International Joint Commission (IJC) Digital Archive

International Joint Commission

1980-11-01

Data Quality Assessment: Report of the Data Quality Work Group to the Water Quality Programs Committee of the Great Lakes Water Quality Board

Great Lakes Water Quality Board. Water Quality Programs Committee. Data Quality Work Group

Follow this and additional works at: https://scholar.uwindsor.ca/ijcarchive

DATA QUALITY ASSESSMENT

REPORT OF THE
DATA QUALITY WORK GROUP

TO THE
WATER QUALITY PROGRAMS COMMITTEE
OF THE
GREAT LAKES WATER QUALITY BOARD

NOVEMBER 1980
WINDSOR, ONTARIO
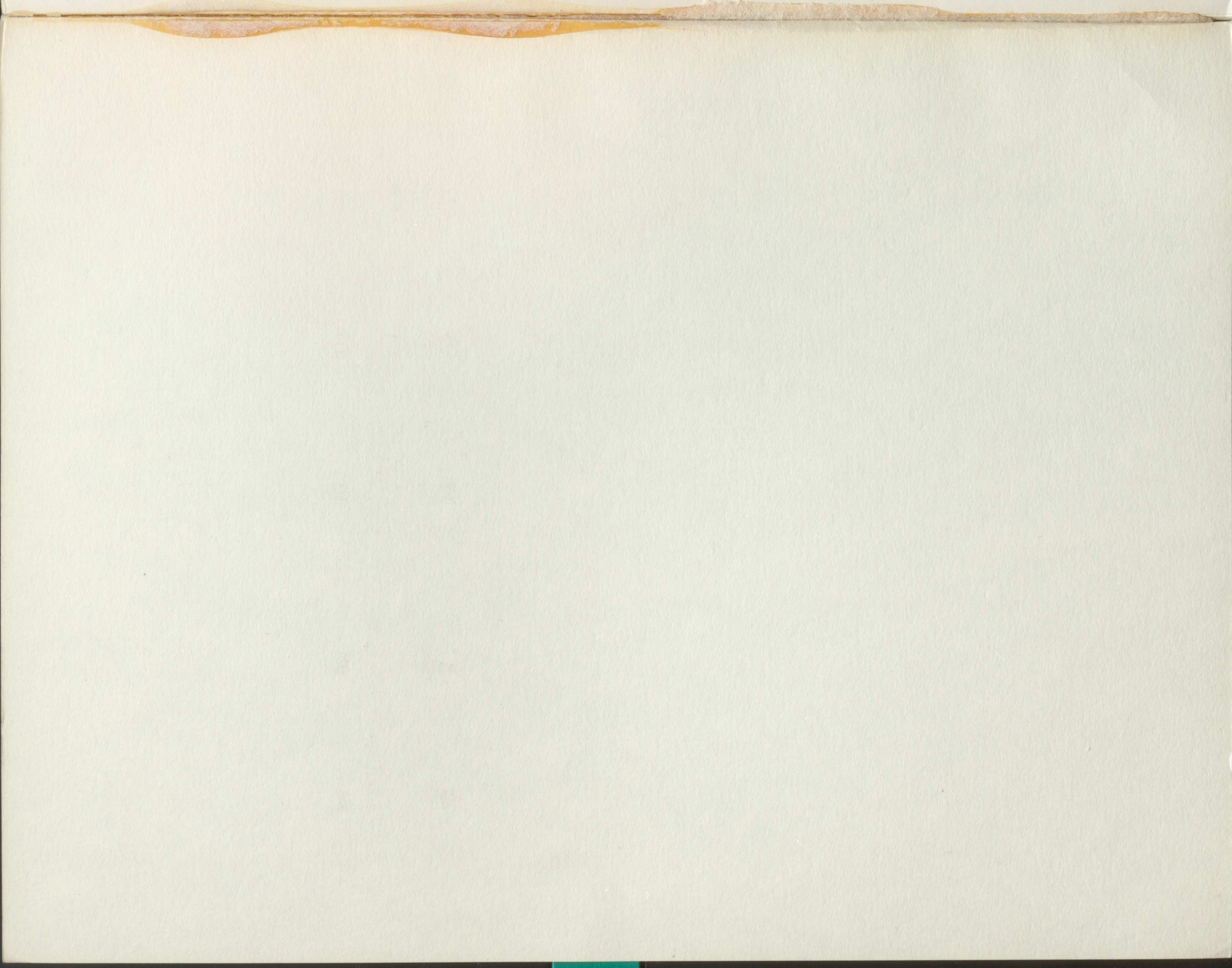
INTERNATIONAL JOINT COMMISSION

GREAT LAKES WATER QUALITY BOARD
DATA QUALITY WORK GROUP OF THE WATER QUALITY PROGRAMS COMMITTEE
100 OUELLETTE AVENUE, 8TH FLOOR
WINDSOR, ONTARIO, CANADA N9A 6T3

November 1980

Water Quality Programs Committee of the
    International Joint Commission
    Great Lakes Water Quality Board

Ladies and Gentlemen:

The Data Quality Work Group, in partial fulfillment of its responsibility,

is submitting the following report on the activities of the Work Group.

Respectfully submitted,

K. I. Aspila

K. I. Aspila
Chairman
Data Quality Work Group

PREFACE

Statements and views presented in this report are those of the Data Quality Work Group and do not necessarily reflect the views and policies of the International Joint Commission, the Great Lakes Water Quality Board, or the Board's Water Quality Programs Committee.

TABLE OF CONTENTS

TABLE OF CONTENTS

SUMMARY
ACTIVITIES AND RECOMMENDATIONS

<u>ACTIVITIES:</u>

For the period 1979-80, the Work Group has developed and distributed recommended guidelines for intralaboratory quality control; endorsed and distributed recommended procedures for reporting low level data; conducted an analytical chemists' meeting to foster an understanding of and commitment to the Great Lakes International Surveillance Plan and the need for adequate data; conducted 14 interlaboratory tests (4 still in progress); evaluated laboratory performance for total phosphorus in water from 5 interlaboratory studies; and reviewed phosphorus methods for comparability.

For phosphorus measurements, some of the laboratories providing much of the Great lakes phosphorus data have compared well. By comparison, of the remaining laboratories, those that have produced less data for Great Lakes studies and surveillance have often demonstrated less competency in making such measurements. Some laboratories have participated in only a minimal way and cannot be appraised on performance.

For the other interlaboratory tests, excluding the PCBs sediment study, laboratories have performed similarly to the findings for phosphorus. For the PCBs in sediment and ampuls study, almost all demonstrated low recoveries in sediment, but many laboratories demonstrated a high degree of precision.

<u>RECOMMENDATIONS:</u>

The Work Group's ability to fulfill its responsibilities depends not only on its own activities but on full participation, including funds and manpower, from the jurisdictions. Agency management within these jurisdictions has a responsibility to support Work Group quality assurance endeavors in order to achieve sufficient laboratory analysis, field sampling, and data assessment comparability.

To meet the various goals which may be subsumed under the heading of assuring data quality and comparability, the Work Group recommends:

o   that the Quality Assurance and Methods Section of the Division of Analytical Methods at the Canada Centre for Inland Waters be formally recognized as the principal laboratory for preparation, storage, and distribution of interlaboratory study samples for the Work Group;

o   that the responsibility for providing interlaboratory study samples be equitably shared by the appropriate agencies in both countries through transfer of funds and/or provision of personnel;

o   that meetings of analytical chemists be held on at least an annual basis to foster understanding of the Agreement and the role laboratories will play in its fulfillment, and to identify and resolve common problems;

o   that agencies and the appropriate bodies within Agreement institutions identify data users so that those responsible for preparing reports may be properly involved in data quality concerns;

o   that either uniform techniques or techniques shown to be equivalent be used in the taking of samples;

o   that participation in Work Group interlaboratory studies be made mandatory for all laboratories providing environmental data for the assessment of contamination in the Great Lakes system so that biases may be detected and resolved;

o   that all major sewage treatment plants providing loading information be included in Work Group studies which are specifically designed to provide samples at typical effluent levels; and

o   that agencies accept the responsibility of having laboratories participating in International Joint Commission programs implement the recommended Guidelines for Intralaboratory Control.

TERMS OF REFERENCE

The following Terms of Reference were approved by the Great Lakes Water Quality Board on July 15, 1980.

Under the direction of the Water Quality Programs Committee, the Data Quality Work Group is responsible for assessing the quality of data reported in support of the Great Lakes Water Quality Agreement and, as necessary, providing advice about needed improvement in data quality.

Specifically the Data Quality Work Group will:

1. Recommend the quality assurance requirements for field, laboratory and data management activities undertaken in support of the Agreement and monitor the meeting of the requirements.

2. Conduct interlaboratory studies and, as needed, assist participants in implementing remedial action.

3. Compile and update information on the analytical characteristics of methods used by participants in interlaboratory studies. Evaluate and recommend necessary changes to these methods if such methods are not comparable.

4. Compile and update information on sample collecting and handling procedures used by organizations providing data in support of the Agreement. Evaluate and recommend necessary changes to those procedures if such procedures yield non-compatible data.

5. Report findings and recommendations to the Water Quality Programs Committee.

6. Respond to specific requests of Water Quality Agreement institutions.

Although the new Terms of Reference confer upon the Work Group some
additional responsibilities which are only now being initiated, activities to
assess laboratory performance, to improve the quality of reporting by
laboratories, and to make recommendations have been in progress for several
years.

Highlights of efforts made over this past year include the adoption and
dissemination of recommended guidelines for the control of analytical
procedures in an intralaboratory control program; the affirmation and
circulation of guidelines for reporting low level data including results of
less than "zero"; a two-year assessment of laboratories' performances for the
measurement of total phosphorus in water; the evaluation of laboratories'
performances by interlaboratory studies for several constituents through 5
separate studies; and the holding of an analysts' meeting to foster a better
understanding of Work Group functions and the data requirements of the Great
Lakes International Surveillance Plan.

A brief discussion of the listed highlights follows:

GUIDELINES FOR THE CONTROL OF ANALYTICAL PROCEDURES IN AN INTRALABORATORY PROGRAM

The Work Group, at its Analytical Chemists' meeting of February 27-28,
1980, presented recommended procedures with examples for an intralaboratory
quality control program. The Guidelines assume that analytical methods used
are suitable for the task at hand, that field procedures for taking and
transporting samples are adequate, and that laboratory quality assurance
responsibilities including adequate reporting to managers are in place.

The Guidelines describe how to estimate analytical procedure variability
using duplicate analyses or stable standards, how to test for change in
analytical variability and how to pool estimates; all these procedures are
aimed toward setting control limits on a Shewhart type control chart. The
goals of setting control limits are stressed; i.e., they should be close
enough to signal when there is trouble with a system, but distant enough to
discourage tinkering with a system that is operating within its capability.
These Guidelines are appended.

The Data Quality Work Group also recommended the following minimum frequencies for the use of control samples:

To monitor accuracy: 1 quality control sample of known value should be included with every 15 analyses or with each batch, whichever results in the greater frequency.

To monitor precision: 1 quality control sample should be included with every 15 analyses or with each batch, whichever results in the greater frequency. If duplicates are used to monitor precision, they should be analysed in different runs when a between run measure of variability is employed in setting control limits.

REPORTING LOW LEVEL DATA

The Work Group endorsed and later distributed to Great Lakes analysts a revised portion of the PLUARG Quality Control Handbook for Pilot Watershed Studies - Reporting Low Level Data. The distributed material provides an explanation of Type I and Type II errors, urges chemists to use codes in reporting low level data rather than ambiguous "less thans," defines the criterion of detection, and developes a rationale for reporting all results including findings of less than zero.

Further, the discussion illustrates the danger of the analyst censoring low level data, thereby causing high biases and providing information which is useless in drawing valid inferences from surveillance data. This full discussion, Reporting Low Level Data, is appended.

ANALYTICAL CHEMISTS' MEETING

On February 27-28, 1980, the Work Group held an analysts' meeting at the Canada Centre for Inland Waters, with the National Water Research Institute as host. About 100 analysts from throughout the Great Lakes Basin attended. The meeting was intended for chemists and technicians who actually perform analyses and for their immediate supervisors.

The principal purpose of the meeting was to assist representatives of laboratories supporting the Great Lakes International Surveillance Plan (GLISP) in providing adequate data to meet Plan objectives. While attendees were somewhat acquainted with Work Group goals through participation in round robin studies, the meeting provided an opportunity for the Work Group to make them explicit. It also allowed analysts to question and discuss Work Group procedures. More important, participants had an opportunity to discuss common problems and possible solutions.

The opening sessions of the meeting included presentations on the International Joint Commission and the Great Lakes Water Quality Agreement of 1978, the GLISP, and the quality control activities and recommendations of the Work Group. After the opening sessions, the attendees split into three analytical task groups depending on individual interests.

The task groups were Major Ions, Nutrients, and Physical Measurements in Water; Metals in Water, Sediment, and Biota; and Organics in Water, Sediment, and Biota.

After the task group meetings, participants assembled for a summary session of task group reports and general conclusions. Although the summary task group presentations did not necessarily present consensus information, the following specific points for management and the Work Group to consider were brought forth:

- Many laboratory chemists are unaware of the purposes of various programs including IJC surveillance work and therefore there is a need for better communication, including demonstration of a data quality requirement and overall usefulness of data collection.

- Participation in Data Quality Work Group round robin studies is not universally viewed as mandatory, but more as an educational process to assist poor performing laboratories in identifying the cause(s) of poor performance including the need for additional laboratory personnel, equipment or better methods.

INTERLABORATORY PERFORMANCE STUDIES

The The Work Group is convinced that the only way to determine if laboratories produce comparable results is through interlaboratory studies.

The Work Group has conducted ten such studies over the past two years with natural and spiked samples with three studies also using ampul references. In addition to these, four studies are now in progress.

The studies are:

1978-79: Study #21 - Major Ions, Trace Metals and Nutrients in Water

Study #22 - Major Ions and Nutrients in Water

Study #23 - Trace Metals in Water

Study #24 - Total Phosphorus in Water

Study #25 - Reactive Silica in Water

1979-80: Study #26 - Arsenic and Selenium in Water

Study #27 - Major Ions, Nutrients and Physical Measurements in Water

Study #28 - Total Phosphorus in Water

Study #29 - Trace Metals in Water

Study #30 - PCBs in Ampuls and Sediments

Study #31 - Metals in Fish (in progress)

Study #32 - Major Ions, Nutrients, and Physical Measurements in Water (in progress)

Study #33 - Total Metals in Water (in progress)

PCB-F-#1 - PCBs and Chlorinated Hydrocarbon Pesticides in Fish and Ampuls (in progress)

The following general procedures were used for these studies: The laboratories to participate were identified based upon the kinds of analytical data they would supply to the Surveillance Program. For each study, transmittal letters were enclosed with the test samples. Also included were forms for reporting results, tables listing the expected range of sample levels, special instructions and information on study sample preparation, and a methods questionnaire.

PROCEDURE USED FOR THE EVALUATION OF RESULTS

Ranking Results

Results were ranked and ranks totalled for each laboratory's results on a specific characteristic. Totals which were unduly low or high were interpreted as an indication of biased results, that is, consistently high or low analyses. Testing was done at an approximate α level of 0.05 so there was about one chance in twenty that an unbiased set of results would be deemed biased.

Flagging Results

The rationale underlying the classification of certain results as errant was as follows: If several participating laboratories demonstrated the ability to achieve a certain level of performance, then all participating laboratories should have been able to achieve that performance level, and results that are outside the achievable level of performance were judged errant.

Specifically, the median of all reported values on a sample was used as a target value, and a basic acceptable difference from the target value was determined from the performance of the participating laboratories. The basic acceptable difference was increased for samples of higher concentrations. Reported values which were more than the acceptable difference from the target value were flagged, and doubly flagged if they were more than one and a half times the acceptable difference from the target value.

Each laboratory received an individual report on its performance. The report consisted of suggestions for areas of improvements, a list of all data and their ranks, individual flags, overall study conclusions, and information on study sample preparation (not in all cases).

Findings for Studies #21 through #25 are reported in Great Lakes Water Quality Report of 1978, Appendix B, Surveillance Subcommittee Report, and will not be repeated here. However, with emphasis on phosphorus loadings to the

- 8 -

Great Lakes there has been the need to evaluate laboratory performance over time particularly for this constituent. Five interlaboratory studies have been completed to date, and one is planned for later this year.

LABORATORY EVALUATIONS FOR PHOSPHORUS IN WATER

Laboratory performances for phosphorus in Water over these past two years have been evaluated and are reported here.

STUDY DESCRIPTIONS AND SUMMARY OF FINDINGS

Study #21 - Major Ions and Nutrients in Water, January 1978

Two water samples and four ampul concentrates were distributed for total phosphorus measurements. Fifteen laboratories reported results. The samples being tested ranged from 0.003 to 0.016 mg/L. Most laboratories performed adequately and the median value obtained by all laboratories' data for each sample agreed well with its corresponding design value.

Study #22 - Major Ions and Nutrients in Water, May 1978

Six natural water samples and two suitably spiked distilled water samples were distributed. The natural waters were mixtures of Lake Ontario tap water and Hamilton Harbour water. Seventeen laboratories reported results. The samples ranged from .001 to .09 mg/L. Several laboratories displayed difficulties in analyzing such low level samples and two laboratories clearly reported exceptionally high results indicating a probable contamination problem.

Study #24 - Total Phosphorus in Water, September 1978

Fourteen samples were prepared for this study. The samples comprised of rain water, natural waters from lakes Superior, Erie, and Ontario, selected nearshore waters, a distilled water blank, and spiked distilled water. The samples ranged from 0.0025 to 0.090 mg/L. Seventeen laboratories reported results.

Improvements relative to Studies #21 and #22 were noted by laboratories.
Some laboratories failed to report results within the capability of the
method, apparently an arbitrary decision. Most laboratories were able to
analyze the samples adequately. Three laboratories' results were deemed
biased and three laboratories produced erratic results. However, the
principal laboratories contributing data to open lake and neashore
surveillance information were in agreement.

Study #27 - Major Ions and Nutrients in Water, November 1979

Twelve samples were prepared for this study. The samples comprised of
rainfall, lakes Huron, Ontario, Erie and Michigan waters, Ottawa and Fraser
river waters, and laboratory reference waters. The samples ranged from 0.002
to 0.85 mg/L. Twenty laboratories reported results.

Several laboratories displayed difficulties. Two laboratories were deemed
to be biased low and four biased high.

Study #28 - Total Phosphorus in Water, August 1979

Fourteen samples were used for this study. The samples were prepared at
the same time Study #24 samples were prepared. Other than how the samples
were coded, they were identical to those used for Study #24.

Median results on each sample as derived from the test results were
compared with Study #24 median values by a two-tailed t-test using the
differences ($\Sigma = 6.7$). The null hypothesis of no difference between means
for the medians reported in the two studies was not rejected when tested at an
α-level of 0.1. It appears that the test samples are stable and
laboratories can collectively agree on a median value, over a period of time.

Laboratory performance in Study #28 was slightly better than Study #24.
The criteria used for flagging in Study #28 was somewhat more stringent. A
table summarizing laboratory performance by laboratory is presented below:

TOTAL PHOSPHORUS IN WATER
SUMMARY EVALUATION BY LABORATORY*

Lab. No.	Study No. 21	22	24	27	28
1	S	S-1FH	S	S-1FH	S
2	S	S	S	S	4FL-LB
3	S	NR	S-1FH	4FL,1FH	S
4	VHB	VHB	NR	5FL,1FH	3FH,14FL
5	S	S	S	4FL-LB	S
6	S	S	2FH	1FL,1FH	2FL
7	S	NR	NR	10FH-HB	6FH-HB
8	S	S-1FH	LB	7FH-HB	S-1FH
9	S	S	S	S	5FH-HB
11	S	NR	NR	8FH-HB	NR
12	Q	Q	6FL-LB	1FL,6FH	5FH,13FL
14	S-1FL	S-1FH	S	S	2FH
15	S	NR	NR	NR	7FH,1FL
16	1FH	1FH	2FL	S	S-2FH
17	S	NR	4FH-HB	1FL,3FH	3FH,1FL
18	S	DNP	DNP	DNP	DNP
20	DNP	S	HB	2FH,1FL	S
22	DNP	NR	3FH,1FL	NR	5FL,4FH
23	DNP	Q	NR	NR	NR
24	DNP	3FH	HB	6FL-LB	6FH-HB
25	DNP	Q	S-1FL	2FL,2FH	DNP
26	DNP	NR	S	S	S
27	DNP	5FH	3FH,3FL	10FH,1FL	10FH-HB
31	DNP	DNP	DNP	1FL,1FH	1FH,5FL-LB

*Full details on individual studies are available from the Data Quality Work
Group on request.

Key: DNP - did not participate
      NR - received samples but never reported values
      HB - high bias
      LB - low bias
      XFL - sum of low flags
      XFH - sum of high flags
      Q   - method insufficient for the test at hand
      S   - satisfactory
      VHB - very high bias

- 11 -

To conclude: the Work Group believes that through conducting these past 5 studies for total phosphorus there is sufficient evidence to draw some conclusions on a laboratory's performance as compared with others. Laboratories 1, 2, 5, 9, 14, 16, and 26 have agreed most often in producing comparable data. Laboratories 4, 7, 12, 17, 22, 24, 25, 27, in three or more studies have shown to be errant in their phosphorus measurements. It is disheartening that so many laboratories have demonstrated difficulty with the measurement of total phosphorus in water, and that performance information on some laboratories is so sketchy. A few laboratories which have provided much of the open lake and nearshore phosphorus measurements have performed well in these described interlaboratory tests. However, the Work Group is cognizant that interlaboratory tests may not be "blind" tests at the bench level and therefore may provide the best a laboratory can do rather than its typical work. Only a well planned and implemented program of intralaboratory quality control will demonstrate day-to-day competent laboratory performance.

METHODS FOR PHOSPHORUS MEASUREMENTS

For several years the question of methods differences has been raised relative to the two types of reducing agents, stannous chloride and ascorbic acid, that are used to develop the color of phosphorus in water measurements. The Work Group has been requested to evaluate as to whether there is a bias present between the two reduction procedure types. Preliminarily, from comparing data derived from the described 5 interlaboratory tests and an additional 184 paired data reported by one laboratory would indicate no statistically significant difference. One method type is somewhat more sensitive and thereby provides more measurable variability at very low levels, but also more accuracy. The other procedure suffers less from some possible interferances such as salts in water. For Great Lakes' waters either procedure appears suitable if great care is taken in its use.

LABORATORY EVALUATIONS FOR CONSTITUENTS OTHER THAN PHOSPHORUS

Although the Work Group has particularly attempted to get information on laboratory competency for the measurement of phosphorus as quickly as

- 12 -

possible, it has not neglected interlaboratory studies for other constituents. Four tests for major ions, nutrients (other than phosphorus) and physical measurements in water, three tests for trace metals in water, one specific test each for silica, arsenic, and selenium, and other tests including trace organics have been completed. Descriptions of some of these tests follow:

Study #26 - Arsenic and Selenium in Water
April 1979, 16 Laboratories Participated

Study #26 comprised of 15 samples derived from waters of the Hamilton Harbour, distilled water, and various spike combinations of each. Expected ranges for the samples were from ≃ 1 to 100 μg/L for each constituent.

Those laboratories that used direct measurement by carbon rod or graphite furnace atomization found the six samples that were preserved with 0.2% v/v sulfuric acid to be quite troublesome for selenium measurements. Also, some laboratories using direct measurement through nickel nitrate addition found their results to be suppressed on these samples to the extent that some reported only "W" codes, essentially no selenium. Others obtained results that approached those laboratories using a hydride reduction to hydrogen selenide technique. One laboratory indicated that it made up acidified standards (.2% H_2SO_4) to compensate, and their results were found to be closest to the hydride results. Upon inquiry on what was different for some using carbon rod or graphite furnace to get more comparable results with that of the hydride technique, it was suggested that using electrodeless discharge lamp (EDL) sources and increasing ashing time help overcome the suppression from H_2SO_4.

Some analysts remarked that extra efforts were required because of the acid addition. The decision to acid preserve those six samples was because they were freshly prepared. Experience has shown that freshly prepared samples deteriorate at a fairly fast rate and then, after time, somewhat stabilize. The other samples had been prepared sometime ago and had therefore become suitable for use without a preservative.

- 13 -

The choice of sulfuric acid rather than nitric acid was based upon the known difficulty in using some hydride procedures in the presence of nitric acid. More than half the laboratories used a hydride procedure.

In future interlaboratory studies for which selenium is to be measured, hydrochloric acid will be used as a preservative since it is believed that it will not interfere with either technique.

With rather generous limits to escape flagging and removal of some laboratories from the data set, three laboratories received no flags, three received 1 flag, two received 2 flags, one received 5 flags, one received 12 flags, and one received 13 flags.

Study #27 - Major Ions, Nutrients, and Physical Measurements in Water
June 1979, 23 Laboratories Participated

Study #27 comprised of 12 samples prepared from rainfall, lakes Huron, Ontario, Erie, and Michigan waters, Ottawa and Fraser river waters, and laboratory reference waters. Sample constituents levels were designed to represent open lake, nearshore, and some tributary waters.

Results were reported for calcium, magnesium, sodium, potassium, chloride sulphate, pH, alkalinity, hardness, total phosphorus, nitrate and nitrite, total Kjeldahl nitrogen, reactive silicate, fluoride, specific conductivity, total organic carbon, total inorganic carbon, organic nitrogen, total nitrogen, and ammonia.

The findings for total phosphorus were earlier discussed in this report under the section on Laboratory Evaluations for Phosphorus.

For calcium, of the eighteen laboratories reporting 6 escaped flagging altogether, three labs had only single flags, and four labs had double flags. One laboratory was judged to be biased high and no labs were judged as biased low. Generally most laboratories have demonstrated a competency for measuring calcium in the study sample range of 0.5 to 61 mg/L.

Of the eighteen laboratories reporting on magnesium, seven escaped flagging altogether, three labs were doubly flagged and several labs had several flags. Three labs were judged to be biased low.

For sodium, of the twenty laboratories reporting, seven labs escaped flagging, one laboratory received 1 flag, while twelve laboratories received 3 or more flags. Two laboratories were judged to be biased low while one was judged high.

Of the nineteen laboratories reporting results for potassium, four labs escaped flagging, three labs were doubly flagged, and the remainder 3 or more flags. One laboratory was judged biased low and two were judged high.

For chloride, of the twenty laboratories reporting, eight escaped flagging, one was singly flagged, and three were doubly flagged. Other labs received 3 or more flags and two labs were judged biased high, while one biased low.

Results for the other constituents follow the pattern described above. It should be noted that for the most part those laboratories that escaped flagging for a few constituents also generally did well for the others. Put another way, a laboratory that performs well for a few tests generally does well overall. Also on the whole, the laboratories that contributed the bulk of the data for the GLISP were also the laboratories that produced the most comparable data for this interlaboratory study.

Study #29 - Total Metals in Water
February 1980, 20 Laboratories Participated

Study #29 comprised of 12 samples of rain water, lake waters, and laboratory standards. The metals to be measured were aluminum, cadmium, cobalt, chromium, copper, iron, manganese, molybdenum, nickel, lead, vanadium, and zinc.

Many laboratories had difficulty with the low level samples and some had difficulties with all samples. For aluminum, two labs were judged as biased high and one laboratory was judged low among the thirteen labs reporting. For cadmium, of fourteen reporting labs, two labs were judged as biased high and two as biased low. For the thirteen labs that reported for cobalt, three labs were judged as biased low and two as biased high. The above examples are typical throughout for the other constituents.

Of the samples distributed, sample #8 was of near blank and low in organic content, while sample #10 was low in metalic content but high in organic chemical content. To provide some indication of whether the sample matrix affects the results, sample #8 water was spiked to 1/20 of sample #1 to provide sample #9 and sample #10 water was spiked also 1/20 of sample #1 to provide sample #11. Therefore, if little difference in recovery is observed for the two sets, matrix effects would appear minimal. Results follow:

DIFFERENCES FROM MEDIAN VALUES (μg/L)

Metal	"Expected"	(9-8) Low Organics	(11-10) High Organics
Al	106.3	110	103
Cd	4.4	4.0	4.2
Co	23.33	23.6	24.1
Cr	4.5	4.54	3.10
Cu	5.24	5.45	5.00
Fe	10.2	10.0	7.0
Mn	6.50	4.0	7.0
Mo	32.5	31.0	32.5
Ni	27.25	27.0	30.2
V	43.45	43.5	41.5
Zn	5.825	1.0	5.0
Pb	18.5	15.0	16.0

From comparing the recovery differences in median value for the sample with high organics with the sample containing low organics little if any matrix effect appeared present. Future studies of the Work Group will include a similar technique to gather further evidence on sample matrix effects.

Sample #1 of the study was also used in another study among laboratories in which all but three are not Great Lakes International Surveillance Program participating laboratories. A comparison of median values obtained is given below. Between twelve and twenty-nine laboratories provided results for the individual constituents on the other study.

MEDIAN VALUES (μg/L)

	Al	Co	Cd	Cu	Fe	Mn
Study #29	2130	463	87.7	105	200	131.5
Other Study	2100	480	94.6	100	200	130

	Mo	Ni	Pb	V	Zn
Study #29	610	550	375	885	181
Other Study	655	580	360	896	120

Although some laboratories reported scattered data for these studies, median results from the two studies demonstrated adequate overall agreement.

Study #30 - PCBs in Sediments and Ampuls
April 1980, 21 Laboratories Were Submitted Study Samples

This study comprised of 4 ampuls and 8 sediment samples. All ampuls and 7 of the samples were spikes from a standard solution of an equal mixture of PCB Aroclors 1254 and 1260. One bulk blank sediment was used as the basis of spiking with 1 subsample of it reserved in each set as a blank. This design allowed for comparison of the participant's standards with the ampul standards plus recovery information on the sediment samples.

Comparison of the medians to the target values of the sediment samples gave an average recovery slightly above 70%, suggesting wide spread difficulty with recoveries. Since there appeared to be recovery problems, all flagging was done from target values rather than medians. However, it was noted that several of the participating laboratories' results exhibited a commendable degree of within laboratory precision.

Of the nineteen laboratories reporting, three labs escaped flagging. Two labs were judged to be biased high and two were judged low. Some laboratories had difficulting in identifying that the ampuls and sediments contained equal mixtures of Aroclors 1254 and 1260. However, most provided a reasonable approach to quantifying what was present.

As an example of an interlaboratory study evaluation, the findings on Study #30 that were reported to the participating laboratories are attached to this report as an appendix.

Also attached is a copy of the recommended guidelines for control of analytical procedures in an intralaboratory quality control program and a discussion on reporting low level data.

INTERLABORATORY STUDIES IN PROGRESS

Study #31 - Metals in Fish; Study #32 - Major Ions, Nutrients and Physical Measurements in Water; Study #33 - Total Metals in Water; and Study PCB-F-#1 - PCBs and Chlorinated Hydrocarbon Pesticides in Fish and Ampuls, are in progress and will be reported on at a later date.

CRITIQUE OF PAST WORK AND A FORECAST FOR THE FUTURE

The Work Group believes that interlaboratory tests are essential in determining whether laboratories produce comparable data, and in some cases they have documented improvement on the part of some laboratories. However, the frequency of testing and the testing sequence within a field sampling analytical season have been very difficult to optimize. Furthermore, not

enough tests have been conducted to establish with confidence that laboratories are producing comparable data.

Due to the enormous cost in manpower and money to obtain environmental data, it is essential that such data are of sufficient quality to detect environmental trends, and to isolate the source of specific contaminants entering the Great Lakes waters. Therefore, the Work Group believes it imperative that interlaboratory studies be conducted before and during the analytical portions of monitoring and surveillance programs. Results from studies in advance of the field season are needed to provide Great Lakes laboratories the opportunity to take corrective measures to reduce out of control analyses and to eliminate identified bias. A second study during the analytical schedule will assess the effectiveness of remedial measures or reaffirm adequate performance, thereby providing data users and program managers an informative document on the probable comparability of large data sets.

The Work Group has encouraged and will continue to encourage participation by the jurisdictions in developing bulk material, Great Lakes' reference materials such as whole fish, sediments, and waters. These essential materials, when demonstrated as homogeneous and stable, will provide Great Lakes laboratories over time a continuous source of quality control material. To date, the Canada Centre for Inland Waters has developed an initial materials bank for some constituents.

The Work Group is just now beginning to address the capability of laboratories to identify and quantify toxic organic chemicals. These future tests along with the more traditional ones require many man years of work in planning, preparation, and execution. To date almost all work components of the Work Group's interlaboratory study programs have been provided by personnel of Environment Canada located at the Canada Centre for Inland Waters. Without the support of these dedicated people at CCIW, the Work Group would have little to report on interlaboratory studies.

The Work Group has found through experience that the use of a single facility with the necessary physical arrangements for sample preparation,

- 19 -

storage, and distribution, coupled with the necessary expertise to conduct interlaboratory studies, provide the best means to conduct its round robins. The Canada Centre for Inland Waters has consistently demonstrated this capability. Given current financial and manpower restraints it is not known whether the degree of past support provided at CCIW can be maintained for the future let alone increased as appears necessary. Members of the Work Group from other jurisdictions are seeking ways to either indirectly support the program through funds or equipment transfer to CCIW or directly provide some assistance in preparing and distributing interlaboratory test samples.

In order to meet its Terms of Reference, the Work Group will: prepare a compendium of analytical methods used by Great Lakes' laboratories; maintain the current compendium of field procedures; continue to test laboratory performance and promote improvement; provide recommendations on methods used; and serve others on specific requests pertaining to data generation, handling, and management.

The Data Quality Work Group will report its progress and findings to the Water Quality Programs Committee at least once a year and more often if deemed desirable or requested.

RECOMMENDATIONS

The Work Group's ability to fulfill its responsibilities depends not only on its own activities but on full participation, including funds and manpower, from the jurisdictions. Agency management within these jurisdictions has a responsibility to support Work Group quality assurance endeavors in order to achieve sufficient laboratory analysis, field sampling, and data assessment comparability.

To meet the various goals which may be subsumed under the heading of assuring data quality and comparability, the Work Group recommends:

o   that the Quality Assurance and Methods Section of the Division of Analytical Methods at the Canada Centre for Inland Waters be formally

recognized as the principal laboratory for preparation, storage, and distribution of interlaboratory study samples for the Work Group;

o   that the responsibility for providing interlaboratory study samples be equitably shared by the appropriate agencies in both countries through transfer of funds and/or provision of personnel;

o   that meetings of analytical chemists be held on at least an annual basis to foster understanding of the Agreement and the role laboratories will play in its fulfillment, and to identify and resolve common problems;

o   that agencies and the appropriate bodies within Agreement institutions identify data users so that those responsible for preparing reports may be properly involved in data quality concerns;

o   that either uniform techniques or techniques shown to be equivalent be used in the taking of samples;

o   that participation in Work Group interlaboratory studies be made mandatory for all laboratories providing environmental data for the assessment of contamination in the Great Lakes system so that biases may be detected and resolved;

o   that all major sewage treatment plants providing loading information be included in Work Group studies which are specifically designed to provide samples at typical effluent levels; and

o   that agencies accept the responsibility of having laboratories participating in International Joint Commission programs implement the recommended Guidelines for Intralaboratory Control.

April 28, 1982

File No. 2600-7

(Addressees: see list)

Dear :

Here is the evaluation of Interlaboratory Study #30, PCBs in ampuls and sediments. The evaluation includes general comment on the round robin, specific comments on your laboratory's performance, five tables and a set of graphs which display results in various ways, and three appendices - one describing sample preparation, another explaining the procedure for flagging aberrant results, and

General Comment

Target values for the ampuls - samples A, B, C and D - are the amounts of total PCBs in an equal mixture of Aroclors 1254 and 1260 in ng/ml which the ampuls were prepared to contain. The medians are in good agreement with these target values.

Target values for the sediment samples, excepting sample E which was not spiked, are the amounts of total PCBs in an equal mixture of Aroclors 1254 and 1260, added to the sediment, in ng/gram, plus 0.020 ug/g, the median for sample E. Prior analyses of the unspiked sediment also support the value of 0.020 for the sediment.

Comparison of the medians to the target values of the sediment samples gives an average recovery slightly above 70%, suggesting wide-spread difficulty with recoveries. Since there appear to be recovery problems, all flagging was done from target values rather than medians. Criteria for flagging are quite generous, 62% error is allowed at the 0.1 µg level with an increment for increased concentration of 30%. These criteria are derived from overall performance on the round robin and indicate the need for considerable improvement if production and elimination of bias between comparable results are achieved by the participants.

APPENDIX 1

STUDY #30 - PCBs IN SEDIMENT AND AMPULS

- 23 -

INTERNATIONAL JOINT COMMISSION

GREAT LAKES WATER QUALITY BOARD
DATA QUALITY WORK GROUP OF THE SURVEILLANCE SUBCOMMITTEE
100 OUELLETTE AVENUE, 8TH FLOOR
WINDSOR, ONTARIO, CANADA N9A 6T3

April 25, 1980

File No. 2400-7

(Addressees: see list)

Dear   :

Here is the evaluation of interlaboratory Study #30, PCBs in ampuls and sediments. The evaluation includes general comment on the round robin, specific comments on your laboratory's performance, five tables and a set of graphs which display results in various ways, and three appendices - one describing sample preparation, another explaining the procedure for flagging aberrant results, and the third summarizing methods.

General Comment

Target values for the ampuls - samples A, B, C and D - are the amounts of total PCBs in an equal mixture of Aroclors 1254 and 1260 in µg/ml which the ampuls were prepared to contain. The medians are in good agreement with these target values.

Target values for the sediment samples, excepting sample 6 which was not spiked, are the amounts of total PCBs, in an equal mixture of Aroclors 1254 and 1260, added to the sediment, in µg/gram, plus 0.020 µg/g, the median for sample 6. Prior analyses of the unspiked sediment also support the value of 0.020 for the sediment.

Comparison of the medians to the target values of the sediment samples gives an average recovery slightly above 70%, suggesting wide spread difficulty with recoveries. Since there appear to be recovery problems, all flagging was done from target values rather than medians. Criteria for flagging are quite generous, 50% error is allowed at the 0.1 µg level with an increment for increased concentration of 30%. These criteria are derived from overall performance on the round robin and indicate the need for considerable improvement in precision and elimination of bias before comparable results are achieved by the participants.

However, it should be noted that several of the participating laboratories' results exhibited a commendable degree of <u>within</u> laboratory precision.

Instead of reporting less thans, please use the W code to indicate the lowest level at which the analytical procedure followed allows confidence in identification.

Specific Comments

These comments are directly related to the results as displayed in Tables 2 and 3. Table 2 displays results corrected for known miscalculations and errors in reporting. For reported total PCBs, this table has flagged individual results and summarized ranked results for evaluation of bias. Table 3 gives the ratios of Aroclors 1260/1254 reported, differences between total PCBs reported and target values, and percent recoveries. This table contains results uncorrected for known miscalculations and reporting errors; footnotes discuss these errors. The graphs plot results on individual samples and give a visual display of performance.

Your laboratory number is (see separate list for laboratory results inserted).

Tables 1, 4, and 5

Table 1 gives results as originally submitted with reporting and calculation errors. Table 4 lists individual sample recoveries. Table 5 summaries percent recovery by laboratory for ampuls and sediments separately; the standard deviations are included as an indication of consistency.

Appendices 1, 2 and 3

Appendix 1 contains a thorough description of the wet sediment sample preparation. Appendix 2 gives the rationale for flagging results and a glossary of terms. Appendix 3 is a summary of methods.

We appreciate your participation and invite comment on the evaluation.

Sincerely yours,

K. I. Aspila
Chairman, Data Quality Work Group

KIA:REW:JLC:hk

Enclosures as stated

LABORATORY ADDRESSEES DO NOT CORRESPOND WITH NUMERICAL LISTING OF LAB RESULTS

Dr. David E. Armstrong
Water Chemistry Laboratory
University of Wisconsin-Madison
Madison, Wisconsin

Mr. G. A. V. Rees
Laboratory Services Branch
Ontario Ministry of the Environment
Toronto, Ontario

Dr. J. J. Delfino
State Laboratory of Hygiene
University of Wisconsin
Madison, Wisconsin

Mr. Curtis Ross
Director, Central Regional Laboratory
U.S. EPA, Region V
Chicago, Illinois

Mr. Carlton M. Duke
City of Chicago
Bureau of Water Operations
Central Water Filtration Plant
Chicago, Illinois

Dr. Ron Rossmann
Great Lakes Research Division
University of Michigan
Ann Arbor, Michigan

Mr. David E. Erdmann
National Water Quality Lab.-Atlanta
U.S. Geological Survey
Doraville, Georgia

Mr. R. F. Showalter
Asst. Director
Water & Sewage Laboratory Division
Indiana State Board of Health
Indianapolis, Indiana

Dr. Richard Frank
Pesticide Laboratory
Ont. Ministry of Agriculture and Food
c/o University of Guelph
Guelph, Ontario

Mr. T. R. Schwartz
Fish and Wildlife Service
U.S. Dept. of the Interior
Columbia National Fisheries Res. Lab.
Columbia, Missouri

Mr. Samuel J. Jackling
Senior Analytical Chemist
Hale Creek Field Station
NYS Dept. of Environmental Conservation
Gloversville, New York

Mr. A. Tupy
Minnesota Department of Health
Section of Health Risk Assessment
  and Analytical Services
Minneapolis, Minnesota

Dr. C. Stephen Kim
Director, Division of Labs and Research
New York State Dept. of Health
Empire State Plaza Labs
Albany, New York

Mr. Ralph Weigelt
Chief
Environmental Laboratories
Ohio Department of Health
Columbus, Ohio

Dr. M. D. Mullin
Chemist Manager
Large Lakes Research Station
U.S. Environmental Protection Agency
Grosse Ile, Michigan

Mr. W. A. Willford
U.S. Department of the Interior
Fish and Wildlife Service
Great Lakes Fishery Laboratory
Ann Arbor, Michigan

Mr. F. J. Philbert
Chief, Water Quality Branch, Ontario Region
Inland Waters Directorate
Canada Centre for Inland Waters
Burlington, Ontario

Mr. Tung Kai Wu
Laboratory Director
Michigan Dept. of Natural Resources
Environmental Laboratory
Lansing, Michigan

LABORATORY RESULTS

Your laboratory number is 2. Satisfactory results on ampuls. Results on sediment samples 3, 4, and 5 flagged high. Ranking indicates a high bias. The bias is virtually a constant on the sediment samples.

Your laboratory number is 3. Satisfactory results on all samples. There is some indication of a possible recovery problem on sediment samples 2, 5, and 7. The discerned ratios of 1254/1260 in samples 7 and 8 are a bit off.

Your laboratory number is 4. Satisfactory results on ampuls. Low results flagged on sediment samples 1, 2, 3, 5, 7, and 8; results on 1, 2, 7, and 8 are very low. Result on low level sample 6 is flagged high. Results on samples 2, 3, 4, and 6 are virtually identical although target values had a twenty fold range. The PCB ratios discerned for the ampuls are much closer to the target ratio of 1:1 than are the ratios discerned for the sediment samples. The ratio discerned for sample 5 is anomalous compared to the ratios discerned on the other sediment samples.

Your laboratory number is 5. Result flagged high on ampul A due to unexplained 1016 contamination. Results on sediment samples 1 and 2 flagged low, and there is also some indication of a recovery problem on samples 7 and 8. Results were reported as either 1254 or 1260; the 1:1 mixture of 1254 and 1260 in all samples was not discerned.

Your laboratory number is 9. Satisfactory results on ampuls. Results on sediment samples were originally misreported due to calculation errors. The corrected results on sediment samples 1 and 8 flagged low. There may also be recovery problems with samples 2, 5, and 7.

Your laboratory number is 11. Ampul results misreported as total µg/ampul which cannot be corrected since ampul volumes were not measured. Results on all sediment samples satisfactory, but without results on ampuls, the possibility of offsetting biases leading to satisfactory results must be considered.

Your laboratory number is 14. Result on ampul A originally reported for ampul B and vice versa. Corrected results flagged low on 3 of 4 ampuls and 4 of 8 sediment samples. Ranking indicates low bias. Chromatogram quality below average, perhaps due to presence of polar solvents. The possible presence of Aroclor 1254 was noted, but the correct 1:1 ratio of the two Aroclors was not discerned.

Your laboratory number is 13. Sediment and ampul contents were not identified as to what PCBs were present. Flagged low on Ampul C and samples 1, 2, 7, and 8. Flagged doubly high on samples 4 and 6. Sample 3 not reported because of error in sample handling. Results are quite erratic with extremely high values for samples 4 and 6.

Your laboratory number is 16. Result on ampul D flagged high. Results on sediment samples 1, 2, 5, and 8 flagged low. 1254 was misidentified as 1248.

Your laboratory number is 17. Results flagged low on ampuls A and C. Results flagged low on sediment samples 1, 3, 5, 7, and 8. Some results were misidentified as 1242.

Your laboratory number is 23. Satisfactory results on ampuls. Satisfactory results on sediment samples except for flagged result on sample 1 which is quite low.

Your laboratory number is 25. Satisfactory results on ampuls though a tendency to be low. Results on 6 of 8 sediment samples flagged low. Ranking indicates a clear low bias. Low bias may be due to chromic acid oxidation, since it has been shown that for Aroclor 1016 such oxidation results in serious losses. This is reported by Michael J. Szelewski, David R. Hill, Stewart J. Speigil, and Edwin C. Tifft, Jr. Loss of Polychlorinated Biphenyl Homologues During Chromium Trioxide Extraction of Fish Tissue, Anal. Chem., Vol. 51, 14, Dec. 1979, pp. 2405-2407. The correct 1:1 ratio for the two Aroclors was more nearly discerned for the ampuls than for the spiked sediment samples.

Your laboratory number is 27. Results on ampuls B and C flagged high. Result on sediment sample 1 flagged low and results on samples 4, 5, 6, 7, and 8 flagged high. Ranking indicates a clear high bias. Perhaps standards were low? Review of chromatograms suggests that detector response may not be linear over the range of attenuations used.

Your laboratory number is 40. Satisfactory results on ampuls though a tendency to be low. Results on sediment samples 2, 5, 7, and 8 flagged low. Perhaps standards were high.

Your laboratory number is 42. Results on ampuls satisfactory both qualitatively and quantitatively. Results on sediments qualitatively satisfactory, however, it is not possible to assess quantitative findings since an unknown amount of sample weight was discarded through pouring off separated water.

Your laboratory number is 44. Satisfactory results on all samples. Recoveries appear a bit low on sediment samples, except for the result on sample 5, which is a bit incongruous with the other sediment results.

Your laboratory number is 45. Satisfactory results on ampuls. Results on sediment samples 1, 2, and 8 flagged low. Result on sample 1 is very low.

Your laboratory number is 47. Your laboratory's results are not incorporated within the tables and graphs because this material was prepared several months ago. Results on ampuls are satisfactory. Flagged low on sediment sample #1 for total PCBs. Flagged high on sediment Sample #6 for total PCBs. PCB Aroclor 1242 should not have been found present in any of the sediment samples. Perhaps a laboratory contamination was present in samples 2 through 8. Qualitatively the ratio of 1260 to 1254 should have been 1, as reported for the ampuls. Whether storage of the samples has somewhat altered the PCB profiles remains a moot question.

TABLE 1. UNCORRECTED RESULTS AS ORIGINALLY SUBMITTED.   APRIL 25, 1980

PARAMETER: PCB IN SEDIMENT & AMPULS   DATA FILE: PCB30   YEAR: 1979   IJC GREAT LAKES REGIONAL OFFICE

SERIES IDENTIFICATION: STUDY 30   LABORATORIES YET TO REPORT: 13, 41, 42, 43, 46

LOWER LIMIT FOR USE OF BASIC ACCEPTABLE ERROR= 0. BASIC ACCEPTABLE ERROR= 1. CONCENTRATION ERROR INCREMENT= 1
LABORATORY RESULTS OMITTED ARE NONE.

SAMPLE LAB NO.	A REPORTED VALUE	A RANK	B REPORTED VALUE	B RANK	C REPORTED VALUE	C RANK	D REPORTED VALUE	D RANK	1 REPORTED VALUE	1 RANK
2	2.02	9.50	0.52	8.00	0.27	7.50	0.99	6.00	2.29	15.00
3	2.25	12.00	0.54	9.50	0.31	12.00	1.26	12.00	2.01	14.00
4	1.9970	8.00	0.4467	3.00	0.3137	13.00	0.8505	4.00	0.60	3.00
5	3.60	14.00	0.51	7.00	0.22	4.00	1.12	11.00	1.33	8.00
9	1.91	7.00	0.49	5.00	0.24	5.00	0.90	5.00	0.682	4.00
11	----	0.00	----	0.00	----	0.00	----	0.00	1.76	13.00
14	0.29	1.00	1.20	14.00	0.27	7.50	0.55	1.00	1.41	10.00
16	2.02	9.50	0.54	9.50	0.28	10.00	1.42	14.00	0.78	5.00
17	1.13	2.00	0.604	12.00	0.148	1.00	1.00	8.50	0.59	2.00
23	1.9	6.00	0.56	11.00	0.28	10.00	1.0	8.50	1.23	6.00
25	1.58	3.00	0.40	2.00	0.203	3.00	0.77	2.00	1.39	9.00
27	2.4	13.00	0.79	13.00	0.63	14.00	1.37	13.00	1.39	9.00
40	1.6190	4.00	0.3750	1.00	0.1978	2.00	0.7998	3.00	1.4288	11.00
44	2.1	11.00	0.50	6.00	0.86	6.00	1.0	8.50	1.6	12.00
45	1.785	5.00	0.486	4.00	0.280	10.00	1.000	8.50	0.379	1.00
MEDIAN IS	1.954		0.515		0.270		1.000		1.330	

SAMPLE LAB NO.	2 REPORTED VALUE	2 RANK	3 REPORTED VALUE	3 RANK	4 REPORTED VALUE	4 RANK	5 REPORTED VALUE	5 RANK	6 REPORTED VALUE	6 RANK
2	0.53	14.00	0.55	15.00	0.56	15.00	0.86	14.00	0.02	4.50
3	0.29	8.00	0.18	11.00	0.11	10.00	0.37	8.00	<0.05	0.00
4	0.09	1.00	0.12	4.00	0.12	12.00	0.29	4.00	0.10	10.50
5	0.20	3.50	0.15	8.00	0.11	10.00	0.52	10.50	0.02	4.50
9	0.681	15.00	0.112	3.00	0.063	2.00	0.199	2.00	----	0.00
11	0.31	9.50	0.246	13.00	0.110	10.00	0.55	12.00	0.027	8.00
14	0.24	6.00	0.13	5.00	0.06	1.00	0.43	9.00	0.02	4.50
16	0.20	3.50	0.15	8.00	0.10	8.00	0.34	5.00	0.02	4.50
17	0.41	13.00	0.101	1.00	0.074	4.00	0.156	1.00	0.0473	9.00
23	0.34	11.00	0.25	14.00	0.15	13.00	0.52	10.50	----	0.00
25	0.179	2.00	0.109	2.00	0.068	3.00	0.220	3.00	0.020	7.00
27	0.39	12.00	0.19	12.00	0.28	14.00	0.87	15.00	0.10	10.50
40	0.2520	7.00	0.1417	6.00	0.0791	5.00	0.3409	6.00	0.0153	2.00
44	0.31	9.50	0.15	8.00	0.093	6.00	0.68	13.00	<0.01	1.00
45	0.220	5.00	0.164	10.00	0.097	7.00	0.360	7.00	----	0.00
MEDIAN IS	0.290		0.150		0.100		0.370		0.020	

SAMPLE	7		8	
	REPORTED		REPORTED	
LAB NO.	VALUE	RANK	VALUE	RANK
2	1.27	13.00	1.57	14.00
3	0.80	10.00	1.23	12.00
4	0.35	2.00	0.39	2.00
5	0.74	7.00	0.96	10.00
9	0.345	1.00	0.378	1.00
11	0.95	11.00	0.89	9.00
14	0.48	4.00	0.73	5.00
16	----	0.00	0.80	7.00
17	0.49	5.00	0.42	3.00
23	1.00	12.00	1.25	13.00
25	0.47	3.00	0.57	4.00
27	1.44	14.00	1.94	15.00
40	0.5597	6.00	0.7425	6.00
44	0.77	9.00	0.97	11.00
45	0.748	8.00	0.844	8.00
	MEDIAN IS	0.744	MEDIAN IS	0.844

LAB NO.	TOTAL RANK	AVERAGE RANK	NO. OF SAMPLES RANKED	SUMMARY OF FLAGGING
2	135.50	11.292	12	BIASED HIGH
3	118.50	10.773	11	
4	66.50	5.542	12	
5	97.50	8.125	12	
9	50.00	4.545	11	
11	85.50	10.688	8	
14	68.00	5.667	12	
16	86.00	7.818	11	
17	64.50	5.375	12	
23	111.00	10.091	11	
25	40.00	3.333	12	BIASED LOW
27	154.50	12.875	12	BIASED HIGH
40	59.00	4.917	12	
44	101.00	8.417	12	
45	73.50	6.682	11	

DATA ARE STORED ON DATA FILE PCB30, REVISED 800313.

,OVERALL AVERAGE RANK IS   7.667

SAMPLE	7		8	
	REPORTED		REPORTED	
LAB NO.	VALUE	RANK	VALUE	RANK
2	1.27	13.00	1.57	14.00
3	0.80	10.00	1.23	12.00
4	0.35bb	1.00	0.39bb	1.00
5	0.74	6.00	0.96	10.00
9	0.754	8.00	0.834b	7.00
11	0.95	11.00	0.89	9.00
14	0.48bb	3.00	0.73b	4.00
16	----	0.00	0.80b	6.00
17	0.49bb	4.00	0.42bb	2.00
23	1.00	12.00	1.25	13.00
25	0.47bb	2.00	0.57bb	3.00
27	1.44#	14.00	1.94##	15.00
40	0.5597b	5.00	0.7425b	5.00
44	0.77	9.00	0.97	11.00
45	0.748	7.00	0.844b	8.00

→ TARGET IS 1.030   TARGET IS 1.240 ←
MEDIAN IS 0.751   MEDIAN IS 0.844

13   *0.671*   *0.867*

LAB NO.	TOTAL RANK	AVERAGE RANK	NO. OF SAMPLES RANKED	SUMMARY OF FLAGGING	DATA ARE STORED ON DATA FILE PCB30, REVISED 800313.
2	137.50	11.458	12	######	BIASED HIGH
3	116.50	10.591	11		
4	63.50	5.292	12	bbbbbb##bbbb	
5	95.50	7.958	12	##bb	
9	85.00	7.727	11	bb	
11	85.50	10.688	8		
14	53.00	4.417	12	bbbbbbbb	BIASED LOW
16	82.00	7.455	11	#bbbb	
17	61.50	5.125	12	bbbbbbbbbb	
23	113.00	10.273	11	bb	
25	36.00	3.000	12	bbbbbbbbbb	BIASED LOW
27	155.50	12.958	12	####b#########	BIASED HIGH
40	55.00	4.583	12	bbbb	
44	101.00	8.417	12		
45	70.50	6.409	11	bbbb	,OVERALL AVERAGE RANK IS 7.667

April 25, 1980

TABLE 3.

QUANTIFICATION OF SPECIFIC AROCLORS AND PERCENT RECOVERY OF TOTAL PCBs
STUDY #30 - PCBs IN SEDIMENT AND AMPULS

SAMPLE		LAB. #2						LAB. #3						LAB. #4					
Target*	No.	1254	1260	1260/1254	Total¹	Δ	%	1254	1260	1260/1254	Total²	Δ	%	1254	1260	1260/1254	Total²	Δ	%
0.260	C	-	-	-	0.27	.01	104	0.12	0.19	1.58	0.31	+.05	119	0.19	0.12	0.63	0.31	+.05	119
0.520	B	-	-	-	0.52	-	100	0.23	0.31	1.35	0.54	+.02	104	0.223	0.223	1.00	0.45	-.07	87
1.04	D	-	-	-	0.99	-.05	95	0.5	0.76	1.52	1.26	+.22	121	0.41	0.44	1.07	0.85	-.19	82
2.08	A	-	-	-	2.02	-.06	97	0.98	1.26	1.29	2.25	+.17	108	1.16	0.83	0.71	2.0	-.08	96
0.02	6	-	-	-	0.02	-	100	-	-	-	<.05	-	-	0.08	0.02	0.25	0.10	+.08	500
0.118	4	-	-	-	0.56	+.44	475	0.06	0.05	0.83	0.11	-.01	93	0.08	0.04	0.50	0.12	.00	102
0.209	3	-	-	-	0.55	+.34	263	0.08	0.10	1.25	0.18	-.03	86	0.09	0.03	0.33	0.12	-.09	57
0.418	2	-	-	-	0.53	+.12	127	0.12	0.17	0.71	0.29	-.13	69	0.07	0.02	0.29	0.09	-.33	22
0.528	5	-	-	-	0.86	+.33	163	0.20	0.17	1.18	0.37	-.16	70	0.09	0.20	2.2	0.29	-.24	55
1.030	7	-	-	-	1.27	+.24	123	0.28	0.52	1.86	0.80	-.23	78	0.26	0.09	0.34	0.35	-.68	34
1.240	8	-	-	-	1.57	+.33	127	0.43	0.80	1.86	1.23	-.01	99	0.28	0.11	0.39	0.39	-.85	32
2.022	1	-	-	-	2.29	+.27	113	0.91	1.1	1.21	2.01	-.01	99	0.44	0.16	0.36	0.60	-1.42	30

*Total PCB, a 1:1 mixture of 1254 and 1260 in µg/ml in ampuls and µg/g in sediments.
¹All peaks quantified using mixture of 1:1:1 1260/54/48.
²Used 1254/60 matches for quantification.

Δ = TOTAL PCBs - TARGET.
% = % RECOVERY, CALCULATED FROM TARGET VALUE.
RESULTS ARE AS ORIGINALLY SUBMITTED; FOOTNOTES INCLUDE EXPLANATIONS OF MISCALCULATIONS AND REPORTING ERRORS.

- 34 -

TABLE 3. Cont'd.

QUANTIFICATION OF SPECIFIC AROCLORS AND PERCENT RECOVERY OF TOTAL PCBs
STUDY #30 - PCBs IN SEDIMENT AND AMPULS

- 35 -

SAMPLE Target No.	LAB. #5						LAB. #9						LAB. #11					
	1254	1260	1260/1254	Total	Δ	%	1254	1260	1260[4]/1254	Total[5]	Δ	%	1254	1260	1260/1254	Total[6]	Δ	%
0.260 C	0.22	-		0.22	-.04	85				0.24	-.02	92			1.22			
0.520 B	0.51	-		0.51	-.01	98				0.49	-.03	94			1.03			
1.04 D	1.12	-		1.12	+.08	108				0.90	-.14	87			2.25			
2.08 A	-[3]	2.00		2.00	-.08	96				1.91	-.17	92			1.9			
0.02 6	0.02	-		0.02	0	100				<.02	≃0		0.027	-	-	0.027	≃.01	135
0.118 4	0.11	-		0.11	-.01	93				0.063	-.05	53	0.053	0.057	1.08	0.11	-.01	93
0.209 3	0.15	-		0.15	-.06	72				0.112	-.097	46	0.086	0.16	1.86	0.25	+.04	120
0.418 2	-	0.20		0.20	-.22	48				0.681	+.263	163	0.13	0.18	1.38	0.31	-.11	74
0.528 5	0.52	-		0.52	-.01	98				0.199	-.33	38	0.25	0.30	1.2	0.55	+.02	104
1.030 7	0.74	-		0.74	-.29	72				0.345	-.685	34	0.46	0.49	1.07	0.95	-.08	92
1.240 8	0.96	-		0.96	-.28	77				0.378	-.862	30	0.39	0.50	1.28	0.89	-.35	72
2.022 1	1.33	-		1.33	-.69	66				0.682	-1.34	34	0.86	0.90	1.04	1.76	-.26	87

[3]Contamination of 1.6 µg/ml of PCB 1016 reported.
[4]Identified as 1:1 mixture of 1260/1254 but not quantified individually.
[5]Results originally reported as above were miscalculated and later corrected by the laboratory as follows: 4(.115), 3(.207), 2(.309), 5(.424), 7(.754), 8(.834) and 1(1.26) yield 97, 99, 74, 80, 73, 68, and 63% recovered respectively.
[6]Data for ampuls, other than 1260/1254 ratio, are inappropriate because reported as µg/ampul and cannot be corrected since volumes are unknown. Sediment and ampul findings qualified and quantified using peack ratios for 1254/1260.

TABLE 3. Cont'd.   April 25, 1980

QUANTIFICATION OF SPECIFIC AROCLORS AND PERCENT RECOVERY OF TOTAL PCBs
STUDY #30 - PCBs IN SEDIMENT AND AMPULS

SAMPLE		LAB. #14					LAB. #16						LAB. #17					
Target	No.	1254	1260[7]	1260[8]/1254	Total	Δ	%	1248[9]	1260	1260/1254	Total[10]	Δ	%	1254	1260	1260/1254	Total	Δ	%
0.260	C		0.27		0.27	+.01	104	0.05	0.20		0.28	+.02	108	0.14	0.008	0.06	0.148	-.11	57
0.520	B		1.20		1.20	+.68	231	0.14	0.38		0.54	+.02	104	0.54	0.064	0.12	0.604	+.08	116
1.04	D		0.55		0.55	-.49	53	0.20	1.16		1.42	+.38	137	0.74	0.26	0.35	1.00	-.04	96
2.08	A		0.29		0.29	-1.79	14	0.45	1.62		2.02	-.06	97	0.86	0.27	0.31	1.13	-.95	54
0.02	6			0.02	0.02	0	100	0.01	0.01		0.02	0	100	.015[11]	0.032[11]		0.047	+.27	235
0.118	4			0.06	0.06	-.058	51	0.03	0.06		0.10	-.02	85	0.033[12]	0.041		0.074	-.04	63
0.209	3			0.13	0.13	-.079	62	0.04	0.11		0.15	-.06	72	0.046[12]	0.055		0.101	-.11	48
0.418	2			0.24	0.24	-.178	57	0.04	0.15		0.20	-.22	48	0.18[12]	0.23		0.41	-.01	98
0.528	5			0.43	0.43	-.098	81	0.07	0.25		0.34	-.19	64	0.036[12]	0.12		0.156	-.37	30
1.030	7			0.48	0.48	-.55	47	-	-		-	-	-	0.31	0.18	0.58	0.49	-.54	48
1.240	8			0.73	0.73	-.51	59	0.13	0.62		0.80	-.44	65	0.26	0.18	0.69	0.42	-.82	34
2.022	1			1.41	1.41	-.612	70	0.32	0.89		1.24	-.78	61	0.29[12]	0.49		0.78	-1.24	39

[7]Reporting ampuls A and B were mixed up, A should have been reported as B and B as A. Above results are as originally reported.
[8]All results quantified to 1260, however, analysts recognized the likely presence of 1254.
[9]Used 1248 rather than 1254 to qualify and quantify. Later by phone acknowledge that had suspected 1254 rather than 1248. Expressed an interest in re-doing.
[10]Reported a combined calculation rather than a simple sum of individual PCB results.
[11]Identified as 1254 and 1242; the 1254 figure is in the 1260 column.
[12]Results are based on 1242 identification and quantified from 1242.

- 36 -

TABLE 3. Cont'd.   April 25, 1980

QUANTIFICATION OF SPECIFIC AROCLORS AND PERCENT RECOVERY OF TOTAL PCBs
STUDY #30 - PCBs IN SEDIMENT AND AMPULS

SAMPLE		LAB. #23						LAB. #25						LAB. #27					
Target No.		1254	1260	1260[13]/1254	Total	Δ	%	1254	1260	1260/1254	Total	Δ	%	1254	1260	1260/1254	Total	Δ	%
0.260	C	0.14	0.14	1.0	0.28	+.02	108	0.11	0.093	0.85	0.203	-.057	78	0.31	0.32	1.03	0.63	+.37	242
0.520	B	0.29	0.27	0.93	0.56	+.04	108	0.21	0.19	0.90	0.40	-.12	77	0.42	0.37	0.88	0.79	+.27	152
1.04	D	0.5	0.5	1.0	1.0	-.04	96	0.41	0.36	0.88	0.77	-.27	74	0.72	0.65	0.90	1.37	+.33	132
2.08	A	0.9	1.0	1.11	1.9	-.18	91	0.85	0.73	0.86	1.58	-.50	76	1.3	1.1	0.53	2.4	+.32	115
0.02	6			-	-	-	-	0.026	<.01	-	0.026	+.006	130	0.10	<.10	-	≈.10		
0.118	4				0.15	+.03	127	0.050	0.018	0.36	0.068	-.05	58	0.15	0.13	0.87	0.28	+.162	237
0.209	3				0.25	+.04	120	0.075	0.034	0.45	0.109	-.10	52	0.19	<.10	-	≈0.19		
0.418	2				0.34	-.08	81	0.11	0.069	0.63	0.179	-.24	43	0.24	0.15	0.63	0.39	-.028	93
0.528	5				0.52	-.01	98	0.14	0.08	0.57	0.220	-.308	42	0.40	0.47	0.89	0.87	+.342	165
1.030	7				1.00	-.03	97	0.28	0.19	0.68	0.47	-.56	46	0.81	0.63	0.78	1.44	+.41	140
1.240	8	0.60	0.65	1.08	1.25	+.01	101	0.33	0.24	0.73	0.57	-.67	46	0.84	1.1	0.89	1.94	+.70	156
2.022	1				0.59	-1.43	29	0.74	0.49	0.66	1.23	-.792	61	0.68	0.71	1.04	1.39	-.632	69

[13] Sediment samples 1 through 5 and No. 7 qualified as 1:1 mixture of 1254/1260 and calculated as such.

TABLE 3. Cont'd.   April 25, 1980

QUANTIFICATION OF SPECIFIC AROCLORS AND PERCENT RECOVERY OF TOTAL PCBs
STUDY #30 - PCBs IN SEDIMENT AND AMPULS

SAMPLE		LAB. #40						LAB. #44						LAB. #45					
Target	No.	1254	1260	$\frac{1260[14]}{1254}$	Total	Δ	%	1254	1260	$\frac{1260[15]}{1254}$	Total	Δ	%	1254	1260	$\frac{1260[15]}{1254}$	Total	Δ	%
0.260	C				0.198	-.062	76				0.26	0	100				0.280	+.02	108
0.520	B				0.375	-.145	72				0.50	-.02	96				0.486	-.034	93
1.04	D				0.800	-.24	77				1.0	-.04	96				1.00	-.04	96
2.08	A				1.62	-.46	78				2.1	+.02	101				1.785	-.295	86
0.02	6				0.0153	-.005	76				<0.01						N/D		
0.118	4				0.079	-.039	67				0.093	-.025	79				0.097	-.021	82
0.209	3				0.142	-.067	68				0.15	-.059	72				0.164	-.045	78
0.418	2				0.252	-.166	60				0.31	-.108	74				0.220	-.198	53
0.528	5				0.341	-.187	65				0.68	+.152	129				0.360	-.168	68
1.030	7				0.560	-.47	54				0.77	-.26	75				0.748	-.282	73
1.240	8				0.743	-.51	60				0.97	-.27	78				0.844	-.396	68
2.022	1				1.43	-.592	71				1.6	-.422	79				0.379	-1.64	19

[14] Laboratory qualitatively verified 1:1 PCB mixture of 1254/1260 and quantified using a 1:1 mixed standard.
[15] Laboratory identified the PCBs as a 1:1 mixture of 1254/1260 and quantified as total with 1:1 standards.

April 25, 1980

TABLE 4.

INDIVIDUAL SAMPLE RECOVERIES BY LABORATORY - STUDY #30

Lab. Code	% Recovery on Ampuls				% Recovery on Wet Sediments						
	C	B	D	A	4	3	2	5	7	8	1
2	103.85	100.00	95.19	97.12	474.58	263.16	126.79	162.88	123.30	126.61	113.25
3	119.23	103.85	121.15	108.17	93.22	86.12	69.38	70.08	77.67	99.19	99.41
4	120.65	85.90	81.78	96.01	101.69	57.42	21.53	54.92	33.98	31.45	29.67
5	84.62	98.08	107.69	173.08	93.22	71.77	47.85	98.48	71.84	77.42	65.78
9	92.31	94.23	86.54	91.83	97.46	99.04	73.92	80.30	73.20	67.26	62.31
11	-	-	-	-	93.22	117.70	74.16	104.17	92.23	71.77	87.04
14	103.85	55.77	52.88	57.69	50.85	62.20	57.42	81.44	46.60	58.87	69.73
16	107.69	103.85	136.54	97.12	84.75	71.77	47.85	64.39	-	64.52	61.33
17	56.92	116.15	96.15	54.33	62.71	48.33	98.09	29.55	47.57	33.87	38.58
23	107.69	107.69	96.15	91.35	127.12	119.62	81.34	98.48	97.09	100.81	29.18
25	78.08	76.92	74.04	75.96	57.63	52.15	42.82	41.67	45.63	45.97	60.83
27	242.31	151.92	131.73	115.38	237.29	90.91	93.30	183.71	139.81	156.45	68.74
40	76.08	72.12	76.90	77.84	67.03	67.80	60.29	64.56	54.34	59.88	70.66
44	100.00	96.15	96.15	100.96	78.81	71.77	74.16	128.79	74.76	78.23	79.13
45	107.69	93.46	96.15	85.82	82.20	78.47	52.63	68.18	72.62	68.06	18.74
*Target Value	μg/ml				μg/gm						
	0.26	0.52	1.04	2.08	0.118	0.209	0.418	0.528	1.03	1.24	2.022

*% Recovery based on (spike + x) x = 0.02.

RESULTS IN THIS TABLE ARE CORRECTED FOR KNOWN MISCALCULATIONS AND REPORTING ERRORS.

TABLE 5.   April 25, 1980

PERCENT RECOVERY BY LABORATORY* FOR PCB RESULTS
STUDY #30

LAB.	AMPULS		WET SEDIMENTS		DIFFERENCE
	% Recovery (4 Ampuls) R_1	Std. Dev. on % Recovery $n=4$	% Recovery (7 Sediments) R_2	Std. Dev. on % Recovery $n=7$	$R_1 - R_2$
2	99.04	3.77	198.65	132.22	-99.61
3	113.10	8.41	85.01	12.90	28.09
4	96.07	17.44	47.24	27.45	48.83
5	115.87	39.30	75.19	16.99	40.68
9	91.23	3.29	79.07	14.26	12.16
11	-	-	91.47	16.13	-
13	-	-	-	-	-
14	67.55	24.28	61.02	11.71	6.53
16	111.30	17.39	65.77 (n=6)	12.17	45.53
17	80.89	30.31	51.24	23.38	29.65
23	100.72	8.29	93.38	32.12	7.34
25	76.25	1.71	49.53	7.47	26.72
27	160.34	56.65	135.90	57.59	24.44
40	75.73	2.52	63.51	5.64	12.22
41	-	-	-	-	-
42	-	-	-	-	-
43	-	-	-	-	-
44	98.32	2.53	83.66	20.09	14.66
45	95.78	9.07	62.99	21.68	32.79
46	-	-	-	-	-
MEAN	98.73 (n=14)		82.91 (n=15)		15.82
MEDIAN	97.20		75.19		

*Results corrected for known miscalculations and reporting errors.

SAMPLE B

| 27 | | 3 | 44 | 16 | 2 | 4 | 9 | 23 | 24 | 45 | 40 | 25 | 14 | 17 | |

1.11 ↑##
1.44 ↑#
TARGET 2.08
2.72 ↓#
3.05 ↓##

5

SAMPLE A

- 41 -

GRAPH OF RANKED RESULTS

PCB IN SEDIMENT & AMPULS   SERIES I.D. = STUDY 30

SAMPLE B

LABORATORY NUMBERS MARK GRAPHED VALUES

GRAPH OF RANKED RESULTS

PCB IN SEDIMENT & AMPULS   SERIES I.D.= STUDY 30

SAMPLE I

```
##↑
2.96  -       -       -       -       -       -       -       -       -       -

 #↑
2.65  -       -       -       -       -       -       -       -       -       -

                                                                           2
- 45 -

TARGET 2.02 ┬───────────────────────────────────────────────────────────────
                                                                   •      •
                                                                   3
                                                                       11
                                                            44
1.40  -       -       -       -       •       -       -      •40     -       -
 B↓                                          27    14
                                   25    16    9    5
1.08  -       -       -       -            -       -       -       -       -
 BB↓

                            17


                     23    4

                     ↓
                    45
```

LABORATORY NUMBERS MARK GRAPHED VALUES

GRAPH OF RANKED RESULTS

PCB IN SEDIMENT & AMPULS   SERIES I.D.= STUDY 30

SAMPLE 3

LABORATORY NUMBERS MARK GRAPHED VALUES

GRAPH OF RANKED RESULTS

PCB IN SEDIMENT & AMPULS   SERIES I.D.= STUDY 30

SAMPLE 6

```
##↑
0.10  -              -              -              -         -      4    27
                                                                  -
##↑
0.07     -           -           -              -           -           -

- 50 -

                                                              17
                                                    25    11
                              16   2   14   5
TARGET  0.02 ──────────────●───────●────●────────●────●────●────●─────●──●──
                    44    40
-0.03
B↓      -           -              -           -           -           -
-0.06
BB↓     -           -              -           -           -           -
```

LABORATORY NUMBERS MARK GRAPHED VALUES

GRAPH OF RANKED RESULTS

PCB IN SEDIMENT & AMPULS   SERIES I.D. = STUDY 30

SAMPLE 7

```
##↑
1.52  -           -        -           -           -           -           -           -
                                                                                    27

#↑
1.36  -           -        -           -           -           -           -    -
                                                                         2

                                                                                -51-

TARGET 1.03 _____
                                                                              11
                                                                       23
                                                            3
                                               44
0.70  -           -        -        9   .      -           -           -
B↓                              5    .45
0.54  -           -        -           -           -           -           -    -
BB↓                          17      40
          25      14
       4
```

LABORATORY NUMBERS MARK GRAPHED VALUES

GRAPH OF RANKED RESULTS

PCB IN SEDIMENT & AMPULS   SERIES I.D.= STUDY 30

SAMPLE B

LABORATORY NUMBERS MARK GRAPHED VALUES

STUDY #30 - PCBs IN AMPULS AND SEDIMENTS
APPENDIX 1
SAMPLE PREPARATION

The PCBs in wet sediments study was designed and prepared by B. Lee and A. S. Y. Chau of the Quality Assurance and Methods Section, Analytical Methods Division, National Water Research Institute located at the Canada Centre for Inland Waters, Burlington, Ontario, Canada.

The bulk reference sediment used for this study was collected in 1978 with a double Shipex sediment sampler from Station 24 in Lake Ontario, about 10 miles north of Niagara River.

The lake's bottom was sampled to a depth of 10 cm, thereby including material thought to be about 10,000 years old, which minimizes possible contamination from high concentrations of PCBs and chlorinated hydrocarbon pesticides that might be at the sediment water interface.

Sample 6 of this study represents a portion of the originally sampled sediment. The other samples were spiked aliquots of the original sediment. The procedure to spike sediments for interlaboratory tests or other laboratory quality control purposes is described by Alfred S. Y. Chau, John Carron, Hing-Bin Lee, Analytical Reference Materials. II. Preparation and Sample Integrity of Homogeneous Fortified Wet Sediment for Polychlorinated Biphenyl Quality Control Studies, J. Assoc. Off. Anal. Chem. (Vol. 62, 6, 1979), pp. 1312-1314.

The authors report in the summary portion of the paper:

"A simple method for the preparation of a large quantity of homogeneous wet sediment spiked with polychlorinated biphenyls (PCBs) is presented. By using a large blender and adding water to form a thick slurry, more than 2 kg spiked wet sediment was homogenized and 100-200 subsamples of 10-15 g, each suitable for checking precision and accuracy of a method or a laboratory, could be obtained.

Eighteen Lake Ontario sediment subsamples were analyzed to check homogeneity. The mean recovery was 97.9% for a 1:1 mixture of Aroclors 1254 and 1260 fortified at 1 ppm; the coefficient of variation was 5.4%. For the Georgian Bay sediment, recovery for 5 replicates was 97.8% with a coefficient of variation of 3.2% for the same Aroclor mixture fortified at 0.629 ppm. Subsamples store at -20°C for up to 3 months showed no losses of PCBs for these 2 sediments. Stability data also provided additional support for the homogeneity of the subsamples for these 2 sediments."

The solutions in the ampuls and the solutions used to spike the sediments were from the same laboratory stock solutions to provide the following concentrations of 1:1 Aroclors 1254 and 1260 as total PCB:

SEDIMENTS

Sample	Spike Level, µg/g	Sample	Spike level, µg/g
1	x + 2.002	5	x + 0.508
2	x + 0.398	6	x (not spiked)
3	x + 0.189	7	x + 1.01
4	x + 0.098	8	x + 1.22

AMPULS

Sample	Concentration µg/ml	Sample	Concentration µg/ml
A	2.08	C	0.26
B	0.52	D	1.04

STUDY #30 - PCBs IN AMPULS AND SEDIMENTS
APPENDIX II
CRITERIA USED FOR DETERMINING BIAS AND FLAGGING RESULTS
AND A GLOSSARY OF TERMS USED

Bias and Flagging:

A set of results is said to be biased when the <u>set</u> exhibits a tendency to be either higher or lower than some standard - the standard which has been used in the analysis of our studies thus far has been the performance of all other participating laboratories. The ranking procedure employed in testing for bias is described in W. J. Youden's paper, "Ranking Laboratories by Round-Robin Tests" from <u>Precision</u> <u>Measurement</u> <u>and</u> <u>Calibration</u>, H. H. Ku, Editor, NBS Special Publication 300-Volume 1, U.S. Government Printing Office, Washington, D.C., 1969. In this paper, Youden establishes the rationale for evaluating laboratories' performance by ranking results. In our use of the procedure there is about 1 change in 20 of deeming a set of results biased when in fact it is not, that is, $\alpha = 0.05$.

When there are unreported results or results which are not ranked, the procedure is modified slightly for determining high bias. An adjustment is made for the decreased possible total score. In addition, when a particular laboratory has unreported or unranked results, an adjustment is made for the fewer number of samples ranked.

Lower Limit for Use of Basic Acceptable Error; Basic Acceptable Error; and Concentration Error Increment:

These terms define the acceptable difference from the target value that is allowed without a result being flagged either low or high. For a sample whose target value is at or below the lower limit for use of basic acceptable error the basic acceptable error is used to determine the range of acceptable deviations.

Since for almost all substances it appears that the variability of results increases with increases in concentration an allowance is made for the increased variability for those sampes whose target values are above the lower limit for use of basic acceptable error. The allowance is added to the basic acceptable error, and it is calculated by multiplying the concentration error increment by the difference between the target value and the lower limit for use of basic acceptable error.

For example: The lower limit for use of basic acceptable error for Total PCBs in sediment (this study) is 0.1 µg/g, the basic acceptable error is 0.05 µg/g, and the concentration error increment is 0.3 µg/g. The target value for sample #5 was determined to be 0.528 µg/g. The difference between the target value and the lower limit for use of basic acceptable error is 0.528-0.10 = 0.428 µg/g. Multiplying this difference (0.428 µg/g) by the concentration error increment (0.300) equals 0.1284 µg/g. This allowance is added to the basic acceptable error of 0.05 µg/g to determine the acceptable difference of 0.1784 µg/g for the sample. Therefore, any reported result within the range 0.528 ± .1784 or 0.3496 to 0.7064 µg/g would be considered acceptable and not flagged.

A result is flagged high # when its value is greater than the target value plus the acceptable difference but not greater than the target plus 1.5 times; a result greater than 1.5 times the acceptable difference is flagged with ##. Similarly, a result less than the target minus the acceptable difference but not less than the target minus 1.5 times the acceptable difference is flagged b; a lower result is flagged bb.

In general, the values chosen for the basic acceptable error and the concentration error increment are derived primarily from the results received for the range of samples analyzed, augmented by our judgment of reasonable performance. The underlying concept is that if several laboratories are found to perform adequately with the values chosen, then all laboratories participating should be capable of that level of performance. In a sense the values represent the present state of the art for analysis of this kind of round-robin samples. As experience is gained, the values will be modified

performance and that over time, the acceptable ranges for results will become
somewhat tighter.

Out of Control:

An analytical system is said to be out of control when it has demonstrated
the ability to perform adequately and produces an extreme result or results.
For an example, consider the set of results by laboratory #3 on total
phosphorus in Study #24.

Sample No.	Reported Value	Median	Difference
1	9	9.5	-.5
2	5	4.5	.5
3	2T	3	-1
4	8	8	0
5	2T	2.5	-.5
6	9	8	1
7	28	28	0
8	18	17	1
9	23	23.7	-.7
10	16	15	1
11	35	35.8	-.8
12	75	78.7	-3.8
13	58	59	-1
14	110	90	20

Given the excellent results obtained on samples 1 through 13, the result
on sample 14 indicates that the analytical system was out of control.

W: A "W" code is used with a reported result when no measurement was possible
   due to no response of the instrument to the sample. The "W" is preceded
   by the smallest determinative division that can be used in the units used
   in reporting.

T: The "T" code is used for results with values between the Criterion of
   Detection and the "W" value. The Criterion of Detection is commonly
   thought of by many as the limit of detection.

Satisfactory: Quite acceptable, "good results."

Erratic: A set of results for a given characteristic is deemed erratic when
   both high and low flags are assigned.

LABORATORY #2

Extraction:

Ultrasonic probe 10+1 solvent - sample with 1+1 hexane-acetone.

Activated copper used in extraction to remove sulphur.

Partition organic phase with water and back extract water phase with benzene.

Dry combined organic phases (by passing through sodium sulphate?), add isooctane and evaporate to ca. 5 ml.

Clean-up:

Fractionate the extract on standardized 28 g Florisil column.

Use Hg in final extracts to remove residual sulphur.

Concentrate on rotary evaporator followed by vortex evaporation.

Qualification-Quantification:

Webb-McCall - all samples quantified against 1+1+1 mixture of Aroclors 1242, 1254, and 1260.

Standards - 100 mg - pure - from U.S. EPA Research Triangle Park.

Column - glass 1.8 m x 2.5 mm, OV-101 - 3%, Chrom. W - 80/100.

Detector - Ni-63 ECD.

LABORATORY #3

Extraction:

Entire contents Soxhlet reflux 2 hr. using 1+1 acetone-hexane. 2nd reflux 22 hr. using 1+2 acetone-hexane.

Water partitioned from combined organic phases and extract dried using sodium sulphate.

Extract concentrated using Kuderna-Danish evaporator with six ball Snyder column.

Clean-up:

Extract shaken with Hg to remove residual sulphur and passed through Florisil (solvent used?).

Qualification-Quantification:

Unique peak selection to match individual standards.

Standards - 100% from Monsanto - 1221, 1254, 1260, and 1242.

Column - glass - 6 ft. x 2 mm ID, SE-30 - 3%, Chrom. Q - 80/100.

Detector - ECD

LABORATORY #4

Extraction:

Sediments air-dried, weighed, Soxhlet extracted 7 hr. with 1+1
acetone-hexane.

Extract concentrated in a Kuderna-Danish evaporator.

Clean-up:

Extract partitioned with acetonitrile, back partitioned, concentrated and
passed through a Florisil column.

Concentrated on Kuderna-Danish and brought to volume.

Gas chromatographed and if sulphur present concentrate treated with
mercury and then re-chromatographed.

Qualification-Quantification:

Standards prepared specifically to bracket the sample g.c. responses.

Specific peaks used to quantify PCBs found.

Standards - pure - obtained from U.S. EPA Research Triangle Park.

Column - glass 6 ft. x 1/4" OD, SE-30/OV-210, 4/6% on Chromosorb WHP
80/100.

Detector - HECD.

LABORATORY #5

Extraction:

Samples directly transferred to Soxhlet extractor and extracted 10 hr.
with 1+1 Acetone-Hexane.

Extracts dried over sodium sulphate and concentrated in a Kuderna-Danish
evaporator.

Clean-up:

Samples passed through Florisil column, eluting with 200 ml 6% ethyl
ether in hexane. Florisil supplied from Research Triangle Park and
conditioned at 130°C before use.

Eluates concentrated with Kuderna-Danish evaporator.

Qualification-Quantification:

Peak ratio techniques via data system used to identify and quantify PCBs
found present.

Standards - from U.S. EPA Research Triangle Park.

Column - glass 10' x 1/8", SP-2100, 3% on Supelcoport 100/120.

Detector - Ni-63 ECD.

LABORATORY #9

Extraction:

Ultrasonic washed twice with 50 ml portions of acetone and then vacuum
filter sample + extracts through Celite.

Water backwash and extracted with methylene chloride (1x100 ml + 2x50 ml).
Concentrate (?) and take representative aliquot.

Clean-up:

Aliquot passed through Florisil column (6 mm ID x 24 cm) and eluted with
pesticide free grade hexane.

Samples concentrated in Kontes rotary evaporator, 3 ml isooctane added
and then evaporated with clean air and 30°C to ca. 3 ml. Extract
re-constituted to 5 ml using hexane.

Qualification-Quantification:

Standardize with 4+1 mixture 1254-1260 and use individual Aroclor
standards (1242, 48, 54, and 60) as needed to quantify PCB sample
components.

Standards - pure - obtained from Monsanto.

Column - glass, 12 ft. x 2 mm, Dexil 300 GC - 3%, Chrom. WHP 60/80.

Detector - Ni-63 ECD

Clean-up:

    Extract placed on pre-calibrated charge of Florisil and eluted with 30 ml
        of 6% ethyl ether in petroleum ether for PCB and other non-polar
        residues.

    Evaporated on steam bath to ca. 7 ml and transferred to culture tubes and
        brought to 1 ml for analysis.

Qualification-Quantification:

    Identify PCBs present by pattern recognition and quantify by peak heights
        using major peaks for each PCB found.

    Standards - pure - from U.S. EPA Research Triangle Park.

    Column - ? , 6 ft. x 2 mm ID, OV-210 and QF-1 - 5%, Supelcoport 80/100 and
        100/120.

    Detector - Ni-63 ECD.

LABORATORY #16

Extraction:

    Sample placed in glass stoppered flask and shaken with 40 ml acetone for
        20 min. 80 ml hexane then added and sample shaken another 10 min.
        Solvent layer decanted in separatory funnel and extraction process
        was repeated 2 more times.

    800 ml water added to separatory funnel and acetone partitioned from
        hexane layer. Extract then placed over sodium sulphate for 18 hrs. to
        dry.

    Extract transferred and concentrated in Kuderna-Danish evaporator and a 1
        ball Snyder column.

Clean-up:

    An aliquot placed over de-activated alumina, eluted with hexane, and
        concentrated to 1 ml in K-D.

    Treated with mercury to remove sulphur and placed over a silica gel column
        for further clean-up. Eluted with hexane and subsequently
        concentrated in K-D.

Qualification-Quantification:

Peak matching to standards and response factors applied. All peaks used
to sum totals for quantity present.

Standard - pure - from EPA.

Column - glass coil, 6 ft. x 2 mm ID, SP-2100 and SP-2250/2401 - 3% and
1.5/1.95%, Supelcoport 80/100 and 100/200.

Detector - Ni-63 ECD

LABORATORY #17

Extraction:

Sample stirred with 300 mg acetonitrile plus 10 ml sulphuric acid;
liquid decanted and 2 more extractions made with 100 ml each
acetonitrile. Extracts filtered through Celite pad.

Filtrate diluted 1:1 with distilled water and partitioned 3 times with
petroleum ether, 150, 90, and 60 mls, respectively. Petroleum ether
layer dried by filtering through sodium sulphate.

Extract concentrated to 2 ml with stream of pure air.

Clean-up:

Extracts placed over Florisil PR (30 g) and eluted with 300 ml hexane and
concentrated to 2 ml; then diluted to 10 ml with benzene.

Extract subjected to mercury to remove sulphur.

Qualification-Quantification:

Three Aroclors used for standardization (1242, 1254, and 1260) in
duplicate for each PCB analysis sequence, and Webb-McCall, J. Chrom.
Sci. 11, 3666 (1973) technique used for quantification.

Standards - purchased in solution, Nanogens, and diluted appropriately.

Column - stainless steel, 6 ft x 1/8" OD, SP-2100 - 3%, Supelcoport
100/120.

Detector - pulsed Ni-63 ECD.

LABORATORY #23

Extraction:

Total contents + 60g sodium sulphate mixed and extracted by Soxhlet for 8
hrs. with 300 ml 1+1 acetone-hexane.

Extract blown down using gentle stream of air.

Clean-up:

Extracts passed through ca. 20 gm Florisil with 200 ml 6% ethyl ether in
hexane.

Extracts injected and those containing interferring

sulphur were treated with several ml of ethanol and potassium
hydroxide pellets. Sample extracts placed in volumetric flasks for
analysis.

Qualification-Quantification:

Aroclor mixtures most closely approximating the specific sample
chromatogram were prepared from existing standards. All of the peaks in
the sample which match the standard for both retention time and peak shape
were summed and compared to those of the standard.

Standards - pure - from EPA Research Triangle Park or local State Dept. of
Agriculture.

Column - glass, 6 ft. x 1/4", SE-30 and OV-210 - 4% and 6%, Chromosorb W
80/100.

Detector - ECD.

LABORATORY #25

Extraction:

Samples air dried, ground, and an aliquot Soxhlet extracted with 1+1
acetone-hexane for 8 hrs. Evaporated in Kuderna-Danish to 10 ml.
Transferred to separatory funnel.

Twice partitioned hexane solution with acetonitrile and back partioned in
to hexane from acetonitrile - water phase. Washed hexane layer twice
with water. Dried hexane with sodium sulphate and evaporated in a
Kuderna-Danish evaporator.

Clean-up:

Sample passed through Florisil using hexane as eluant. Sample eluate
concentrated in Kuderna-Danish evaporator. Samples re-cleaned up
with a chromic acid oxidation coupled with a micro-Florisil column.

Qualification-Quantification:

Peaks for each Aroclor given a weight percent value, for 1254 first six
peaks and for 1260 last six peaks, peaks prominent and in common for
both PCBs not used.

Standards - pure - from U.S. EPA Research Triangle Park.

Column - glass, 6 ft x 2 mm ID, OV-210/OV-17 - 1.95/1.5%, Varaport 100/120.

Detector - Ni-63 ECD.

LABORATORY #27

Extraction:

Samples were air dried then extracted with homogenizer and 1+1
acetone-hexane (sample/solvent ratio?). Extracts were washed with
water to remove acetone; the water was back extracted with 15%
methylene chloride in hexane. Extracts combined, passed through
sodium sulphate and evaporated with a Kuderna-Danish evaporator.

Clean-up:

Sample extracts cleaned up with Florisil column using 6% methylene
chloride in hexane as eluant. Eluates then concentrated in
Kuderna-Danish evaporator.

Qualification-Quantification:

Individual standards compared with sample responses. Identified as
mixture of 1254, 1260 and quantified as such by average peak height
of selected peaks.

Standards - pure - from U.S. EPA Research Triangle Park.

Column - glass, 4 ft. x 4 mm, OV-1 3%, Sulpelcoport 80/120.

Detector - Ni-63 ECD.

LABORATORY #40

Extraction:

Wet samples mixed with (4+1) sodium sulphate-sample and extracted by
Soxhlet for 24 hrs. with 41% hexane in acetone. Extracts washed
three times with water to remove acetone, then water-acetone back
extracted with hexane.

Clean-up:

Combined extracts passed through silica gel/sodium sulphate column with 2%
ethyl ether in petroleum ether as eluant. Extracts concentrated in
Kuderna-Danish evaporator with 3 ball Snyder column. Extracts
further cleaned up with a 2% water deactivated silica gel column.
Sulphur removed by adding activated copper.

Qualification-Quantification:

Two standard curves of four concentrations were used. A least squares fit
for peak heights vs. concentration was used to quantify. Sample
results for PCBs 1254 and 1260 were first verified as 1+1 mixture by
individual calculations; then total was calculated using 1+1 standard
mixture of 1254, 1260.

Standards - pure, from U.S. EPA Research Triangle Park.

Column - glass, 6 ft. x 2 mm ID, OV-101 3%, Chromosorb WHP 80/100.

Detector - Ni-63 ECD.

LABORTORY #44

Extraction:

Samples dried to ca. 5% moisture at ambient temperature and extracted by
shaking for 2 hrs. with 1+1 acetone-hexane. Filtered aliquot
partitioned into hexane through water addition.

Clean-up:

Preliminary clean-up with activated Florisil; then PCBs fractionated on
coconut charcoal column. Extracts concentrated in a rotary vacuum
evaporator at 45°C.

Qualification-Quantification:

Individual standards of varying concentration used. Response patterns indicated 1+1 mixture of 1260, 1254. Total PCB quantified using summation of peak heights according to Dr. Reynolds numbering system.

Standards - pure, obtained several years ago from Dr. Lincoln M. Reynolds, Ontario Research Foundation.

Column - glass, 1.8 m x 2 mm ID, OV-17/OV-210 - 1.5/2.0%, Gas Chrom. Q 100/120.

Detector - constant current ECD.

LABORATORY #45

Extraction:

Samples were air dried for one day, mixed with sodium sulphate and extracted with Soxhlet overnight with hexane as solvent. Sample extracts were then concentrated (with KD?).

Clean-up:

Extracts cleaned up using activated Florisil PR grade with 6% ethyl ether in petroleum ether as eluant - 20g Florisil + 1/2" sodium sulphate + 200 ml eluant. Eluates evaporated using Kuderna-Danish apparatus.

Qualification-Quantification:

Individual standards of PCBs and their concentrations were chromatographed to establish linearity and patterns. Sample responses demonstrated 1260-1254 patterns. Using data system, sample responses were compared with stored integrations of standards and then quantified.

Standards - pure, from U.S. EPA Research Triangle Park. Standards varied somewhat from ampuls.

Column - glass, 305 cm x 2 mm ID, GE Viscasil 30,000 - 7.5%, Gas Chrom. Q 80/100.

Detector - Ni-63 ECD.

APPENDIX 2

GUIDELINES FOR CONTROL OF ANALYTICAL PROCEDURES

IN AN

INTRALABORATORY QUALITY CONTROL PROGRAM

GUIDELINES FOR CONTROL OF ANALYTICAL PROCEDURES
IN AN INTRALABORATORY QUALITY CONTROL PROGRAM

Scope:

The following guidelines are applicable to laboratories for all data
provided to the International Great Lakes Surveillance Program.

Assumptions:

1. The analytical methods used are appropriate for the surveillance tasks;
   they are essentially bias free, are capable of being brought into a state
   of statistical control at the precision required, and have adequate
   sensitivity to analyze environmental samples at the levels of interest.

2. Quality assurance procedures for field operations such as sample
   collection, container selection, preservation, transportation and storage
   have been satisfactorily implemented and are therefore not addressed
   herein. However, for ship laboratories and other field laboratories, it
   is understood that intralaboratory quality control may include analysis of
   field blanks and field duplicates.

3. The laboratory has designated the person or persons responsible for
   quality control together with development of an adequate reporting system
   such that the laboratory director and any other senior managers are kept
   apprised of the laboratory's performance and can substantiate it.

General Considerations:

Any analytical procedure that is in a state of statistical control will
have an inherent variability as one of its characteristics. For a given
procedure this variability is irreducible, that is, there is no identifiable
factor which contributes to procedure variation (no assignable cause).

The measure of procedure variability which we will use is the estimate of
the population standard deviation. The specific population of interest is the
population of between run analyses; between run analyses are chosen rather

than within run analyses because we are interested in monitoring performance across runs. However, with highly labile constituents it may be necessary to use an estimate of the standard deviation of the population of within run analyses.

To obtain a reliable initial estimate of the population standard deviation 40 to 50 data are needed. They may be either duplicates analysed in separate runs or analyses of a stable standard in separate runs; examples of both will be given. Highly labile constituents may, however, require an estimate based on duplicates analysed in the same run.

Once the estimate is obtained, control limits can be set for the analytical procedure which, if exceeded, indicate that the procedure is probably out of control. The control limits are commonly set at 3 standard deviations (3σ limits). These limits imply an $\alpha = 0.0027$ or about 3 chances in 1000 of judging an in control procedure to be out of control.

Control limits are generally incorporated into control charts which provide an immediate visual record of performance. If a procedure goes out of control, the point(s) at which control is lost can be easily identified.

Two types of control charts can be differentiated: those that monitor accuracy and those that monitor precision. An example of the former is a chart that monitors results on a known, stable standard; violation of a control limit indicates that the analytical procedure is not producing accurate results. The difficulty may be due to bias, may be due to a loss of precision, or may stem from a combination of the two. An example of the latter is a chart that monitors the range of duplicate analyses on a sample whose value is unknown; violation of a control limit indicates that precision has been lost. However, information regarding possible bias is not provided either by control limit violation or the lack thereof.

Estimating Analytical Procedure Variability:

The essential first step in developing a control system for an analytical procedure is to acquire a sound estimate of procedure variability when the procedure is in a state of statistical control. Once the estimate has been

- 72 -

obtained, it can be used to set control limits for the monitoring of both accuracy and precision.

Example 1 - Using Duplicates

Consider the following 50 pairs of results, in µg/L, on duplicates which were analysed in different runs.

1st Result	2nd Result	Range	1st Result	2nd Result	Range
50	46	4	39	42	3
37	36	1	25	25	1
22	19	3	20	18	2
17	20	3	12	10	2
32	34	2	28	32	4
46	46	0	35	40	5
26	28	2	22	22	0
26	30	4	26	25	1
61	58	3	41	40	1
44	45	1	20	21	1
40	44	4	22	40	18
36	35	1	37	35	2
29	31	2	29	26	3
26	38	12	34	35	1
36	36	0	17	19	2
47	45	2	43	44	1
16	20	4	56	53	3
18	21	3	30	32	2
26	22	4	20	21	1
35	36	1	36	32	4
26	25	1	43	39	4
49	51	2	22	21	1
33	32	1	35	36	1
40	38	2	53	50	3
16	13	3	47	47	0

Two of the ranges obtained, 12 and 18, strongly suggest that the analytical system was out of control; these two values are discarded. The remaining 48 ranges are summed and the average range, $\bar{R}$, found.

$$4 + 1 + 3 \ldots + 3 + 0 = 101$$
$$\bar{R} = 101/48 = 2.104$$

An estimate of the standard deviation, s, is obtained from the average range of duplicate analyses by dividing by 1.128, the proper factor for acquiring a standard deviation estimate from ranges derived from duplicates.

$$s = \frac{2.104}{1.128} = 1.865 \ \mu g/L$$

Example 2 - Using a Stable Standard

Consider the following 50 results, in $\mu g/L$, obtained by analysing a stable standard in separate runs.

35.1	31.8	36.4	33.8	33.0	34.2
33.2	35.0	32.1	34.3	37.2	33.7
33.7	31.4	24.7	32.9	34.3	33.9
35.9	35.6	38.2	34.2	32.7	35.6
33.5	30.2	33.1	35.6	34.1	40.1
34.5	32.7	34.9	31.5	35.8	34.6
34.4	31.1	36.2	36.4	33.9	33.8
49.6	34.8	34.0	32.6	35.5	33.0
34.3	35.3				

Mean of all values = 34.368
Mean of 48 values (omitting 24.7 and 49.6), $\bar{x}$ = 34.252083

The two values 24.7 and 49.6 clearly indicate that the procedure was out of control; they are discarded. The value 40.1 is marginal and represents a

more difficult decision; in this example it is left in, provisionally.

The estimate of the standard deviation, s, is obtained in the usual way.

$$s^2 = \frac{\Sigma Xi - N\overline{X}^2}{N-1}$$

$$s^2 = \frac{56,470.35 - 48 (34.252083)^2}{47}$$

$$s^2 = 3.32978$$

$$s = 1.825 \; \mu g/L \quad \text{(provisional value, see Example 7 below)}$$

If the two omitted values had been included in the calculation, the estimated standard deviation would have been a badly inflated 3.138 μg/L.

It should be noted that s is expressed in absolute rather than relative terms. If variability were proportional to concentration, then the relative standard deviation (coefficient of variation) would be appropriate, but we are not aware of any analytical procedures so characterized. It appears that for any given practical working range variability may be treated as a constant with minimal ill effects. However, if very different ranges are employed to determine the same constituent an estimate of the standard deviation will be required for each range. One would not expect the variability which characterizes analyses in the range 0-100 μg to also pertain to analyses in the range 0-10 mg.

As additional data are obtained initial estimates of variability can be put on a sounder footing by pooling with estimates from the new information, assuming that no substantial change is apparent. If a procedure's variability

appears to change significantly, the procedure should be carefully reviewed to ascertain the cause.

The following method may be used to test for change in procedure variability.

Example 3 - Testing for Change in Variability

Suppose an initial estimate of an analytical procedure's standard deviation is obtained, s_1 = 1.796 µg/L, based on a data set of 61 items and therefore having associated with the estimate 60 degrees of freedom. A new estimate, s_2 = 2.145 µg/L, is then obtained based on 41 additional measurements, and thus having 40 degrees of freedom.

The ratio of the two estimates of the <u>variance</u> is found,

$$\frac{s_1^{\,2}}{s_2^{\,2}} = \frac{1.796^2}{2.145^2} = \frac{3.225616}{4.601025} = 0.701$$

and the ratio compared to appropriate values of the F distribution.

Testing at an α-level = 0.05, the appropriate upper value is simply the tabulated value for the upper 2.5% point of the F distribution with 60 and 40 degrees of freedom; this tabulated value is 1.80. Obtaining the appropriate <u>lower</u> value requires a little arithmetic. The tabulated value for the upper 2.5% point of the F distribution with 40 and 60 degrees of freedom (note the reversal) is found and its reciprocal taken, 1/1.74 = 0.575, to give the required value.

Since the ratio of the two estimates of the analytical procedure variance, 0.701, lies between the values 0.575 and 1.80, we would <u>not</u> conclude that the variability of the procedure had changed.

This test differs from the usual F test in that it is two-tailed, there being no a priori reason for assuming that one variance estimate would be greater than the other.

When it appears that the variability of an analytical procedure has not changed, a pooled estimate of variability may be obtained as follows:

Example 4 - Pooling Estimates of Variability

The pooling method consists of weighing the two _variance_ estimates by the size of the respective data sets from which they were obtained, summing the weighed variance estimates, and dividing the sum by the sum of the degrees of freedom associated with the two estimates. The quotient which results is the pooled variance estimate, s^2, from which the new, pooled estimate of the standard deviation, s, is obtained.

Using the data of Example 3 we have

$$s^2 = \frac{61(1.796)^2 + 41(2.145)^2}{60 + 40}$$

$$s^2 = \frac{196.7626 + 188.6420}{100}$$

$$s^2 = 3.854$$

$$s = 1.963 \ \mu g/L$$

When a pooled estimate of the procedure standard deviation is obtained, new control limits should be calculated using the revised estimate.

Setting Control Limits:

There are two goals in setting control limits. They should be close enough to signal when there is trouble with a system, and they should be distant enough to discourage tinkering with a system that is operating within its capabilities. Since these two goals are antithetical, a compromise is necessary. The compromise which has been found satisfactory in a great many applications is the use of 3σ control limits, and they are illustrated here.

Example 5 - Use of a Known

A known sample whose concentration is 32.7 µg/L is analysed by a procedure whose estimated standard deviation is 2.131 µg/L. The control limits are 32.7 ± 3 x 2.131 or 26.31 and 39.09. Assuming that results can be read to tenths of a microgram, a result ≥26.3 and ≤39.1 is judged acceptable.

Example 6 - Use of an Unknown Duplicate

An unknown duplicate sample is analysed in separate runs by a procedure whose estimated standard deviation is 1.537 µg/L. The control limit for the range of the two analyses is 1.537 x 3.686 or 5.67; 3.686 is the proper factor for duplicate ranges. Assuming that results can be read to tenths of a microgram, a pair of results whose range is ≤5.7 is judged acceptable.

Example 7 - Correcting an Initial Estimated Standard Deviation

In Example 2 the value 40.1 µg/L was provisionally allowed to remain in the data set for which an estimated standard deviation of 1.825 µg/L was obtained. We now determine whether the 40.1 should remain in the data set.

From the results of Example 2 we can calculate the 3σ control limits 34.252 ± 3 x 1.825 or 28.8 and 39.7.

Since 40.1 is larger than the upper control limit 39.7, there is sufficient evidence to discard this value also.

The estimate of the standard deviation is now recalculated from the 47 item data set to give s = 1.626 µg/L. The new sample mean is 34.128, resulting in new control limits of 29.3 and 39.0 which encompass the 47 values remaining in the data set.

Example 8 - A Special Case, Use of Recovery Data

The use of recovery data for control purposes presents some special problems which are dealt with in this example. We begin with estimation of the variability associated with the determination of recoveries.

Consider the following data set, values in mg/L:

1.	2.	3.	4.	5.	6.
Spiked Recovery	Unspiked Result	Apparent Recovery	True Spike	Deviation From Expected	% Recovery
1.91	0.68	1.23	1.30	-0.07	94.615
1.78	0.57	1.21	1.30	-0.09	93.077
1.53	0.23	1.30	1.30	0	100
1.74	0.15	1.59	1.30	0.29	122.308
2.10	0.53	1.57	1.30	0.27	120.769
1.82	0.61	1.21	1.30	-0.09	93.077
2.07	0.54	1.53	1.30	0.23	117.692
1.39	0.14	1.25	1.30	-0.05	96.154
1.16	0.20	0.96	1.30	-0.34	73.846
1.55	0.19	1.36	1.30	0.06	104.615
2.02	0.41	1.61	1.30	0.31	123.846
1.58	0.36	1.22	1.30	-0.08	93.846
13.01	11.97	1.04	1.30	-0.26	80
1.46	0.17	1.29	1.30	-0.01	99.231
1.63	0.31	1.32	1.30	0.02	101.538
11.95	10.98	0.97	1.30	-0.33	74.615
1.68	0.27	1.41	1.30	0.11	108.462
1.83	0.47	1.36	1.30	0.06	104.615
1.62	0.43	1.19	1.30	-0.11	91.538
5.04	3.96	1.08	1.30	-0.22	83.077
2.53	1.22	1.31	1.30	0.01	100.769
2.69	1.09	1.60	1.30	0.3	123.077
1.50	0.25	1.25	1.30	-0.05	96.154
2.73	0.24	2.49	1.30	1.19	191.538
2.86	0.23	2.63	1.30	1.33	202.308
1.77	0.51	1.26	1.30	-0.04	96.923
1.88	0.55	1.33	1.30	0.03	102.308
0.90	0.57	0.33	1.30	-0.97	25.385
2.22	0.95	1.27	1.30	-0.03	97.692
1.99	0.85	1.14	1.30	-0.16	87.692
1.54	0.26	1.28	1.30	-0.02	98.462
1.47	0.15	1.32	1.30	0.02	101.538
1.43	0.09	1.34	1.30	0.04	103.077
1.65	0.35	1.30	1.30	0	100
1.91	0.68	1.23	1.30	-0.07	94.615
2.06	0.93	1.13	1.30	-0.17	86.923
5.24	4.02	1.22	1.30	-0.08	93.846
1.58	0.27	1.31	1.30	0.01	100.769
1.63	0.28	1.35	1.30	0.05	103.846
1.52	0.23	1.29	1.30	-0.01	99.231
1.70	0.35	1.35	1.30	0.05	103.846
1.77	0.31	1.46	1.30	0.16	112.308
1.93	0.49	1.44	1.30	0.14	110.769
2.30	1.13	1.17	1.30	-0.13	90

In column 5 there are 3 deviations from expected recoveries which appear
extreme: 1.19, 1.33 and -0.97; these results are discarded. From the
remaining 41 results in the 5th column of the data set an estimate of the
standard deviation of the spiking recovery procedure is calculated in the
usual way and found to be s = 0.1532 mg/L. (Since the deviations from
expected results represent the difference between two analytical
determinations, we would expect the standard deviation of the spiking recovery
procedure to be greater than the standard deviation of a single determination
by a factor of $\sqrt{2}$.)

The mean of the deviations from the expected results is -0.0061 mg/L. Since
the absolute value of this mean is less than the standard error of the mean of
the spiking recovery procedure, s_m (= 0.1532/$\sqrt{41}$ = 0.024 mg/L),
the spiking recovery procedure appears to be unbiased with complete recovery a
reasonable expectation. Control limits may therefore be set around the
expectation of complete recovery with allowable deviations of 0 ± 3 x 0.1532
or -0.46 mg/L and 0.46 mg/L. The remaining 41 members of the data set are all
within these limits.

Had the spiking recovery procedure demonstrated a bias, the control limits
would have been calculated from the estimate of the bias.

In this example the data in column 6 may be used to obtain equivalent
control limits in terms of percent recovery. With the omission of the 3
questionable results, the estimate of the standard deviation of the spiking
recovery procedure is 11.782% on a spike of 1.3 mg/L; 11.782% of 1.3 mg/L is
0.1532 mg/L, which is the same estimate as obtained from column 5. However,
the equivalency holds because identical spikes were employed in all
recoveries. If variable spikes are used, then the estimate of the standard
deviation and the ensuing control limits must be made in absolute units such
as mg/L and not in percent recovery.

Frequency of Control Sample Use:

The following minimum frequencies are recommended for the use of control
samples:

To monitor accuracy, 1 quality control sample of known value should be
included with every 15 analyses or with each batch, whichever results in
the greater frequency;

To monitor precision, 1 quality control sample should be included with
every 15 analyses or with each batch, whichever results in the greater
frequency. If duplicates are used to monitor precision, they should be
analysed in different runs when a between run measure of variability is
employed in setting control limits.

Concluding Remarks:

The control charts which employ control limits as illustrated above are
examples of Shewhart control charts. The reference used in preparing these
guidelines is:

ASTM Manual on Presentation of Data and Control Chart Analysis
ASTM Special Technical Publication 15D, 1976

The factors 1.128 and 3.686 used in examples 1 and 6, respectively, were taken
from this manual.

Recognition that problems exist is, of course, but the essential first
step toward their solution; one authority on quality control has suggested
that it represents only 10% of the effort which will be required. For the
intermittent problems which often occur in analytical chemistry, the
identification of causes will typically be arduous. For such recurrent
problems, careful record keeping will be required to determine whether rates
of occurrence have in fact diminished when putative causes are addressed.

APPENDIX 3

REPORTING LOW LEVEL DATA

REPORTING LOW LEVEL DATA

There are specific problems in the reporting of low level data which are associated with the question: is a substance present? While this question is seldom germane in IJC work concerned with loading estimation, it has so influenced thinking about reporting low level data that it seems best to consider it in some depth before dealing with how such data are to be reported for IJC purposes.

In answering the question "is a substance present?", there are two possible correct conclusions which may be reached. One may conclude that the substance is present when it is present, and one may conclude that the substance is not present* when it is not present. Conversely, there are two possible erroneous conclusions which may be reached. One may conclude that the substance is present when it is not, and one may conclude that the substance is not present when it is. The first kind of error, finding something which is not there, is called a TYPE I ERROR. The second kind of error, not finding something which is there, is called a TYPE II ERROR.

These two types of errors are illustrated in the material that follows, using the result which might be obtained from a single analysis when the substance is not present to illustrate Type I error and the inferences that might be drawn from a single analysis at two different actual concentrations to illustrate Type II error.

Of course inferences as to water quality are seldom, if ever, based on the result of a single analysis. A single result is used here to simplify the exposition.

---

*Since Avogadro's number is very large, a pedant could argue that one should never claim that a substance is not present. A common sense meaning of not present is intended here, i.e. if measurement is being made in micrograms per litre the presence of a few nanograms per litre is irrelevant.

If the standard deviation, σ, of an analytical procedure has been determined at low concentrations including 0, then the probability of making a Type I error can be set by choosing an appropriate α level to determine the Criterion of Detection.†

For example, suppose that the standard deviation, σ, of an analytical procedure is 6 μg/litre and that an α of 0.05 is deemed acceptable so that the probability of making a Type I error is set at 5%. The Criterion of Detection can then be found from a table of cumulative normal probabilities to be 1.645 σ = 1.645 x 6 μg/litre ≃ 10 μg/litre.
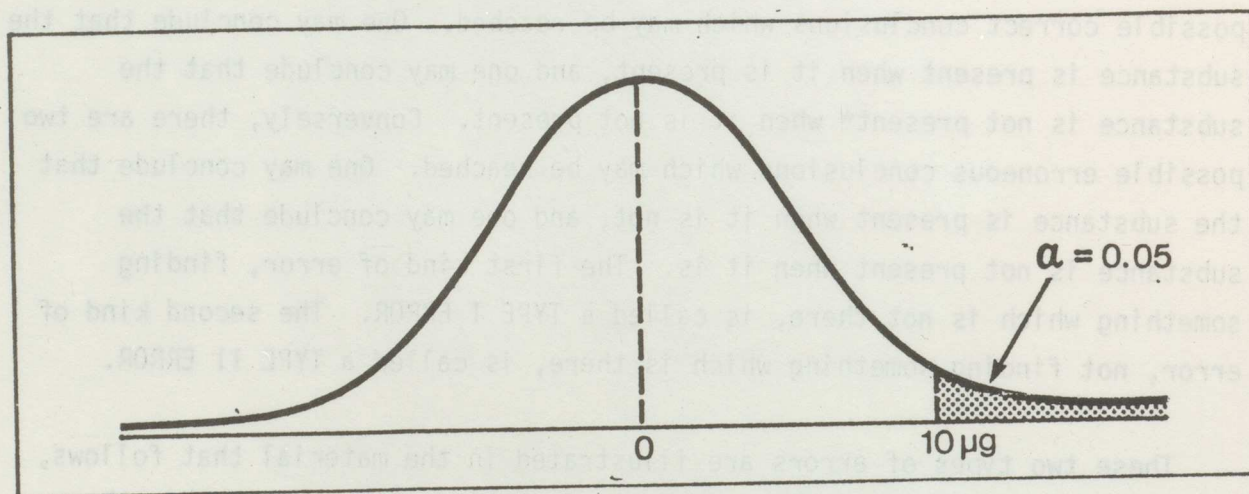
Any value observed below 10 μg/litre would be reported as less than the Criterion of Detection, since to report such a value otherwise would increase the probability of making a Type I error beyond 5%.

Note that the context of decision is the analytical result produced by the laboratory. A result is obtained and a response made to it. Nothing has been said concerning the ability to detect a substance which is present at a specified concentration.

---

†Criterion of Detection may be a new term to some. It refers to the minimum analytical result which must be observed before it can be stated that a substance has been discerned with an acceptable probability that the statement is true. The terms Detection Limit or Limit of Detection are often used with this meaning, but in this Handbook they are reserved for a more appropriate usage.

Once the Criterion of Detection has been set, the probability of making a Type II error, β, or its complement 1-β, the probability of discerning the substance when it is present, can be determined for <u>given</u> <u>true</u> <u>situations</u>. (The probability 1-β is sometimes called the power of the test).

Consider the same analytical procedure as above with a Criterion of Detection of 10 µg/litre. Suppose that the concentration of the sample being analyzed is 10 µg/litre, i.e. the concentration is equal to the Criterion of Detection. If, all analytical results below the Criterion of Detection were reported as such, then the probability of discerning the substance would be 0.5 or 50%.

Do not discern

Discern

10 µg

$\beta = 0.5$ $1 - \beta = 0.5$
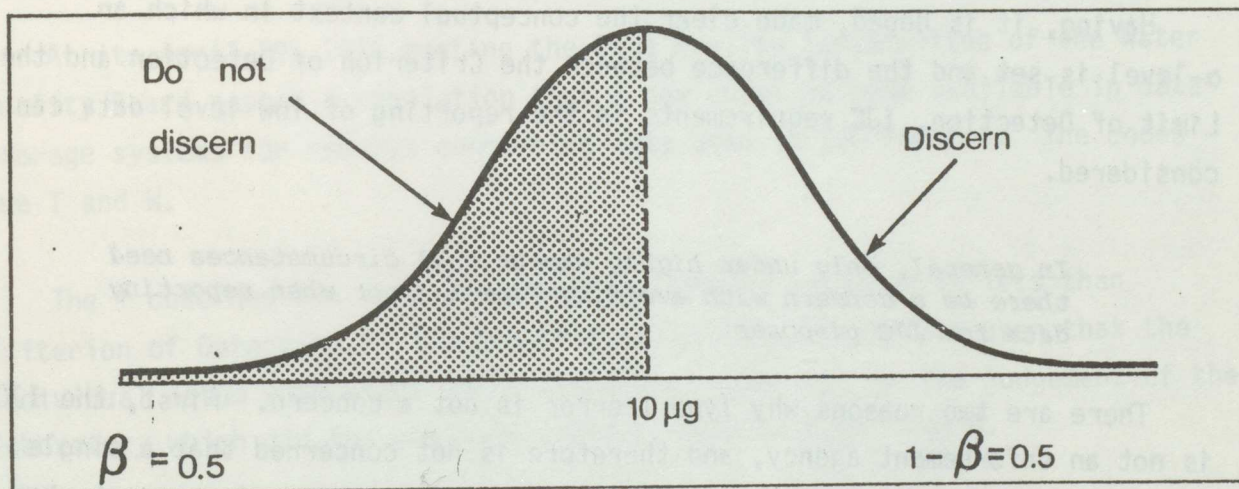
Conversely, the probability of making a Type II error and failing to discern the substance would also be 0.5. From this example it can be seen that the probability of discerning a substance when its concentration is equal to the Criterion of Detection is hardly overwhelming. In order for the probability of a Type II error to be equal to the probability of a Type I error, β = α, then the concentration of the sample being analyzed must be twice the Criterion of Detection.

$\beta = 0.05$

10 µg   20 µg
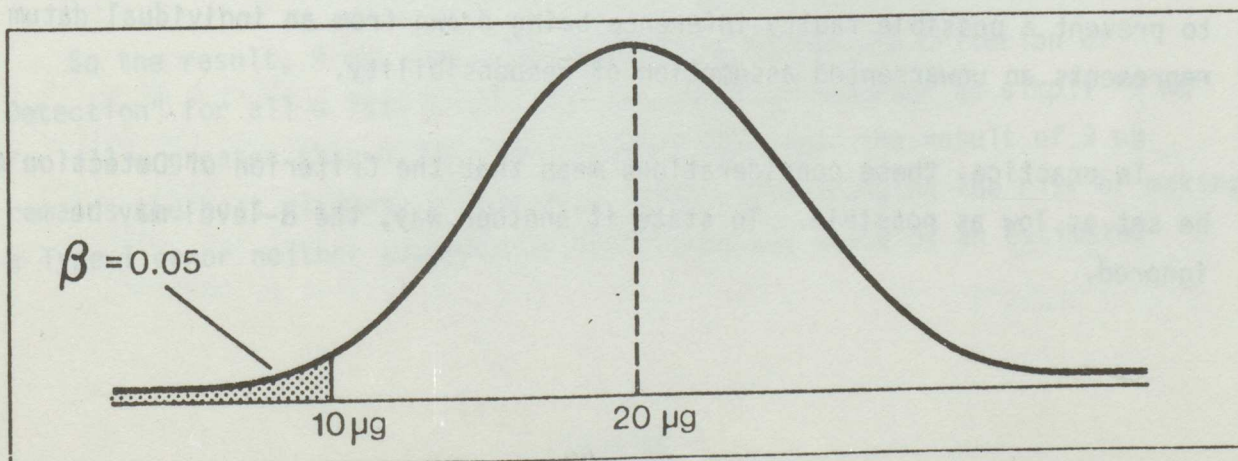
This concentration of twice the Criterion of Detection <u>is</u> <u>the</u> <u>Limit</u> <u>of</u> <u>Detection</u> when it has been decided that the risk of making a Type II error is to be equal to the risk of making a Type I error.

The concept of Type II error has been emphasized because it is usually ignored. Generally, attention is paid to the avoidance of Type I error with no consideration given to the probability of making a Type II error. It should also be recognized that when the probability of making a Type I error is decreased by selecting a lower α-level, the probability of making a Type II error is increased.

Having, it is hoped, made clear the conceptual context in which an α-level is set and the difference between the Criterion of Detection and the Limit of Detection, IJC requirements in the reporting of low level data can be considered.

> In general, only under highly exceptional circumstances need there be a concern with avoiding Type I error when reporting data for IJC purposes.

There are two reasons why Type I error is not a concern. First, the IJC is not an enforcement agency, and therefore is not concerned that a single datum will lead it into a false accusation that a substance is present when it is not. Second, in virtually all cases data are aggregated for IJC purposes in order to provide estimates of loadings and/or concentrations; therefore the avoidance of Type I error relates to data sets and not to the individual datum.

This second point is crucial. Rarely, if ever, will the analytical chemist have responsibility for inference from data sets or even be in a position to know which data may be combined. Therefore, censoring of results to prevent a possible faulty inference being drawn from an individual datum represents an unwarrented assumption of responsibility.

In practice, these considerations mean that the Criterion of Detection may be set as low as possible. To state it another way, the α-level may be ignored.

*On the other hand, when reporting data for IJC purposes
every effort must be made to avoid Type II error.*

The reason is obvious. When results are reported as "less than" or "below
the Criterion of Detection," they are virtually useless for either estimating
loadings or concentrations.

In practice, this consideration means that if a number can be obtained, it
is to be reported.

CODES TO BE USED IN REPORTING LOW LEVEL DATA

At its April 12, 1976 meeting the Data Quality Subcommittee of the Water
Quality Board passed a resolution that 2 new codes be made available in data
storage systems for remarks concerning data used in IJC reports. The codes
are T and W.

The T code has the following meaning: "Value reported is less than
Criterion of Detection." 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 which did the analysis, differ significantly from 0.

It should be recognized than an implied significance test which fails to
reject the null hypothesis that a result does not differ from a standard value
in no way diminishes the value of the result as an estimate. To illustrate:
a result of 9 μg on a test whose σ = 6 μg can not be regarded as
significantly different from 0 for any α-level less than 0.067; however, if
a significance test were made with α = 0.1, then the null hypothesis would
be rejected and the result deemed significantly different from 0.

So the result, 9 μg, could be reported as "Below the Criterion of
Detection" for all α less than 0.067 and could be reported as simply "9 μg"
for all α greater than 0.067. But however reported, the result of 9 μg
remains the best estimate of the true value since changing the risk of making
a Type I error neither augments or diminishes the value of an estimate.

It may be added that low level results are better estimates, in the sense of being more precise, than higher results since for all analytical tests with which we are acquainted the standard deviation of the test increases with the concentration.

The W code has the following meaning: "Value observed is less than 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.

The following example illustrates the use of the codes:

Suppose that a laboratory has determined that its Criterion of Detection for total phosphorus is 10 µg/litre, and suppose in addition that the smallest increment that can be read on the analytical device corresponds to a concentration of 2 µg/litre. Given these conditions, any value observed >10 µg/litre would be reported without an accompanying code; any value observed >2 µg and <10 µg would be reported with the T code; if no instrument response were observed, the result would be reported as 2W.

REPORTING NEGATIVE RESULTS

With many analytical procedures there will always be an instrument response, so the W code will not apply. In particular, this lack of applicability will occur when a result is obtained through subtraction of a blank correction. In this case negative results will often be obtained; in fact, if the constituent of interest is not present, one would expect negative results to occur as often as positive.

In order that valid inferences may be made from surveillance data, it is important that negative results be reported as such. Consider the following three different ways of reporting the same results. The left hand column gives results in a heavily censored form; the center column has negative results censored; the right hand column gives the results as obtained.

<3 µg	2 µg	2 µg
<3	0	-2
<3	0	-1
4	4	4
3	3	3
<3	0	-3
<3	1	1
<3	0	-1
<3	0	0
<3	2	2

Nothing can be done with the results in the left hand column except to conclude that we don't know whether the constituent is present or not; the sampling and analytical effort have been wasted.

If the results in the center column were taken at face value, one could conclude that the mean concentration was 1.2 µg with a standard error of the mean of 0.467 and 95% confidence limits for the mean of 0.14 µg and 2.26 µg. Since the confidence limits do not include zero, it would appear that the evidence supports the presence of the constituent.

Analysis of the uncensored results of the right hand column gives a mean concentration of 0.5 µg, a standard error of the mean of 0.719, and 95% confidence limits for the mean of -1.13 µg and 2.13 µg. The correct conclusion can be drawn that the evidence is insufficient to support the presence of the constituent.

Note that the censored data of the center column distort both the mean and the standard error of the data, making the data appear more precise than they are.

Of course any result of 0 or less which is reported should be reported with the T code.

APPENDIX 4

DATA QUALITY WORK GROUP MEMBERSHIP LIST

MEMBERSHIP LIST

DATA QUALITY WORK GROUP

Mr. Keijo I. Aspila (Chairman)
Quality Assurance and Methods Section
Analytical Methods Division, NWRI
Canada Centre for Inland Waters
Burlington, Ontario

Mr. James H. Adams, Jr.
Head
Quality Assurance Office
Surveillance and Analysis Division
U.S. EPA, Region V
Chicago, Illinois

Mr. J. N. Bishop
Assistant Director
Laboratory Services Branch
Ministry of the Environment
Rexdale, Ontario

Mr. James C. Daly
Quality Assurance Director
New York State Dept. of Health
Albany, New York

Dr. M. D. Mullin
Chemist Manager
Large Lakes Research Station
U.S. Environmental Protection Agency
Grosse Ile, Michigan

Mr. John H. Peck
Environmental Laboratory
Michigan Department of Natural Resources
Lansing, Michigan

Mr. C. Ross
Director
Central Regional Laboratory
U.S. EPA, Region V
Chicago, Illinois

Secretariat Responsibilities
Mr. R. E. White
Senior Scientist
Great Lakes Regional Office
International Joint Commission
Windsor, Ontario