and typical samples (all fractions analyzed) must be included in the report along with a basic description of the analytical method (extraction and analysis).

12. Original chromatograms (or other output), if appropriate (or digitally stored data sufficient to regenerate the original chromatograms) must exist for all analyses, and be retained by the contractor unless otherwise authorized in writing by the Scientific Authority. At the time of dispositions, CWS shall have the right to take possession of all chromatograms which the contractor wishes to discard.

13. Control charts, both method and instrument, covering the period during which the samples were analyzed must be provided.

14. All results, including controls, RMs, duplicates, spikes and blanks will be presented in tabular form. The dates on which each sample was extracted and analyzed must be included. All results must be made available in an acceptable IBM PC format on 5¼" floppy discs.
15. All low-level results will be reported according to the protocol outlined in ASTM Standard D-4210-83. Criteria of detection will be calculated for each residue or metal in accordance with the procedure in the same ASTM Standard. Additionally, all results which may have been subject to chromatographic interference will be prefixed I and those whose numerical accuracy is in doubt for any other reason will be prefixed A. Metals which were in high concentration and had to be diluted will be prefixed D.

16. At least one sample as determined by the Scientific Authority will be analyzed (if appropriate) by GC-MS or other technique to confirm identity of all measured residues.