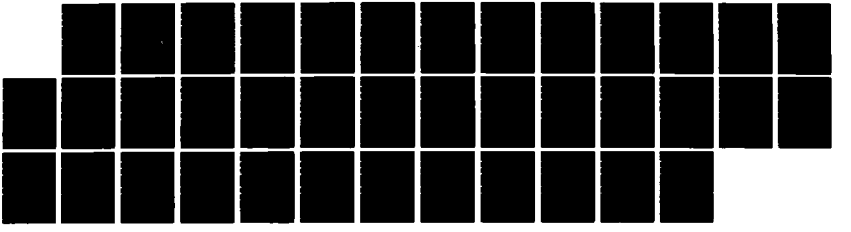
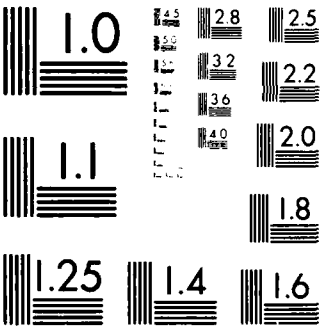
AD-A170 180 PROTOCOL DEVELOPMENT AND EQUIVALENCY TESTING OF THE FACTS PROCEDURE FOR C... (U) FLORIDA INTERNATIONAL UNIV MIAMI DRINKING WATER RESEARCH CENT... W J COOPER

1/1

UNCLASSIFIED 15 MAR 84 DAMD17-80-C-0164 F/G 7/4 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

(12)

AD

PROTOCOL DEVELOPMENT AND EQUIVALENCY TESTING OF THE FACTS PROCEDURE FOR CHLORINE RESIDUAL DETERMINATION IN DRINKING WATER

AD-A170 180

FINAL REPORT

William J. Cooper

March 15, 1984

Supported by

U.S. ARMY MEDICAL RESEARCH AND DEVELOPMENT COMMAND
Fort Detrick, Frederick, Maryland 21701-5012

Contract No. DAMD17-80-C-0164

Drinking Water Research Center
Florida International University
Miami, FL 33199

DTIC
ELECTE
JUL 2 3 1986
S
A

DOD DISTRIBUTION STATEMENT
Approved for public release; distribution unlimited

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents

DTIC FILE COPY

86 0

SECURITY CLASSIFICATION OF THIS PAGE

AD-A170 181

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Florida International University 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION U.S. Army Medical Bioengineering Research and Development Laboratory
6c. ADDRESS (City, State, and ZIP Code) Drinking Water Research Center Miami, FL 33199	7b. ADDRESS (City, State, and ZIP Code) Fort Detrick Frederick, Maryland 21701-5010
8a. NAME OF FUNDING / SPONSORING ORGANIZATION U.S. Army Medical Research & Development Command 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER DAMD17-80-C-0164
8c. ADDRESS (City, State, and ZIP Code) Fort Detrick Frederick, Maryland 21701-5012	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. 62777A; PROJECT NO. 3E1-62777A878; TASK NO. 00; WORK UNIT ACCESSION NO. 297

11 TITLE (Include Security Classification)
Protocol Development and Equivalency Testing of the FACTS Procedure for Chlorine Residual Determination in Drinking Water

12. PERSONAL AUTHOR(S)
William J. Cooper

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Final Report	FROM 80/9/1 TO 83/8/31	1984 March 15	34

16. SUPPLEMENTARY NOTATION

17 COSATI CODES		
FIELD	GROUP	SUB-GROUP
07	03	
06	09	

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Chlorine Analysis, FACTS, DPD, Free available chlorine Equivalency Testing, Drinking Water, Field Army

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The FACTS test procedure was tested at 16 water treatment laboratories throughout the U.S. It was found to be equivalent for the determination of free available chlorine to both the DPD test and amperometric titration and is approved by the U.S. environmental Protection Agency for use in drinking water treatment plants.

Several instrumental procedures for chlorine residual analysis were also tested. An on line continuous amperometric procedure was shown to be operator independent and could be used for control and monitoring of chlorine in water treatment processes.

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL Mrs. Virginia M. Miller	22b. TELEPHONE (Include Area Code) 301/663-7325 22c. OFFICE SYMBOL SGRD-RMS

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE

AD

PROTOCOL DEVELOPMENT AND EQUIVALENCY TESTING OF THE FACTS PROCEDURE FOR CHLORINE RESIDUAL DETERMINATION IN DRINKING WATER

FINAL REPORT

William J. Cooper

March 15, 1984

Supported by

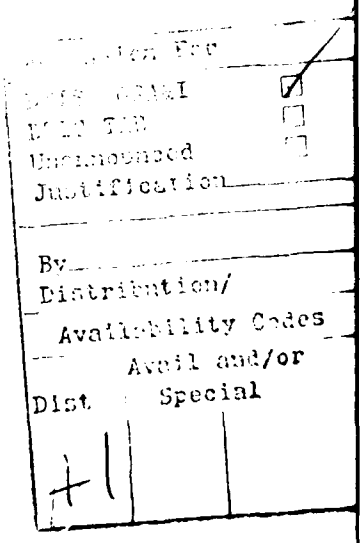
U.S. ARMY MEDICAL RESEARCH AND DEVELOPMENT COMMAND
Fort Detrick, Frederick, Maryland 21701-5012

Contract No. DAMD17-80-C-0164

Drinking Water Research Center
Florida International University
Miami, FL 33199

DOD DISTRIBUTION STATEMENT
Approved for public release; distribution unlimited

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents



EXECUTIVE SUMMARY

The objective of this project was to provide the US Environmental Protection Agency (EPA) data necessary for nationwide approval of the FACTS test procedure for determining free available chlorine for "National Interim Primary Drinking Water Regulations" (NIPDWR) compliance monitoring. Based on the minimum requirement for NIPDWR nationwide approval comparability testing, a detailed protocol was developed and testing conducted.

In all, 16 water treatment plants cooperated in the equivalency testing. Ten compared the FACTS test to the DPD, and six compared the FACTS test to the amperometric titration. The range of concentrations obtained for the FACTS and DPD comparison was from 0.4 to 2.3 mg/L as Cl_2; the range for the FACTS and amperometric titration comparison was from 0.55 to 2.7 mg/L as Cl_2. At most treatment plants a prechlorination and postchlorination site were used in the testing.

Sixty samples were analyzed at each plant, yielding 192 data points per plant. Forty-eight of these samples were paired comparisons of the FACTS with either the DPD or amperometric titration. The paired comparisons were randomly assigned to two operators, at two sites (where possible) for each plant. For the remaining twelve samples, six for each operator, four replicate analyses were run using each method.

The data summary for the comparison between FACTS and DPD is shown in Figure 1 and in for the comparison between the FACTS and amperometric titration in Figure 2.

No statistical difference was observed in the analysis of the summary data. Results are summarized in the report for the individual water treatment plants. The report details the results for the comparison of FACTS and DPD, and of FACTS and amperometric titration.

As a result of these tests the EPA has approved the FACTS test procedure for compliance monitoring of free available chlorine at water treatment plants.

It is recommended that the US Army review the requirements for a free available chlorine test for field use. Consideration should be given to adopting the FACTS test procedure for field Army use, because of its superior specificity for free chlorine, its equivalency to the presently used DPD test, and its availability in test kit form.

In addition to the protocol development and equivalency testing of the FACTS test procedure, a study was conducted to compare several instrumental methods for determining chlorine in drinking water. It was concluded that the total chlorine analyzer was analyst independent and is capable of monitoring chlorine continuously. This has potential applications in water purification units where chlorine is added as the disinfectant and can be used to monitor and control chlorine.

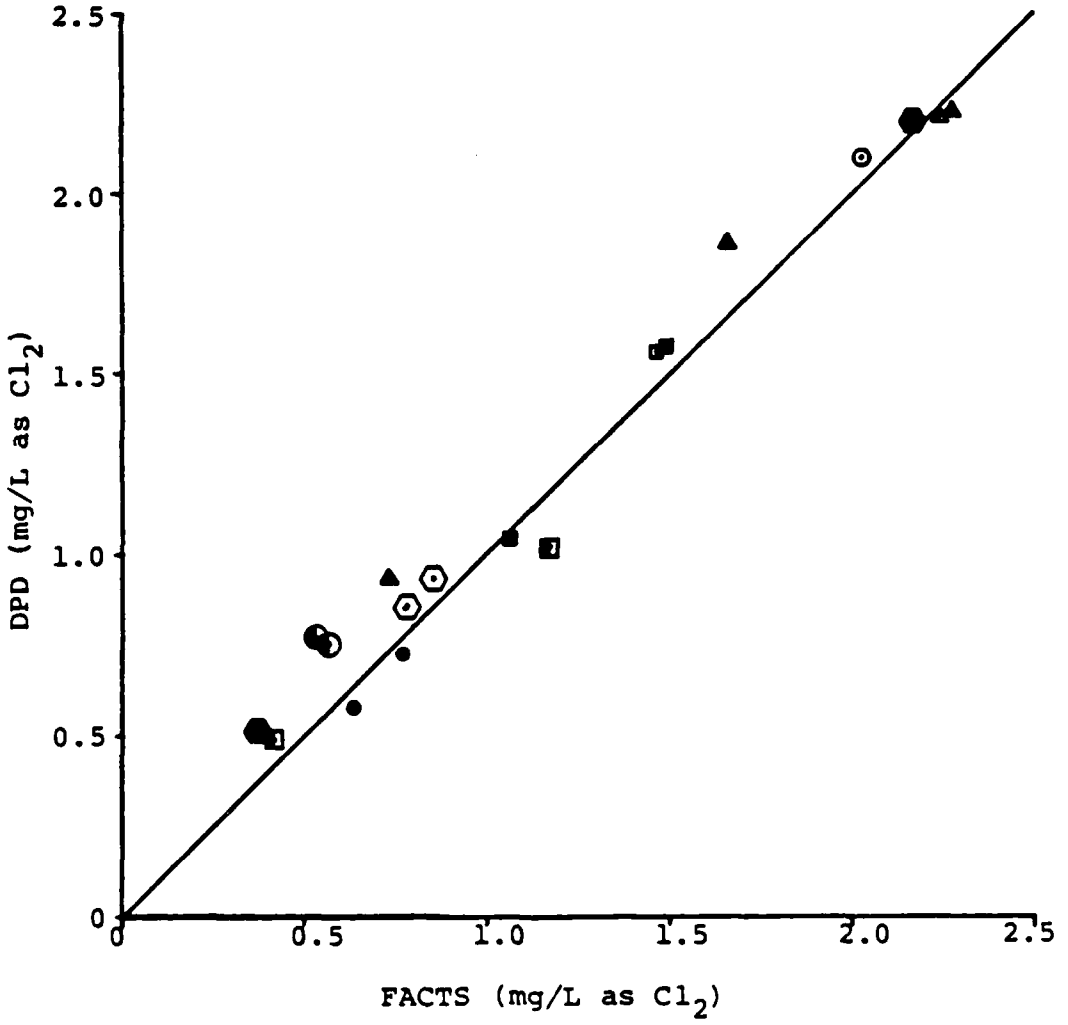


Figure 1. Comparison of Means for FACTS and DPD Test Procedures for Free Available Chlorine Obtained at Ten Water Treatment Plants.

Reprinted from JOURNAL AWWA, Vol. 75, No.12 (December 1983), by permission. Copyright© 1983, The American Water Works Association.

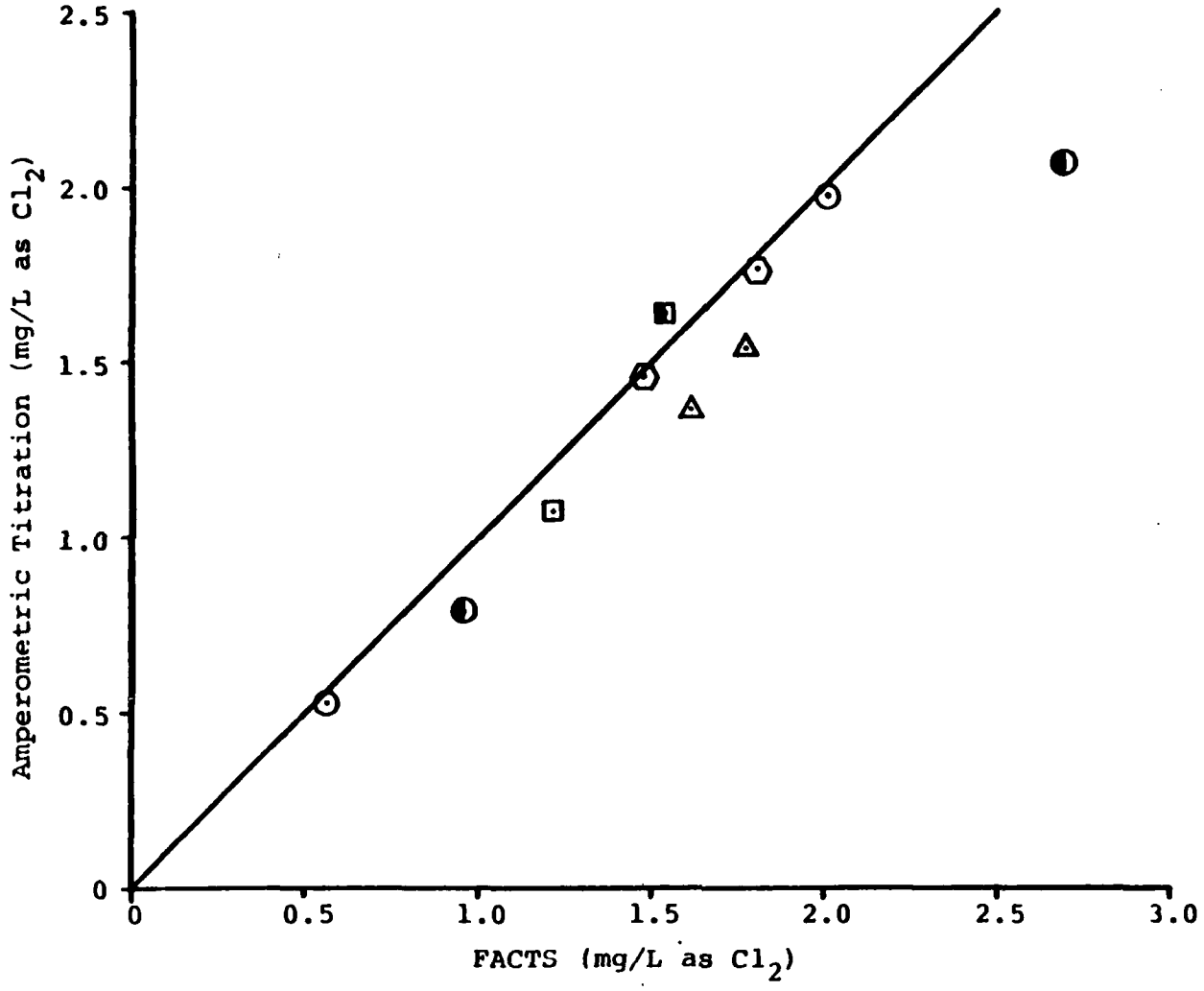


Figure 2. Comparison of Means for FACTS and Amperometric Titration for Free Available Chlorine Obtained at Six Water Treatment Plants.
Reprinted from JOURNAL AWWA, Vol. 75, No.12 (December 1983), by Permission. Copyright© 1983, The American Water Works Association.

TABLE OF CONTENTS

INTRODUCTION

Chlorine is used in water and wastewater treatment for various purposes. Of these uses, the most common are taste and odor control, color removal, and disinfection in drinking water, and disinfection in wastewater treatment.

There are numerous methods for determining chlorine residuals in aqueous solutions (1,2). Briefly, these include the iodometric titration, amperometric titration, and several colorimetric procedures. The iodometric titration is limited to total residuals above 1 mg/L as Cl_2. The amperometric titration is capable of differentiating free chlorine ($HOCl/OCl^-$), monochloramine (NH_2Cl), and dichloramine ($NHCl_2$) and is generally the method of choice in the laboratory. Field measurements are limited by the complexity of the instrumentation. The colorimetric determinations, which find application in both the laboratory and the field are the DPD (N,N-diethyl-p-phenylenediamine), LCV (leuco crystal violet) and FACTS (free available chlorine test with syringaldazine) procedures.

More recently an electrode method has been published which was reported to be selective for HOCl (3).

All test procedures used for National Interim Primary Drinking Water Regulation (NIPDWR) compliance monitoring must be approved by the U.S. Environmental Protection Agency (4). Collaborative testing of each procedure is required for inclusion in Annual Book of Standards, Part 31, Water, of American Society for Testing and Materials (ASTM) (2), and Standard Methods for the Examination of Water and Wastewater (Standard Methods) (1).

The U.S. Army Medical Bioengineering Research and Development Laboratory has been involved in the evaluation of chlorine residual test procedures and the development of a colorimetric test for chlorine residuals, FACTS, free available chlorine test with syringaldazine. It has been shown that the FACTS test procedure is more specific for free chlorine in the presence of common interferences, NH_2Cl, $NHCl_2$, Mn^{+4} and Fe^{+3} than is DPD (5-15). However, prior to this research the FACTS procedure had not been approved by EPA for use in drinking water and had never been subjected to an extensive collaborative comparison.

This report presents the results of the comparison of the FACTS test procedure and the approved standard test, DPD. Data are also included showing the comparison of the FACTS test procedure with amperometric titration.

OBJECTIVES

1. To compare several instrumental methods for determining chlorine residuals in drinking water.

2. To develop a detailed protocol for equivalency testing of methods for drinking water.

3. To obtain equivalency test data from a minimum of six water treatment plants using the FACTS test kit procedure and approved methods.

METHODS AND MATERIALS

The instrumental methods that were compared are described in detail in Appendix A. A detailed discription of the methods and material used in this study are provided in the Appendices. The main objective of study was to compare the FACTS test procedure, the FACTS® test kit was the HI FACTS Test Kit for Measuring Free Available Chlorine, obtained from Ames Division, Miles Laboratories, Inc., P.O. Box 70, Elkhart, Indiana 46515. The DPD or amperometric procedure normally used at each plant participating in this comparison served as the standard test procedure.

RESULTS AND DISCUSSION

As a portion of our overall evaluation of chlorine test procedures, several instrumental methods for determining chlorine residuals in drinking water were evaluated. The results of this experimentation have been published and are detailed in Appendix A (16).

A detailed research protocol was developed, which when followed provides sufficient data for the comparison of analytical procedures for determining free chlorine in aqueous samples. This protocol has been published and is reproduce in Appendix B (17).

The experimental design (17) was tested using the FACTS, DPD and amperometric test procedures. Ten water treatment laboratories participated in the comparison of the FACTS and DPD, and six laboratories participated in the comparison of the FACTS and the amperometric titration. The results have been published, Appendix C (18), and the FACTS test has been approved as an alternate test procedure for determining free chlorine, Appendix D.

A study has recently been reported that details the kinetics of monochloramine oxidation of DPD (19) and confirms the previous studies (5-13). From the kinetic expressions it was shown that the monochloramine interference with the DPD test is dependent in pH, monochloramine and DPD concentration. An average of 5.8 percent per minute interference was calculated for the DPD colorimetric test as described in Standard Methods (1).

CONCLUSIONS

The objectives of this project were achieved. A detailed protocol was developed that can be used for comparison testing of new water quality test procedures. This test plan was designed to meet the minimum requirements of the USEPA for equivalency testing of new methods for compliance monitoring.

Subsequent to the protocol development, an extensive test of the FACTS test was conducted at 16 water treatment plants. From the results of these tests it was shown that the FACTS test was equivalent to the DPD test and the amperometric titration procedure. The test also demonstrated that the FACTS test kit procedure had precision equal to the DPD test procedure.

In conclusion the FACTS test has been shown to be equivalent to the DPD test. The FACTS test procedure is more specific for free chlorine that the DPD test presently authorized. Therefore, when chlorine is the only treatment of water prior to consumption, e.g. field Army operations, the FACTS test for free chlorine is superior to the DPD test procedure.

RECOMMENDATIONS

1. An evaluation of the present Army standard procedure for determining available chlorine for the field Army should be initiated.

2. The FACTS test procedure should be considered as a replacement for the Army standard procedure for free available chlorine for field Army.

REFERENCES

1. APHA, AWWA, WPCF. 1980. Standard Methods for the Examination of Water and Wastewater. 15th ed., pp 309-349. American Public Health Association, 1015 Eighteenth Street, NW, Washington, DC 20036.

2. American Society for Testing and Materials. 1980. Annual Book of ASTM Standards, Part 31, Water. ASTM. 1916 Race Street, Philadelphia, PA 19103, pp 366-377.

3. Johnson, J. D., J. W. Edwards, and F., Keeslar. 1978. Chlorine residual measurement cell: The HOCl membrane electrode. J. Am. Water Works Assoc. 70:341-348.

4. U.S. Environmental Protection Agency. 1976. Guidelines establishing test procedures for the analysis of pollutants. Federal Register 41(232):52780-52786, (Dec. 1).

5. Guter, K. J. and W. J. Cooper. 1973. The Evaluation of Existing Field Test Kits for Determining Free Chlorine Residuals in Aqueous Solutions. USAMEERU Report 73-03, AD 752440. U.S. Army Medical Environmental Engineering Research Unit, Edgewood Arsenal, MD.

6. Guter, K. J., W. J. Cooper, and C. A. Sorber. 1973. The Evaluation of Existing Field Test Kits for Determining Free Chlorine Residuals in Aqueous Solutions. Presented at the First Water Quality Technology Conference, Sponsored by the American Water Works Assoc., Dec. 3 and 4, Cincinnati, OH.

7. Guter, K. J., W. J. Cooper and C. A. Sorber. 1974. Evaluation of existing field test kits for determining free chlorine residuals in aqueous solutions. J. Am. Water Works Assoc. 66:38-43.

8. Cooper, W. J., E. P. Meier, J. W. Highfill, and C. A. Sorber. 1974. The Evaluation of Existing Field Test Kits for Determining Free Chlorine Residuals in Aqueous Solutions, Final Report. Technical Report 7402, AD 780051. U.S. Army Medical Bioengineering Research and Development Laboratory, Fort Detrick, Frederick, MD.

9. Meier, E. P., W. J. Cooper, and C. A. Sorber. 1974. Development of a Rapid Specific Free Available Chlorine Test with Syringaldazine (FACTS). Technical Report 7405, AD 780857. U.S. Army Medical Bioengineering Research and Development Laboratory, Fort Detrick, Frederick, MD.

10. Cooper, W. J., C. A. Sorber, and E. P. Meier. 1974. A Rapid Specific Free Available Chlorine Test with Syringaldazine (FACTS). Presented at the 94th Annual American Water Works Association Conference. 1974. June 16-21, Boston MA.

11. Cooper, W. J., C. A. Sorber, and E. P. Meier. 1975. A rapid specific free available chlorine test with wyringaldazine (FACTS). J. Am. Water Works Assoc. 67:34-39.

12. Sorber, C.A., W.J. Cooper, and E.P. Meier. 1977. Selection of a field method for free available chlorine. In J.D. Johnson, ed. Disinfection Water and Wastewater. Ann Arbor Science Publ., Inc. pp. 91-112. Ann Arbor, MI.

13. Meier, E. P., W. J. Cooper, and J. W. Highfill. 1978. Evaluation of the Specificity of the DPD-Glycine and FACTS Test Procedures for Determining Free Available Chlorine. Presented at the Division of Environmental Chemistry, American Chemical Society, Anaheim, CA. March 13-17.

14 Free available chlorine test with syringaldazine (FACTS) (tentative). In APHA. Standard Methods for the Examination of Water and Wastewater, 15th ed., para. 114, Chlorine (residual).

15. Liebermann, J., Jr., N. M. Roscher, E. P. Meier, and W. J. Cooper. 1980. Development of the FACTS procedure for combined forms of chlorine and ozone in aqueous solutions. Environ. Sci. Technol. 14:1395-1400

16. Cooper, W.J., M.F. Mehran, R.A. Slifker, D.A. Smith, J.T. Villate, and P.H. Gibbs. 1982. Comparison of several instrumental methods for determining chlorine residuals in drinking water. J. Am. Water Works Assoc. 74:546-552.

17. Gibbs, P.H., W.J. Cooper, and E.M. Ott. 1983. A statistical experimental design for comparison testing of analytical procedures. J. Am. Water Works Assoc. 75:578-582.

18. Cooper, W.J., P.H. Gibbs, E.M. Ott, and P. Patel. 1983. Equivalency testing of procedures for measuring free available chlorine: amperometric titration, DPD, and FACTS. J. Am. Water Works Assoc. 75:625-629.

19. Moore, H.E., M.J. Garmendia and W.J. Cooper. Kinetics of Monochlorimine Oxidation of N,N-Diethyl-p-phenylenediamine. Environ. Sci. Technol. 18:348-353.

APPENDIX A*

William J. Cooper
M.F. Mehran
R.A. Slifker
D.A. Smith
J.T. Villate
P.H. Gibbs

COMPARISON OF SEVERAL INSTRUMENTAL METHODS FOR DETERMINING CHLORINE RESIDUALS IN DRINKING WATER

Reprinted from JOURNAL AWWA, Vol. 74, No. 10
(October 1982), by permission.
Copyright© 1982, The American Water Works Association

* Note: This Appendix A includes Tables 1 - 12.

Comparison of several instrumental methods for determining chlorine residuals in drinking water

William J. Cooper, M.F. Mehran, R.A. Slifker, D.A. Smith, J.T. Villate, and P.H. Gibbs

The authors evaluated four methods for determining chlorine residuals in water. Two membrane electrodes, a potentiometric electrode, and a continuous total chlorine analyzer were used by analysts to measure free or total chlorine in quality assurance samples and in tap water samples. Amperometric titration was used as a referee method. The statistical analyses of the evaluations are presented in this paper. Average precision of all analysts and of individual analysts is considered for each method.

The realization that the use of chlorine can produce halogenated organic compounds in drinking water has emphasis on research regarding the chlorination process.[1-4] This increased emphasis has also resulted in additional studies of the analytical determination of active chlorine in aqueous solutions. Numerous methods for the determination of hypochlorous acid, hypochlorite ion, monochloramine, dichloramine, and nitrogen trichloride (HOCl, OCl^-, NH_2Cl, $NHCl_2$, and NCl_3) have been presented in *Standard Methods*[5] and evaluated elsewhere.[6-9] New methods utilizing membrane electrodes, which are reported to differentiate between HOCl and OCl^-,[10-12] have been developed. Potentiometric electrodes[13] and an automated amperometric method[14,15] also may be used in the analysis of total residual chlorine. The objective of this study was to evaluate two membrane electrodes, a potentiometric electrode, and a total chlorine analyzer for determining chlorine residuals in water.

Experimental procedures

Membrane electrodes. Two membrane amperometric electrodes, obtained from different manufacturers, were tested. The manufacturers' operating instructions were followed. When not being used, the electrodes were capped to protect the membrane. The meters were connected to a 110-V power source, with the electrode attached to the meter at all times to maintain stability. Calibration was checked daily by preparing a solution of acetate buffered (pH 4) HOCl. (The chlorine solutions were prepared by diluting reagent grade sodium hypochlorite. The buffer solution was prepared by diluting 146 g sodium acetate and 480 g acetic acid to 1 L.) The concentration of HOCl was determined by amperometric titration.

The specific operating directions for the membrane electrodes were as follows:

- Membrane electrode A*

1. Determine the pH of the sample.
2. Place 150 mL of the sample in a 180-mL, tall form beaker, and place the beaker on a magnetic stirrer. (Shake the electrode to dislodge air bubbles when placing it in the test solution.)
3. Lower the electrode into the sample until it almost touches the bottom of the beaker, and turn on the stirrer (a stirring bar is built into the electrode).
4. Take a reading on the appropriate scale—0-1, 0-5, or 0-10 mg Cl_2/L.
5. Calculate the actual free chlorine by using the following equation:

$$K_1 = \frac{[H^+]\,[OCl^-]}{[HOCl]} \qquad (1)$$

for which $K_1 = 2.0 \times 10^{-8}$ at 0°C or $K_1 = 3.3 \times 10^{-8}$ at 20°C. For example, for a temperature of 20°C and a pH of 7.0, then $K_1 = 3.3 \times 10^{-8}$ and $[H^+] = 1 \times 10^{-7}$ M. If [HOCl] = 1.0 mg/L, then by substitution into and rearrangement of Eq 1, one can determine the value of $[OCl^-]$

$$[OCl^-] = \frac{(3.3 \times 10^{-8})\,(1.0)}{1 \times 10^{-7}}$$

$$= 0.33 \text{ mg/L}$$

Therefore, the free available chlorine is $[HOCl] + [OCl^-] = 1.0 \text{ mg/L} + 0.33 \text{ mg/L} = 1.33$ mg/L.

- Membrane electrode B†

1. Determine the pH of the sample.
2. Place 400 mL of the sample in a 600-mL beaker.
3. Place the beaker on a magnetic stirrer, and insert the electrode into the sample, making sure there are no bubbles on the membrane.
4. Take a reading on the appropriate scale—0-1 or 0-5 mg Cl_2/L.
5. Calculate the actual free chlorine by using Eq 1.

Potentiometric electrode. The potentiometric electrode‡ can measure only total chlorine. The electrode was calibrated daily as follows:

1. Pipet 0.20, 1.00, 2.00, and 2.50 mL of 0.00281*N* iodate standard into 100-mL volumetric flasks. (The standard was prepared by dissolving 1.005 g of potassium iodate and diluting it to 1 L.)
2. Add 5 mL of an iodide solution (10 percent potassium iodide [KI]) and 1 mL of acetate buffer (pH 4) to each flask and to a flask containing no iodate (blank). Swirl the flasks to mix the solutions, and let the solutions stand for 5 ± 0.5 min.
3. Dilute each standard to the 100-mL mark, then mix and pour the solution into a 150-mL Erlenmeyer flask. These solu-

*Orion Research Inc, Cambridge, Mass.
†Delta Scientific Envirotech, Hauppauge, N.Y.
‡Orion Research Inc, Cambridge, Mass.

0003-150X/82/100546-07$02.00

© 1982 American Water Works Association

TABLE 1

Descriptive statistics of the test procedures used in analyzing the total chlorine content of quality assurance samples prepared with chlorine-demand-free water

Sample	Amperometric Titrator				Potentiometric Electrode				Total Chlorine Analyzer			
	$\bar{x}$*	s†	RSD‡	n§	$\bar{x}$	s	RSD	n	$\bar{x}$	s	RSD	n
1	0.40	0.013	3.3	12	0.38	0.060	15.8	12	0.42	0.0030	0.7	12
2	0.92	0.022	2.4	12	0.93	0.095	10.2	12	0.94	0.049	0.5	12
3	1.28	0.023	1.8	12	1.28	0.16	12.7	12	1.30	0.013	1.0	12
4	1.29	0.019	1.5	12	1.21	0.14	11.6	12	1.35	0.0076	0.6	12

*The mean
†The standard deviation
‡The relative standard deviation
§The number of observations

TABLE 2

*Summary of the calculated t-statistic, adjusted degrees of freedom, and value of the 95 percent level t-distribution used for comparing data from Table 1**

Sample	Amperometric Titrator Versus Potentiometric Electrode			Amperometric Titrator Versus Total Chlorine Analyzer			Potentiometric Electrode Versus Total Chlorine Analyzer		
	ADF†	t Calculated	t Comparison	ADF	t Calculated	t Comparison	ADF	t Calculated	t Comparison
1	N.D.‡	1.13	N.C.§	12	5.19	2.78	N.D.	2.31	N.C.
2	N.D.	0.36	N.C.	12	3.07	2.78	N.D.	0.36	N.C.
3	N.D.		N.C.	17	5.24	2.66	N.D.	0.88	N.C.
4	N.D.	1.96	N.C.	16	10.16	2.74	12	3.45	2.78

*Statistical significance is indicated when the calculated t-statistic is greater than the value obtained from the t-distribution at the appropriate degrees of freedom (comparison)
†Adjusted degrees of freedom
‡Not determined.
§Not calculated, for a comparison of three test procedures, any calculated t-statistic less than 2.60 at 22 degrees of freedom is not significant, and therefore no calculation was necessary.

TABLE 3

Summary of single analyst statistics and relative standard deviation for test procedures used in analyzing quality assurance samples prepared with chlorine-demand-free water

Sample	Analyst	Amperometric Titrator			Potentiometric Electrode			Total Chlorine Analyzer		
		$\bar{x}$*	s†	RSD‡	$\bar{x}$	s	RSD	$\bar{x}$	s	RSD
1	1	0.40	0.012	3.1	0.42	0	0	0.42	0.0015	0.4
	2	0.40	0.016	3.9	0.42	0	0	0.42	0.0012	0.3
	3	0.40	0.021	5.3	0.41	0	0	0.42	0.0006	0.1
	4	0.40	0.0035	0.9	0.28	0	0	0.42	0.0015	0.4
2	1	0.89	0.012	1.3	1.00	0	0	0.93	0.004	0.4
	2	0.92	0.020	2.2	0.99	0	0	0.94	0.002	0.2
	3	0.92	0.009	1.0	0.97	0	0	0.94	0.003	0.3
	4	0.94	0.005	0.5	0.78	0.035	4.4	0.93	0.002	0.2
3	1	1.24	0.038	3.0	1.07	0.052	4.9	1.30	0.004	0.3
	2	1.27	0.010	0.8	1.25	0.159	12.7	1.32	0.008	0.6
	3	1.28	0.021	1.7	1.30	0	0	1.31	0.003	0.2
	4	1.25	0.012	0.9	1.43	0.081	5.7	1.29	0.001	0.2
4	1	1.28	0.008	0.6	1.35	0.092	6.8	1.34	0.004	0.3
	2	1.29	0.012	0.9	1.11	0	0	1.36	0	0
	3	1.28	0.015	1.1	1.32	0.029	2.2	1.35	0.0006	0.04
	4	1.30	0.035	2.7	1.08	0.064	6.0	1.35	0.0006	0.04
Average RSD				1.9			2.7			0.2

*The mean
†The standard deviation
‡The relative standard deviation

TABLE 4

Descriptive statistics of test procedures used in analyzing the free chlorine content of tap water samples

Sample	Amperometric Titrator				Amperometric Membrane Electrode A				Amperometric Membrane Electrode B			
	$\bar{x}$*	s†	RSD‡	n§	$\bar{x}$	s	RSD	n	$\bar{x}$	s	RSD	n
5	1.05	0.038	3.6	11	0.37	0.045	12.2	12	0.23	0.10	43.5	12
6	1.09	0.028	2.4	12	0.29	0.034	11.7	12	0.24	0.078	32.5	12
7	0.80	0.032	4.0	12	0.16	0.015	9.4	12	0.17	0.037	21.8	12
8	0.87	0.019	2.2	12	0.14	0.025	17.9	12	0.17	0.046	28.2	12

*The mean
†The standard deviation
‡The relative standard deviation
§The number of observations

tions are the equivalent of 0.20, 1.00, 2.00, and 2.50 mg Cl_2/L.

4. Stir each solution at 30–60 rpm with a magnetic stirrer, immerse the electrode into the solution, and read the millivolts (positive) after 30 s.

5. To obtain a calibration curve plot millivolts versus log chlorine concentration (mg Cl_2/L).

To determine the total chlorine content of the samples, the procedure was as follows:

1. Add 5 mL of an iodide solution (10 percent KI) and 1 mL of the acetate buffer (pH 4) to a 100-mL volumetric flask.

2. Fill the volumetric flask to the mark with the sample.

3. Stopper the flask, and mix the sample and reagents by inverting the flask five times. Let the solution stand for 5 ± 0.5 min.

4. Pour the solution into an Erlenmeyer flask, and stir at 30–60 rpm with a magnetic stirrer. Immerse the electrode into the solution, and read the millivolts (positive) after 30 s.

5. Read the concentration from the calibration curve.

Continuous total chlorine analyzer.* The manufacturer's operating procedures were followed. The instrument was zeroed and then calibrated by depressing the appropriate range button and adjusting a multiturn potentiometer to give the appropriate value.

Samples were poured into 400-mL beakers. After the instrument was calibrated, the samples were analyzed for their total chlorine content. Readings were taken from the digital display of the analyzer.

Experimental design

Two different types of water samples—quality assurance samples and tap water samples—were used to test the procedures. Quality assurance samples were obtained from the US Environmental Protection Agency. Four of these samples (8 mL diluted to 1 L with chlorine-demand-free water) were tested. Tap water samples were obtained after allowing the water to run 5–10 min. Four tap water samples were tested.

For each sample, there were four analysts and four test procedures. Analysts were randomly assigned test procedures, and samples were presented blind. Each sample was broken down into four sets, one for each analyst-procedure combination. For each set tested, the analyst was instructed to run triplicates. Therefore, the total number of observations for any one sample was 48 (4 × 4 × 3 = 48).

The referee method was amperometric titration. For this method, an automatic titrator* was used. For each sample set, triplicate analyses were performed.

*[illegible]

Results and discussion

Statistical analysis. A total of eight samples was used. Each sample was analyzed in a similar manner. For each test procedure-sample combination, there were 12 measurements, i.e., four analysts reading triplicates. Only one analyst used the referee procedure—amperometric titration. For all of the procedures, descriptive statistics—mean, $\bar{x}$, and standard deviation, s—were obtained. For the statistical tests, independence of all measurements was assumed. (If the three measurements obtained by each analyst for each sample cannot be assumed to be independently obtained values, then the triplicates should be averaged and the average value should be used in the statistical analysis.) The sample means were compared quantitatively by computing the *t*-statistic:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \qquad (2)$$

where $\bar{x}_1$ and $\bar{x}_2$ represent the means of the samples of interest, s_1^2 and s_2^2 were the variances of the respective samples, and n_1 and n_2 were the number of observations, i.e., measurements, for each sample. Homogeneous variances between samples were assumed for the computation and comparison of the *t*-statistic for any given level of significance and any number of degrees of freedom. The results from these experiments indicated that there was considerable variability among variances of the procedures, i.e., heterogeneity. Therefore, the number of the adjusted degrees of freedom (ADF) for the *t*-statistic was computed for each comparison of interest by using the Satterthwaite approximation:

$$\text{ADF} = \frac{u^2}{\frac{v^2}{n_1 - 1} + \frac{w^2}{n_2 - 1}} \qquad (3)$$

where $v = s_1^2/n_1$, $w = s_2^2/n_2$, $u = v + w$, and n = the number of observations.[16]

The Bonferroni procedure was used to adjust for multiple comparisons.[17] This procedure determines the level of significance for each pairwise comparison by dividing the overall α level by the number of comparisons made on each sample. Therefore, for comparisons of three sample means at the overall 95 percent confidence level, the *t*-statistic chosen for each pairwise comparison is that of α = 0.017 and the appropriate number of adjusted degrees of freedom. Several examples of the statistical calculations are included in the appendix.

Quality assurance. *Precision of all analysts.* A summary of the data obtained for the four quality assurance samples is provided in Table 1. (These samples were unbuffered, and the membrane-covered amperometric electrodes did not function properly. These samples should be buffered to obtain standard data.) Comparisons of the results from the amperometric titrator with those from the potentiometric electrode showed that no statistical difference existed between the means of the samples at the 95 percent level of confidence (Table 2). The precision of all analysts who used the potentiometric electrode was significantly less (greater spread) than the precision of the one analyst who used the amperometric titrator. This was not unexpected because an individual analyst's precision is generally better than the precision of a group of analysts.

The mean of the results obtained with the amperometric titrator was lower than that obtained with the total chlorine analyzer. This difference was statistically significant at the 95 percent level of confidence for all samples (Table 2). However, the means were never different by more than 0.06 mg Cl_2/L. The precision of all analysts who used the total chlorine analyzer was better than the precision of one analyst who used the amperometric titrator. Because the precision obtained with both test procedures was good, statistical differences were observed with very small absolute differences in the two means, i.e., samples 1, 2, and 3. Even in sample 4 (see Table 1) there was only a difference of 0.06 mg Cl_2 L in the means. Once detected, this difference was examined to determine whether it was acceptable. In most water treatment plants, a difference of 0.06 mg Cl_2 L will not be significant, and therefore the difference is acceptable.

A comparison of the results obtained with the potentiometric electrode and those obtained with the total chlorine analyzer showed no statistically significant differences at the 95 percent level of confidence in samples 1, 2, and 3 (Table 2). Sample 4 appeared to be statistically different with $\bar{x}$ = 1.21 for the potentiometric electrode and $\bar{x}$ = 1.35 for the total chlorine analyzer.

Precision of individual analysts. Each sample was divided into four subsamples. Each analyst received a subsample for use in testing each procedure. Triplicate analyses were performed, and the precision of an individual analyst was determined by using the relative standard deviation (the standard deviation was divided by the mean and then multiplied by 100). The data are summarized in Table 3.

It appears from the data that an individual analyst's precision with the potentiometric electrode was best when measuring low chlorine levels (≤1.0 mg Cl_2/L). In comparing the means of samples 1 and 2 obtained by each analyst, it appears that in both cases, analyst 4 obtained low results. If the results obtained for samples 1 and 2 by analyst 4 were eliminated, then the overall precision of the procedure would be increased. The overall means of samples 1 and 2 were very similar to those obtained by the referee method (Table 1).

Over the entire range of chlorine concentrations tested, no differences in precision were evident from the results obtained with the total chlorine analyzer. These results appeared to be independent of the analyst who performed the procedure.

Tap water samples. *Precision of all analysts.* Four samples of tap water were used. Each analyst obtained triplicate results for each of the four procedures. Because the tap water contained both free and combined chlorine, different comparisons were obtained. The comparisons of the free chlorine reading obtained by using the amperometric titrator with those from the two membrane electrodes are shown in Table 4.

The results indicate good agreement between the means of the values obtained from the two membrane electrodes; no statistical difference was detected at the 95 percent level of confidence except for sample 5 (Table 5). In every case, the relative precision was better with membrane electrode A than with membrane electrode B. This may be related to the small size of meter B on which readings were taken, resulting in less accurate interpolation between numbers. Surprisingly, these results obtained from the membrane electrodes are low compared with those obtained from the amperometric titrator. This point will be discussed further under the discussion of active chlorine speciation.

Table 6 summarizes the data for the measurements of total chlorine in the tap water samples. When the results from the amperometric titrator were compared with those from the potentiometric electrode, there was no statistical difference for sample 6 at the 95 percent level of confidence (Table 7). In every sample, except sample 7, the relative precision was less with the amperometric titrator than that with the electrode. Again, it should be pointed out that only one analyst used the amperometric titrator, whereas four analysts used the potentiometric electrode.

A statistical difference at the 95 percent level of confidence was obtained for all samples except sample 8 when measurements from the amperometric titrator were compared with those from the total chlorine analyzer. Statistically different means were obtained with these two methods. This was also the case with the quality assurance samples. The spread was somewhat larger between the means (except for sample 8), but in no case was the difference greater than the 0.08 mg Cl_2/L, which was observed in sample 7. This difference represents an absolute difference of 7.2 percent from the mean concentration determined by amperometric titration. The differences in samples 5 and 6 were 4.5 percent and 3.6 percent.

TABLE 5

Summary of the calculated t-statistic, adjusted degrees of freedom, and value of the 95 percent level t-distribution used for comparison of data from Table 4

	Amperometric Membrane Electrode A Versus Amperometric Membrane Electrode B		
	Adjusted Degrees of	t	
Sample	Freedom	Calculated	Comparison
5	15	4.42	2.13
6	N.D.*	2.04	N.D.
7	N.D.	0.87	N.D.
8	N.D.	1.92	N.D.

*Not determined; for a comparison of two test procedures, any calculated t-statistic less than 2.07 at 22 degrees of freedom is not significant, and therefore, no calculation was necessary.

TABLE 6

Descriptive statistics of the test procedures for determining the total chlorine content of tap water samples

	Amperometric Titrator				Potentiometric Electrode				Total Chlorine Analyzer			
Sample	$\bar{x}$*	s†	RSD‡	n§	$\bar{x}$	s	RSD	n	$\bar{x}$	s	RSD	n
5	1.33	0.036	2.8	10	1.13	0.072	6.4	12	1.27	0.022	1.7	12
6	1.37	0.016	1.2	12	1.28	0.17	13.3	12	1.32	0.012	0.9	12
7	1.11	0.062	5.6	12	1.03	0.043	4.2	12	1.03	0.012	1.2	12
8	1.16	0.026	2.4	12	1.08	0.075	7.1	12	1.15	0.023	2.0	12

*The mean
†The standard deviation
‡The relative standard deviation
§The number of observations

TABLE 7

Summary of the calculated t-statistic, adjusted degrees of freedom, and value of the 95 percent level t-distribution used for comparison of data from Table 6

	Amperometric Titrator Versus Potentiometric Electrode			Amperometric Titrator Versus Total Chlorine Analyzer			Potentiometric Electrode Versus Total Chlorine Analyzer		
		t			t			t	
Sample	ADF*	Calculated	Comparison	ADF*	Calculated	Comparison	ADF*	Calculated	Comparison
5	N.D.†	6.33	N.D.	14	4.42	2.74	N.D.	6.44	N.D.
6	N.D.	1.83	N.D.	20	5.66	2.61	N.D.	0.81	N.D.
7	N.D.	3.87	N.D.	12	4.39	2.78	N.D.		N.D.
8	N.D.	4.33	N.D.	21	0.98	N.D.	N.D.	3.97	N.D.

*Adjusted degrees of freedom
†Not determined; in many cases it was necessary to calculate the adjusted degrees of freedom. The t-distribution value (the least possible in this experiment) at 12 degrees of freedom comparing three tests is 2.78. Therefore, any value greater than 2.78 results in a statistically significant difference at the 95 percent confidence level. On the other hand, any value less than 2.60 results in no statistical difference at the 95 percent confidence level at 22 degrees of freedom, the maximum possible during this experiment.

TABLE 8

Summary of single analyst statistics and relative standard deviation for test procedures used in analyzing the free available chlorine content of tap water samples

		Amperometric Titrator			Amperometric Membrane Electrode A			Amperometric Membrane Electrode B		
Sample	Analyst	$\bar{x}$*	s*	RSD‡	$\bar{x}$	s	RSD	$\bar{x}$	s	RSD
5	1	1.00	0.041	4.1	0.33	0.012	3.6	0.11	0.006	5.5
	2	1.08	0.006	0.5	0.35	0.059	16.7	0.17	0.006	3.7
	3	1.07	0.020	1.9	0.41	0	0	0.10	0.017	5.4
	4	1.08	0.032	2.9	0.40	0.033	8.3	0.33	0.068	20.5
6	1	1.12	0.020	1.8	0.25	0.021	8.2	0.13	0.004	3.5
	2	1.08	0.006	0.5	0.27	0.007	2.4	0.33	0.013	1.8
	3	1.10	0.015	1.4	0.33	0.013	3.8	0.21	0.005	2.5
	4	1.06	0.012	1.1	0.29	0.013	4.4	0.28	0.022	7.7
7	1	0.80	0.011	1.4	0.15	0.013	9.0	0.13	0.014	0.9
	2	0.82	0.022	2.7	0.17	0.011	6.3	0.14	0.004	2.9
	3	0.81	0.045	5.5	0.18	0.018	10.9	0.19	0.021	10.9
	4	0.78	0.009	1.1	0.16	0.020	12.6	0.21	0.027	12.9
8	1	0.89	0.010	1.2	0.15	0.005	3.1	0.16	0.011	6.6
	2	0.86	0.021	2.3	0.11	0.013	12.1	0.13	0	0
	3	0.88	0.007	0.8	0.12	0.009	7.2	0.14	0.009	6.2
	4	0.85	0.011	1.0	0.17	0.021	12.2	0.25	0.017	6.9
Average RSD				1.9			7.6			6.9

*The mean
†The standard deviation
‡The relative standard deviation

respectively. The differences were minimal and would not be significant to the operation of a water treatment plant.

The comparison of the potentiometric electrode with the total chlorine analyzer indicated a statistical difference in samples 5 and 8 (at the 95 percent level of confidence). The difference in the means was 0.14 mg Cl_2/L in sample 5 and 0.09 mg Cl_2/L in sample 8. No statistical difference was observed in samples 6 and 7.

Precision of individual analysts. Each sample was divided into four subsamples. Each analyst received a subsample for use in testing each of the procedures. The tap water contained both a free and a combined residual; therefore, the data were divided into free chlorine and total chlorine measurements. The precision of individual analysts for the free available chlorine determinations are presented in Table 8. An individual analyst's precision obtained with either membrane electrode was considerably less than the overall precision (Table 4). It was also apparent from a comparison of the means for any one sample that considerable variability existed. This variability accounted for the large, overall relative standard deviation.

As a prototype instrument, membrane electrode A and its associated electronics may require additional engineering to reduce the variability in results from analyst to analyst. Also, the electrical noise detected by the electronics could have caused some of the fluctuations in measurements. Stirring of the sample is critical when using membrane electrodes, and slight variations in stirring speed (the velocity of fluid across the membrane) can cause large variations in the

TABLE 9

Summary of single analyst statistics and relative standard deviation for test procedures used in analyzing the total chlorine of tap water samples

Sample	Analyst	Amperometric Titrator			Potentiometric Electrode			Total Chlorine Analyzer		
		$\bar{x}$*	s†	RSD‡	$\bar{x}$	s	RSD	$\bar{x}$	s	RSD
5	1	1.296	0.008	0.7	1.19	0	0	1.26	0.00§	0.5
	2	1.35	0.008	0.4	1.20	0.040	3.4	1.23	0.008	0.6
	3	1.34	0.017§	1.3	1.05	0	0	1.28	0.001	0.1
	4	1.348	0.021	1.6	1.13	0.046	4.1	1.29	0.006	0.4
6	1	1.38	0	0	1.08	0.040	3.7	1.33	0.009	0.6
	2	1.37	0.010	0.7	1.21	0.095	7.9	1.30	0.005	0.3
	3	1.37	0.011	0.8	1.30	0	0	1.32	0.005	0.3
	4	1.35	0.015	1.1	1.52	0	0	1.33	0.003	0.1
4	1	1.10	0.003	0.3	1.02	0.035	3.4	1.01	0.00§	0.0
	2	1.11	0.015	1.4	1.03	0.081	7.8	1.04	0.00§	0.0
	3	1.16	0.114	9.8	1.03	0	0	1.03	0.001	0.1
	4	1.06	0.002	0.2	1.04	0.046	4.4	1.03	0.008	0.8
4	1	1.16	0.057	4.9	1.05	0.040	3.8	1.13	0.011	1.0
	2	1.17	0.016	1.4	1.00	0	0	1.12	0.010	0.9
	3	1.17	0.008	0.5	1.18	0	0	1.18	0.005	0.4
	4	1.15	0.013	1.1	1.01	0	0	1.16	0.00§	0.6
Average RSD				2.0			2.4			0.5

*The mean
†The standard deviation
‡The relative standard deviation
§Only two observations were made

TABLE 10

Summary of analyst precision with test procedures for analyzing quality assurance samples prepared with chlorine-demand-free water

	Amperometric Titrator*	Potentiometric Electrode	Total Chlorine Analyzer
Average single analyst precision	1.9†	2.7	0.7
Average overall analyst precision	2.3	12.6	0.7

*Only one analyst used the amperometric titrator, whereas four analysts used each of the other procedures
†Relative standard deviation

TABLE 11

Summary of analyst precision with test procedures for analyzing the free available chlorine of tap water samples

	Amperometric Titrator*	Amperometric Membrane Electrode A	Amperometric Membrane Electrode B
Average single analyst precision	1.9†	7.8	6.9
Average overall analyst precision	3.1	12.8	31.5

*Only one analyst used the amperometric titrator, whereas four analysts used each of the other procedures
†Relative standard deviation

TABLE 12

Summary of analyst precision with test procedures for analyzing the total chlorine of tap water samples

	Amperometric Titrator*	Potentiometric Electrode	Total Chlorine Analyzer
Average single analyst precision	2.0†	2.4	0.5
Average overall analyst precision	3.0	7.8	1.5

*Only one analyst used the amperometric titrator, whereas four analysts used each of the other procedures
†Relative standard deviation

measurements. Membrane electrode A was calibrated daily.

Membrane electrode B was extremely stable throughout the evaluation. The calibration, as adjusted at the factory, was checked daily but did not require adjustment over a two-month period. The single largest factor that may account for the analyst-to-analyst variation was the size of the meter. Because the meter is small, large errors in interpolation of numbers on the meter are possible.

Table 9 summarizes the precision of individual analysts in measuring the total chlorine content of tap water. Very little difference was observed between the precision obtained with the amperometric titrator and that obtained with the potentiometric electrode.

For the total chlorine analyzer, a relative standard deviation of 0.5 percent was calculated for the average precision of individual analysts. This result was 1.5 percent less than that for the amperometric titrator. Again, as was observed from the analyses of the quality assurance samples, the results obtained with the total chlorine analyzer appear almost independent of the analyst who performed the procedure.

Summary of precision. Summaries of the average precision of individual analysts and all analysts are presented in Tables 10, 11, and 12. The average precision of individual analysts was always better than the average precision of all analysts. The data for the amperometric titrator were obtained by one analyst and were similar for quality assurance samples and for tap water samples.

The average precision of individual analysts with the potentiometric electrode is comparable to that with the amperometric titrator. However, the average precision of all analysts with the potentiometric electrode was not as good as that with the amperometric titrator. This difference in precision may have been due to the nondigital millivolt meter that was used. (When millivolts are read on an armature meter, the third place must be estimated, and this can lead to significant errors.) Moreover, the standard curve was obtained by using only one millivolt reading for each concentration. Precision probably would have been better if duplicate or triplicate millivolt readings had been obtained for each standard and then these data had been analyzed by linear regression.

Both membrane electrodes produced rather large estimates of the average precision of all analysts. The reasons for this have been discussed.

The total chlorine analyzer consistently gave excellent precision. This technique for measuring total chlorine appears to be independent of the analyst who performed the procedure.

Active chlorine speciation in natural systems

The amperometric titrations were very reproducible for analyzing the tap water samples. For the membrane-covered amperometric electrodes, values were obtained that averaged approximately 21 percent of the values obtained from amperometric titration. (The one exception was for electrode A and sample 5, the electrode determined a value that was 35 percent of the value determined by amperometric titration.) Two possible explanations for these low values are either that the electrodes were not functioning properly or that there could have been active chlorine species in the water which mimicked free chlorine in the amperometric procedure but did not permeate the HOCl specific membrane.

It is possible that the membrane electrodes were not functioning properly and there was a chance that both electrodes gave similar readings. The fact that less chlorine was detected by both membranes indicates that very little, if any, inorganic chloramine (presuming this species was present) permeated the membrane. Generally, if the membrane electrodes were

malfunctioning, the error would have been positive (i.e., higher value), not negative, with respect to the amperometric titrator.

It has been shown by several analytical methods that organic nitrogen chloramines appear as free chlorine.[18] The formation of these compounds has been studied.[10-22] Direct evidence for the presence of organic nitrogen compounds in southern Florida groundwater has been confirmed by the formation of dihaloacetonitriles.[23] No routine method exists for determining organic nitrogen chloramines.[24] Recently, a method has been published, which may provide the necessary derivational procedure for determining organic nitrogen chloramines.[25]

Thus, there is no definitive explanation for the low values obtained with the membrane amperometric electrode. However, one cannot rule out the possibility that the amperometric titrator is not specific for free chlorine and that organic nitrogen chloramines account for falsely high levels of free chlorine.

Summary of the statistical analyses

1. The following observations were made regarding the analysis of the quality assurance samples (in all cases, the level of significance was 95 percent):

- There was no statistical difference between results obtained with the amperometric titrator and those obtained with the potentiometric electrode.
- For all samples, there was a statistical difference between results obtained with the amperometric titrator and those obtained with the total chlorine analyzer. The largest absolute difference in the means was 0.06 mg Cl_2/L.
- There was no statistical difference between the results obtained with the potentiometric electrode and those obtained with the total chlorine analyzer for three of four samples. For the one sample, the mean was 1.21 with the potentiometric electrode and 1.35 with the total chlorine analyzer.
- The precision of all analysts was best with the total chlorine analyzer, followed by the amperometric titrator and then the potentiometric electrode.

2. The following observations were made regarding the analysis of the free chlorine in tap water samples (in all cases, the level of significance was 95 percent):

- The results obtained with the membrane-covered electrodes were not significantly different in three of four samples.
- Both membrane-covered electrodes detected chlorine levels that were an average of 21 percent less than the levels detected with the amperometric titrator.

3. The following observations were made regarding the analysis of the total chlorine in tap water samples (in all cases the level of significance was 95 percent).

- For all samples, the results obtained with the potentiometric electrode were lower than those obtained with the amperometric titrator. The difference was statistically significant in three of four samples.
- The results obtained with the total chlorine analyzer were low in three of four samples, with no statistical difference in the fourth sample. The greatest difference between means was 0.08 mg Cl_2/L.
- For two of four samples, the results obtained with the potentiometric electrode were statistically different from those obtained with the total chlorine analyzer.
- The precision of all analysts was best with the total chlorine analyzer, followed by the amperometric titrator and then the potentiometric electrode.

4. Of all the methods tested, including the amperometric titrator, the total chlorine analyzer produced the best precision for both individual analysts and for all analysts.

5. For all of the electrode procedures, precision of individual analysts was better than the precision of all analysts combined.

Conclusions

The total chlorine analyzer was shown to be analyst independent, providing online, continuous measurement. This instrument would be the method of choice when continuous or many total chlorine analyses are required. Although operation of the potentiometric electrode is easy, the results can vary from analyst to analyst and would require additional quality assurance for in-plant, multi-analyst operation. No conclusion can be made regarding the amperometric membrane electrodes because of the variability in the results.

Acknowledgments

This research was supported by contract DAMD 17-80-C-0164 with the U.S. Army Medical Research and Development Command, Ft. Detrick, Md., and by the Drinking Water Research Center, Florida International University, Miami, Fla.

The opinions expressed in this paper are the private views of the authors and should not be construed as official or reflecting the views of the Department of the Army.

Appendix

Calculation of the *t*-statistic. The *t*-statistic is used to test the hypothesis that two sample means are the same. In calculating the *t*-statistic, it is assumed that the data are normally distributed about the mean. The general formula for determining the *t*-statistic is given in Eq 2.

For example, to compare the results obtained from the amperometric titrator (*a*) with those obtained from the potentiometric electrode (*p*), data for sample 1 from Table 1 are substituted into Eq 2. Therefore, $\bar{x}_a = 0.40$, $s_a^2 = 0.000169$, $n_a = 12$, $\bar{x}_p = 0.38$, $s_p^2 = 0.00360$, and $n_p = 12$, then

$$t = \frac{0.40 - 0.38}{\sqrt{\frac{0.000169}{12} + \frac{0.00360}{12}}} = \frac{0.02}{0.0177} = 1.129$$

In another example, data for sample 3 from Table 1 are substituted into Eq 2 to compare the results obtained from the amperometric titrator (*a*) with those obtained from the *total* chlorine analyzer (*c*). For $\bar{x}_a = 1.26$, $s_a^2 = 0.000529$, $n_a = 12$, $\bar{x}_c = 1.30$, $s_c^2 = 0.000169$, and $n_c = 12$, then

$$t = \frac{1.26 - 1.30}{\sqrt{\frac{0.000529}{12} + \frac{0.000169}{12}}} = \frac{-0.04}{0.00763} = -5.24$$

When calculating the *t*-statistic, the sign is disregarded and a positive value is reported (t=5.24) because differences in either direction are being investigated.

Calculating the adjusted degrees of freedom by using the Satterthwaite approximation. The ADF must be used when the means of two samples are to be compared by using the *t*-statistic, but the two samples have different variances. The general formula for determining the ADF is that given in Eq 3.

For example, to compare the results obtained from the amperometric titrator (*a*) with those obtained from the potentiometric electrode (*p*), one must know the ADF. Therefore, data for sample 1 from Table 1 are substituted into Eq 3. For $s_a^2 = 0.000169$, $n_a = 12$, $s_p^2 = 0.00360$, and $n_p = 12$, then

$$ADF = \frac{\left(\frac{0.000169}{12} + \frac{0.00360}{12}\right)^2}{\frac{\left(\frac{0.000169}{12}\right)^2}{11} + \frac{\left(\frac{0.00360}{12}\right)^2}{11}} = \frac{9.86 \times 10^{-8}}{1.803 \times 10^{-10} + 8.2 \times 10^{-8}} = 1.203$$

= 12 (rounded to nearest integer)

In the sample calculations for the ADF that follow the data are for sample 2 from Table 1, for sample 3 from Table 1, and for sample 6 from Table 4, respectively. In these calculations, *a* refers to the

amperometric titrator and c refers to the chlorine analyzer.

For $s_b^2 = 0.000484$, $n_b = 12$, $s_c^2 = 0.0000240$, and $n_c = 12$, *then*

$$ADF = \frac{1.8 \times 10^{-4}}{1.48 \times 10^{-10} + 3.46 \times 10^{-13}}$$

$$= 12.14$$

$$= 12$$

For $s_b^2 = 0.000529$, $n_b = 12$, $s_c^2 = 0.0000169$, and $n_c = 12$, then

$$ADF = \frac{3.4 \times 10^{-4}}{1.77 \times 10^{-10} + 1.80 \times 10^{-13}}$$

$$= 17.4$$

$$= 17$$

Finally, for $s_b^2 = 0.000256$, $n_b = 12$, $s_c^2 = 0.000144$, and $n_c = 12$, then

$$ADF = \frac{1.1 \times 10^{-4}}{4.14 \times 10^{-11} + 1.3 \times 10^{-13}}$$

$$= 20.4$$

$$= 20$$

These calculations for ADF clearly demonstrate that as the variances (s^2) of two samples approach the same value, the degrees of freedom approach the limiting value. Consider the case in which two samples have equal variance, and $n = 12$ for each sample. In this case, the maximum number of degrees of freedom is 22, i.e., $(n_1 - 1) + (n_2 - 1)$. For the sample calculation given above in which the two variances were 0.000169 and 0.00360 and $n = 12$ for both samples, then the ADF was calculated to be 12. However, for the last sample calculation given above, in which the two variances were 0.000256 and 0.000144, the ADF was calculated to be 20.

Calculation of the level of significance for comparing sample means. The researcher chooses the significance level (probability level) to be used. For a 95 percent significance level, $\alpha = 0.05$; for a 99 percent significance level, $\alpha = 0.01$.

In order to use a *t*-statistic to compare sample means, one must make adjustments for multiple comparison. The Bonferroni procedure can be used to make adjustments for each pairwise comparison by dividing the α value by the total number of comparisons. For example, if there are three means—*A*, *B*, and *C*—and one wishes to compare *A* to *B* and *A* to *C*, the α level is divided by 2. However, if one wishes to compare *A* to *B*, *A* to *C*, and *B* to *C*, the α level is divided by 3.[17]

In the discussion that follows, results obtained from the amperometric titrator are compared with those obtained from the total chlorine analyzer. The data are for sample 2 from Table 2.

For $t = 3.074$ and ADF = 12, the 95 percent confidence level ($\alpha = 0.05$) is obtained by evaluating the two-tailed *t* at $\alpha = 0.05/3 = 0.017$. From such an evaluation, it can be determined that $t = 2.78$ for ADF = 12 and $\alpha = 0.017$. Because the calculated value of $t = 3.074$ is greater than 2.78, the two sample means are statistically different at the 95 percent confidence level.

If the data are reevaluated at the 99 percent confidence level ($\alpha = 0.01$), then $t = 3.65$ for ADF = 12 and $\alpha = 0.01/3 = 0.003$. Because the calculated value of t (3.074) is less than 3.65, the two sample means are not statistically different at the 99 percent confidence level.

References

1. *Water Chlorination Environmental Impact and Health Effects.* Vol. 1. (R.L. Jolley, editor). Ann Arbor Science Publishers, Inc., Ann Arbor, Mich. (1978).
2. *Water Chlorination Environmental Impact and Health Effects.* Vol. 2. (R.L. Jolley, G. Grocheu, and D.H. Hamilton, Jr., editors). Ann Arbor Science Publishers, Inc., Ann Arbor, Mich. (1978).
3. *Water Chlorination Environmental Impact and Health Effects.* Vol. 3. (R.L. Jolley, W.A. Brungs, and R.B. Cummings, editors). Ann Arbor Science Publishers, Inc., Ann Arbor, Mich. (1980).
4. Jolley, R.L. Proc. Fourth Conf. on Water Chlorination: Environmental Impact and Health Effects, Pacific Grove, Calif. (Oct. 1981).
5. *Standard Methods for the Examination of Water and Wastewater.* APHA, AWWA, and WPCF, Washington, D.C. (15th ed., 1981).
6. Guter, K.J.; Cooper, W.J.; & Sorber, C.A. Evaluation of Existing Field Test Kits for Determining Free Chlorine Residuals in Aqueous Solutions. *Jour. AWWA*, 66:1:38 (Jan. 1974).
7. Sorber, C.A., Cooper, W.J., & Meier, E.P. Selection of Field Methods for Free Available Chlorine. *Disinfection: Water and Wastewater.* (J.D. Johnson, editor). Ann Arbor Science Publishers, Inc., Ann Arbor, Mich. (1975).
8. Longley, K.E. et al. Measuring Disinfectant Residuals. Proc. AWWA WQTC, Louisville, Ky. (Dec. 1978).
9. Cooper, W.J., Roscher, N.M., & Siefker, R.A. Determining Free Available Chlorine by DPD-Colorimetric, DPD-Steadifac and FACTS Procedures. *Jour. AWWA*, 74:7:362 (July 1982).
10. Johnson, J.D., Edwards, J.W., & Keenlar, F. Chlorine Residual Measurement Cell: The HOCl Membrane Electrode. *Jour. AWWA*, 70:6:341 (June 1978).
11. Soderberg, R. et al. Testing a Portable Polarographic Electrode for the Measurement of HOCl and Free Available Chlorine. *Water Chlorination: Environmental Impact and Health Effects.* Vol. 3. (R.L. Jolley, W.A. Brungs, and R.B. Cummings, editors). Ann Arbor Science Publishers, Inc., Ann Arbor, Mich. (1980).
12. Stanley, J. & Mossner, R. Measurement of Residual Chlorine in Municipal Wastewater With an Amperometric Membrane Electrode. Proc. Fourth Conf. on Water Chlorination: Environmental Impact and Health Effects, Pacific Grove, Calif. (Oct. 1981).
13. Soderberg, J. & Synnott, J. Monitoring Total Residual Chlorine by Potentiometric Analysis. Proc. Fourth Conf. on Water Chlorination: Environmental Impact and Health Effects, Pacific Grove, Calif. (Oct. 1981).
14. Marinenko, G., Heckert, R.C., & Ferns, D.G. An Instrument With Internal Calibration for Monitoring Chlorine Residual in Natural Waters. *Jour. Fisheries Res. Board Canada*, 33:822 (1976).
15. Marinenko, G. & Skogerboe, H. Monitoring Chlorine at Trace Levels. Proc. Fourth Conf. on Water Chlorination: Environmental Impact and Health Effects, Pacific Grove, Calif. (Oct. 1981).
16. Winer, B.J. *Statistical Principles in Experimental Design.* McGraw-Hill, New York (1971).
17. Chew, V. Comparison Among Treatment Means in an Analysis of Variance. Agricultural Res. Service Rept. 0-228-258, US Dept. Agriculture, Washington, D.C. (1977).
18. Watson, J.E. & Morris, J.C. The Analysis of Free Chlorine in the Presence of Nitrogenous Organic Compounds. Proc. ACS Meeting, Div. Envir. Chem., Anaheim, Calif. (Mar. 1978).
19. Mattice, J.S. et al. Total Residual Chlorine as a Regulatory Tool. Proc. Fourth Conf. on Water Chlorination: Environmental Impact and Health Effects, Pacific Grove, Calif. (Oct. 1981).
20. Morris, J.C. et al. Formation and Significance of N-Chloro Compounds in Water Supplies. USEPA Rept. EPA-600/2-80-013, USEPA, Cincinnati, Ohio (1980).
21. Heinz, G.R., Watson, D.A., & Sieger, A.C. Chlorine Demand Studies Concerning Its Chemical Basis. Proc. Fourth Conf. on Water Chlorination: Environmental Impact and Health Effects, Pacific Grove, Calif. (Oct. 1981).
22. Margerum, D.W., Gray, E.T., & Huffman, R.P. Chlorination and the Formation of N-Chloro Compounds in Water Treatment. *Organometals and Organometalloids: Occurrence and Fate in the Environment.* (F.E. Brinckman and J.M. Bellama, editors). ACS Symposium Series 82, ACS, Washington, D.C. (1978).
23. Trehy, M.L. & Bieber, T.I. Formation of Dihaloacetonitriles by Chlorination of Natural Waters. Proc. Fourth Conf. on Water Chlorination: Environmental Impact and Health Effects, Pacific Grove, Calif. (Oct. 1981).
24. Gray, E.T. & Workman, H.J. Simultaneous Kinetic Analysis of Chlorine and Chloramines in Aqueous Solution at Micromolar Concentration Levels. Proc. Fourth Conf. on Water Chlorination: Environmental Impact and Health Effects, Pacific Grove, Calif. (Oct. 1981).
25. Scully, F.E., Jr., Bemporad, M.A., & Daniel, F.B. Analysis of Organic N-Chloramines. Proc. Fourth Conf. on Water Chlorination: Environmental Impact and Health Effects, Pacific Grove, Calif. (Oct. 1981).

William J. Cooper is an associate professor, M.F. Mehran is a research chemist, R.A. Siefker is a research chemist, D.A. Smith is a graduate assistant, and J.T. Villate is a professor, Drinking Water Research Center, Florida International University, Miami, FL 33199. P.H. Gibbs is a statistician, US Army Medical Bioengineering Research and Development Laboratory, Fort Detrick, Frederick, MD 21701.

APPENDIX B*

Paul H. Gibbs
William J. Cooper
Edward M. Ott

A STATISTICAL EXPERIMENTAL DESIGN FOR COMPARISON TESTING OF ANALYTICAL PROCEDURES

Reprinted from JOURNAL AWWA, Vol. 75, No.11
(November 1983), by permission.
Copyright© 1983, The American Water Works Association

* Note: This Appendix B includes Tables 1 - 10.

A statistical experimental design for comparison testing of analytical procedures

Paul H. Gibbs, William J. Cooper, and Edward M. Ott

An experiment has been designed for comparison (equivalency) testing of an approved standard test method and an alternative test method. The design has been applied in tests under actual operating conditions at water treatment plants. Statistical analysis of an example data set is detailed. The estimates of single-analyst precision obtained with this design may reflect the results encountered at any water treatment plant.

The US Environmental Protection Agency (USEPA) promulgated guidelines to standardize test procedures for compliance monitoring. These guidelines have been amended to allow applications for approval of alternative procedures for nationwide use.[1] As part of the application, comparability data are required between the proposed alternative procedure and an approved procedure. The comparability data are intended to test the equivalency of the two methods. The minimum requirements for equivalency testing have been outlined elsewhere.[2]

This article describes a generalized experimental design that can be used for obtaining the data necessary for comparison testing, a particular application of which is equivalency testing. Another article details results of this experimental design as it was applied to free available chlorine test procedures.[3]

The data required for equivalency testing also provide estimates of single-analyst precision for the analytical methods tested. These estimates can be used for purposes other than equivalency testing and may provide valuable data for evaluating the relative merit of analytical procedures. The major difference between the estimates obtained by using equivalency testing and those obtained by analytical chemists who develop the methods is that the former are obtained under actual operating conditions at water treatment plants. These single-analyst *estimates* of precision, therefore, reflect results that would be obtained at other water treatment plants.

The experimental design is based on several general criteria.

- The test should be conducted over a time span that covers a significant change in either the process or sample quality.
- Replicate analyses should be required on samples for the approved and proposed alternative test procedures at multiple laboratories.
- The test program should require a minimum of analyst time at the cooperating laboratory, i.e., the testing program should be self-explanatory, and forms should be provided for data collection.
- Data should be received in a form that facilitates computer-aided statistical evaluation.
- Randomization should be used to assign experimental factors to the samples and to determine the order of testing of the procedures for each sample.

Experimental design

Experimental factors (variables). Several experimental factors were identified and included in the experimental design.

Methods. An alternative test is used in addition to the approved method currently in use at the test location. The approved (standard) method is denoted by SM, and the alternative method is denoted by AM.

Locations and sites. The experiment is conducted at several geographically dispersed locations (water treatment plants). Water quality data may be collected at more than one site in a plant (after different water treatment processes); thus, the design allows for testing at two sites within any one plant. The exact position of the two sites within the water treatment plant is determined for each plant and reported on the forms; i.e., for the duration of the experiment, one position is designated as site 1 and the other position is designated as site 2.

Analysts. To include variability resulting from analyst effects in the chemical analysis, two analysts who normally take the readings are usually available for the duration of the experiment at each location. The two analysts work on the experiment independently of one another. If only one analyst is available, the same total number of observations (the number taken by two analysts) is taken.

Samples. The duration of the experiment (long enough to include variability in plant operation) calls for sampling over a period of several weeks. Samples are obtained in the same manner as those collected for monitoring normal plant operations. Most samples require one reading with the approved test procedure and one reading with the alternative test procedure.

Order of testing. To guard against bias in the results, the sequence in which results are obtained for any given test is randomized and followed carefully by each analyst. It is necessary to emphasize that prior experimental test results should not influence the analyst's subsequent readings because this will invalidate the experiment. To help guard against the influence of previous results, the instructions to the analysts explicitly point out this potential problem.

TABLE 1
Experimental design for equivalency testing at one water plant

Sample Number	Day	Analyst	Site	Randomized Test Sequence
1	1	2	2	SM AM
2	1	2	2	SM AM
3	2	1	2	SM AM
4	2	1	2	SM AM
5	3	1	1	AM SM
6	3	1	1	SM AM
7	4	2	2	AM SM
8	4	2	1	AM SM
9	5	1	1	SM AM AM SM AM SM AM SM
10	5	1	2	SM AM SM AM AM AM SM SM
11	6	2	1	SM AM
12	6	2	2	SM AM SM AM AM SM SM AM
13	7	2	2	SM AM SM AM SM AM SM AM
14	7	1	1	AM SM SM SM AM AM SM AM
15	8	1	1	SM AM
16	8	2	2	SM AM
17	9	2	1	SM AM
18	9	2	2	SM AM
19	10	1	1	AM AM
20	10	2	2	AM SM
21	11	1	1	AM SM AM SM AM SM SM AM
22	11	1	2	SM AM
23	12	1	2	SM AM
24	12	1	2	AM SM
25	13	1	1	SM AM
26	13	1	2	AM SM
27	14	2	1	AM SM
28	14	2	1	AM SM SM AM SM AM AM SM
29	15	2	1	AM SM
30	15	2	1	SM AM AM SM AM SM AM SM
31	16	1	2	AM SM AM SM SM AM AM SM
32	16	1	1	AM SM
33	17	1	2	AM SM AM SM AM SM SM AM
34	17	1	1	AM SM
35	18	1	1	AM SM
36	18	2	1	SM AM
37	19	2	1	AM SM SM AM SM AM SM AM
38	19	2	2	AM SM
39	20	2	2	SM AM
40	20	2	2	AM SM
41	21	1	2	SM AM
42	21	2	1	AM SM
43	22	2	2	SM AM
44	22	2	1	SM AM
45	23	2	1	SM AM
46	23	1	2	SM AM
47	24	1	1	SM AM
48	24	2	2	AM SM
49	25	1	1	AM SM
50	25	2	1	AM SM
51	26	2	1	AM SM
52	26	1	2	AM SM
53	27	1	1	AM SM
54	27	2	2	SM AM
55	28	2	2	SM AM AM SM AM SM SM AM
56	28	1	2	SM AM
57	29	1	2	SM AM
58	29	2	1	AM SM
59	30	1	2	AM SM
60	30	1	1	AM SM

TABLE 2
Sample data forms provided to water treatment plants

Data Form

Sample No. 1 — Date ____ (Mon / Day / Yr)
Plant ID — Time ____ AM PM
Analyst 2 — Day 1 — Morning
Site 2:

Readings

Order	Methods	Results
1	SM	
2	AM	

Comments

Data Form

Sample No. 9 — Date ____ (Mon / Day / Yr)
Plant ID — Time ____ AM PM
Analyst 1 — Day 5 — Morning
Site 1:

Readings

Order	Methods	Results
1	SM	
2	AM	
3	AM	
4	SM	
5	AM	
6	SM	
7	AM	
8	SM	

Comments

Subsampling. To meet the minimum requirements of the experimental design, multiple analyses are made on a preselected, restricted random selection of samples. The selection of samples is balanced with respect to analyst and site combinations so that each combination of analyst and site is represented equally. The multiple analyses of each sample also have a preselected random sequence, such as AM SM SM AM AM SM AM SM. In these cases, again, the analyst exercises care to guard against biasing the results of subsequent determinations by preceding observations in the sequence. These additional sample analyses are used to estimate single-analyst precision.

Duration. The experiment requires 30 working days to complete. This length of time results in minimum daily sampling and encompasses a time span sufficient to observe variability in plant operation.

Design. The 30 days allotted for the experiment are sequential but not necessarily consecutive; i.e., day 5 precedes day 6, but day 5 is not necessarily adjacent to day 6. Each day is split into a morning and an afternoon test session, although no attempt is made to balance the design for day effect. Each segment of the day is treated as a separate sample. The design assumes that two analysts and two sites are available.

If only one site is available, all tests are performed at that site, including those designated "site 2." Having two analysts and two sites available at each plant broadens the scope of the experiment but is not necessarily applicable to all water quality tests. The specific analyst and site combination for each test sample is determined by a random process so that 15 samples are allocated to each combination of analyst and site, for a total of 60 samples. Subsamples for multiple determinations are chosen by random selection from the samples allocated to each analyst and site combination. In this case, three samples from each combination of analyst and site were used. The sequence in which methods are tested with each sample is also random. A completely different randomization is performed for each water plant (regardless of the number of plants).

Computer program. A computer program was written to provide the necessary randomizations to conform to the balanced experimental design required for statistical analysis. The output of this program

TABLE 3
The ANOVA table for estimating single-analyst precision within each location

Source	Degrees of Freedom
Total	48
Mean	1
Samples	11
Within samples	36

TABLE 4
The ANOVA table for testing method effects within location

Source	Degrees of Freedom
Total	120
Mean	1
Between samples	59
Analyst	1
Site	1
Analyst × site	1
Samples, analyst × site	56
Within samples	60
Methods	1
Methods × analyst	1
Methods × site	1
Methods × analyst × site	1
Experimental error	56

TABLE 5
Estimates of the minimum positive detectable true difference

Case		
1	0.00A	0.06
2	0.0[illegible]	0.0[illegible]
3	3.0[illegible]	3.10

TABLE 7
The ANOVA results for single-analyst precision [illegible]

Method	Source	df	SSQ	MSQ
SM	Between samples	[illegible]	[illegible]	[illegible]
	Analysts	[illegible]	[illegible]	[illegible]
AM	Between samples	[illegible]	[illegible]	[illegible]
	Analysts	[illegible]	[illegible]	[illegible]

[illegible]

TABLE 8
Single-analyst precision estimates [illegible]

[illegible]		[illegible]	
SM	AM	SM	AM
[illegible]	[illegible]	[illegible]	[illegible]

TABLE 6
Experimental results[a] obtained for the design in Table 1

Sample Number	Analyst	Site	Day	Test[b]	Results	Test	Results[c]
1	2	2	1	1	2.10	3	2.20
2	1	1	1	1	2.00	3	2.20
3	1	2	1	1	2.20	3	2.30
4	2	1	2	1	2.10	3	2.20
5	1	1	3	3	0.50	5	0.50
6	2	1	3	1	0.70	3	0.80
7	1	2	4	1	1.70	3	1.80
8	2	1	4	3	0.55	5	0.30
9	1	2	5	3	0.60	5	0.50 [illegible]
10	1	1	5	1	2.60	3	2.20 [illegible]
11	2	1	6	1	0.80	3	0.80
12	2	2	6	1	2.40	3	2.20 [illegible]
13	1	2	7	1	2.10	3	2.20 [illegible]
14	2	1	7	3	0.50	5	0.25 [illegible]
15	1	1	8	1	0.50	3	0.30
16	2	1	8	1	1.80	3	1.90
17	1	2	9	3	0.60	5	0.30
18	2	2	9	1	1.90	3	1.70
19	1	2	10	3	1.20	5	0.10
20	2	2	10	1	1.80	3	1.90
21	1	1	11	3	0.10	5	0.20 [illegible]
22	2	2	11	1	2.30	3	2.10
23	1	2	12	1	2.10	3	2.20
24	2	2	12	1	2.40	3	2.40
25	1	1	13	1	0.80	3	0.40
26	2	1	13	1	1.90	3	1.70
27	1	2	14	1	0.30	3	0.50
28	2	1	14	3	0.50	5	0.50 [illegible]
29	1	1	15	3	0.55	5	0.60
30	2	1	15	3	0.30	5	0.10 [illegible]
31	1	2	16	1	2.70	3	2.20 [illegible]
32	2	2	16	1	1.90	3	1.80
33	1	2	17	1	2.20	3	2.40 [illegible]
34	2	2	17	1	1.70	3	1.90
35	1	1	18	3	0.10	5	0.40
36	2	1	18	1	1.60	3	1.80
37	1	2	19	1	2.30	3	2.50
38	2	2	19	1	2.00	3	2.10
39	1	2	20	1	2.20	3	2.40
40	2	2	20	1	1.80	3	1.90
41	1	1	21	3	0.30	5	0.50
42	2	1	21	1	1.80	3	1.90
43	1	2	22	1	2.20	3	2.30
44	2	2	22	1	1.90	3	1.90
45	1	1	23	3	0.50	5	0.40
46	2	1	23	1	1.70	3	1.80
47	1	2	24	1	2.10	3	2.30
48	2	2	24	1	1.90	3	1.90
49	1	2	25	1	2.30	3	2.40
50	2	2	25	1	1.80	3	1.90
51	1	1	26	3	0.30	5	0.40
52	2	1	26	1	1.90	3	1.90
53	1	2	27	1	2.20	3	2.30
54	2	2	27	1	1.80	3	1.90
55	1	1	28	3	0.40	5	0.30
56	2	1	28	1	1.80	3	1.90
57	1	2	29	1	2.30	3	2.40
58	2	2	29	1	1.90	3	1.90
59	1	1	30	3	0.40	5	0.30
60	2	1	30	1	1.90	3	1.90

[a] Results are reported in milligrams per liter.
[b] Test 1 was the approved standard test method (SM), and test 3 was the alternative test method (AM).
[c] Results are listed in the order in which the analyses were performed.

for one treatment plant is summarized in Table 1.

To read Table 1, consider the first line as an example. On day 1, analyst 2 is to obtain a reading at site 2 with the approved standard test method followed by a reading with the alternative test method. The first sample of each day is obtained in the morning; the second is collected in the afternoon. An example of a sample on which multiple analyses are required is sample number 9. On day 5, analyst 1 is to obtain a sample at site 1 and make four analyses using the approved standard test method and four analyses using the alternative method in the following order: SM AM AM SM AM SM AM SM.

The computer program provides printed data forms corresponding to each design. For every location, a total of 60 data forms are printed out. Sample data forms are shown in Table 2.

Thus, a complete layout of the experimental design with data collection forms is provided to each participating treatment plant to assist personnel in planning, scheduling, and recording results.

Statistical analysis

Definition of terms. 1. *Single-analyst precision.* A measure of the variability between repeated measurements made by a single analyst on the same sample using the same method under uniform conditions, usually expressed as a standard deviation.

2. *Experimental error.* A measure of variability in the data not accounted for by experimental factors and composed in part of single-analyst precision, usually expressed as a variance.

3. *ANOVA.* Analysis of variance, a statistical technique for attributing variability in data to certain experimental factors and for assessing the significance of the factors in accounting for the total variability.

4. *Sum of squares.* A mathematical quantity that expresses the amount of the total variability in the observations attributed to each experimental factor.

5. *Degrees of freedom*: Number of cases used to estimate a mean square, decreased by the number of mathematical constraints on the statistic. Usually the number of constraints is one.

6. *Mean square*: The sum of squares divided by the degrees of freedom.

7. *F-test*: A statistical test for determining whether an experimental factor significantly explains the variability in the data. It is calculated by dividing the mean square by the appropriate error term for each experimental factor being tested.

8. *Error mean square*: An estimate of the experimental error variance determined from the data by an ANOVA.

9. *Power*: The probability of rejecting the hypothesis of equality of means, or the sensitivity of the statistical test. Power depends, in part, on the magnitude of the experimental error relative to the true difference in means and on the sample size [see the section on power under "Interpretation of Results"].

10. *Significance level*: The preassigned probability of concluding that the methods are not equivalent when, in fact, the methods are equivalent.

11. *Main effect*: Refers to the effects on the measured response of one of the experimental variables.

12. *Interaction*: A measure of the dependence of one main effect on one or more other main effects.

13. *Within-sample variation*: A measure of the variation of observations made on the same sample.

14. *Between-sample variation*: A measure of the variation between observations on different samples.

15. *Coefficient of variation*: The standard deviation divided by the mean and expressed as a percentage.

16. F_{max} *test*: A statistical test used to compare the variances of two or more samples.

Single-analyst precision. Using only those 12 samples with multiple determinations of eight readings (four by each method), single-analyst precisions are estimated for each method and compared by statistical tests. The estimates are derived within each location and represent the variation between repeated observations averaged over both analysts in each location (Table 3).

Single-analyst precision is estimated from the within-sample variation of repeated determinations on a randomly selected subset of samples made over the span of the experiment. For each location, 48 observations per method are available for analysis (4×12). The subsample selected is balanced with respect to combinations of analyst and site to avoid biasing the estimates of precision.

If the means of the methods vary significantly, it would be advisable to use coefficients of variation to estimate and compare precisions of the methods because precision is often correlated with the mean. Otherwise the single-analyst precisions may be compared by using the F_{max} test.[7]

Comparison testing. The ANOVA analysis (Table 4) is performed for each location. The additional observations for those samples with multiple determinations are discarded, and only the first paired determinations are used. Therefore, for 60 samples and two determinations, 120 observations are analyzed.

There are two error terms: samples (analyst × site), which is the error term for between-sample comparisons, such as analyst, site, and analyst × site, and the experimental error, which serves as the error term for all method effects including interactions. The experimental factors are then tested, using *F*-tests, against their appropriate error terms.[7]

The analysis is comparable to a paired *t*-test on methods in that between-sample variation is removed from the analysis before the method effect from within-sample variation is estimated. The interactions of method with analyst and with site provide tests of whether the method effect is analyst-dependent or site-dependent. Evidence of either dependency calls for investigation of the underlying causes. The three-way interaction of method with analyst and site is usually difficult to interpret. If present, this source of variation sometimes indicates dependence of the method effect on the combination of site and analyst. Under this circumstance, it is impossible to draw valid conclusions about the method effect from the data. Underlying causes should again be examined. The relative magnitude of the interactions should be examined by using tables of means and standard errors. Competent statistical advice should be sought in the interpretations of interactions because the ANOVA may be overly sensitive to minor differences in the data, which could lead to overinterpretation of the results.

An overall assessment of equivalency of methods can be obtained by a simple paired *t*-test on the method means for each plant. This test assumes that the method precisions do not change between locations. If this is not the case, an overall test of equivalency may not be possible, and only tests for equivalency within each location may be done.

Interpretation of results

Results of the ANOVA should be treated cautiously because results of hypothesis testing are largely dependent on the sample sizes and experimental errors. Overly sensitive statistical tests may detect differences that are not meaningful from an operational point of view. Also, insensitive statistical tests may not detect differences that are operationally meaningful. The definition of operationally meaningful differences must be agreed on prior to the experiment, not after the data have been collected, to avoid the introduction of biased definitions.

If possible, the power of the tests should be examined to determine the sensitivity of the ANOVA (see the following section on power). If the statistical tests are not sensitive enough, the minimum detectable difference will be larger than is operationally acceptable, and the tests will not detect differences of operational importance. Increasing sample sizes may alleviate this problem unless the methods have inherently large precisions. On the other hand, if the minimum detectable difference is too small to be operationally meaningful, the statistical tests may signal rejection of equivalency on the basis of small differences in means. In this latter case, one can revise the hypothesis being tested to include an "acceptable" window of difference in the method means and test for departures from this window. Also, and this is true in all cases, the confidence intervals on the mean differences can be examined to assess the magnitude of the differences detected by the ANOVA. This latter step is somewhat subjective unless a prior understanding has been reached before the experiment is run as to the magnitude of the acceptable difference between methods.

Power. As defined previously, power is a measure of the sensitivity of a statistical test of a hypothesis. It can be expressed in several ways, but the most convenient is the minimum detectable difference. This is the minimum true difference in method means that can be detected statistically with a given probability, at a given significance level, for a given number of samples, and for a given estimate of experimental error. In one case, power is set at 0.80, and the significance level is set at 0.05.

Let Δ be the minimum positive detectable true difference. If n paired samples are available and S^2 is the estimated experimental error from the ANOVA for equivalency (based on at least 30 degrees of freedom), then Δ can be estimated approximately by

$$\Delta = 2.80 \sqrt{2S^2/n}$$

Example calculations of Δ, for $n = 30$ and for various estimates of S^2, are shown in Table 5. Thus, for a significance level of 0.05 and 30 paired samples with an estimated experimental error of 0.008, there is an 80 percent chance of detecting a true difference in method means of at least 0.06 or greater by using statistical tests.

Example data set and statistical analysis

To illustrate the procedure that is used in analyzing the data, an example data set is analyzed in this section. Table 6 is the result of the design shown in Table 1. Codes for SM and AM are 1 and 3, respectively. In this example, the pre-

TABLE 9

The ANOVA results for testing method effects (× 100).

Source	df	SS	MS	Fs
Between samples	59	10039.625		
Analyst	1	10.208	10.208	0.64
Site	1	9135.075	9135.075	575.92**
Analyst × site	1	6.075	6.075	0.38
Samples within analyst × site	56	888.272	15.862	
Within samples	60	101.500		
Methods	1	23.408	23.408	18.38**
Methods × analyst	1	0.075	0.075	0.06
Methods × site	1	8.008	8.008	6.29**
Methods × analyst × site	1	0.675	0.675	0.53
Experimental error	56	71.144	1.274	

[illegible]

[illegible] significant effects, $\alpha = 0.05$

TABLE 10

Confidence intervals on mean method differences by site

Site	Sample Size	Means		Difference in Means	95 Percent Confidence Interval on Difference

chlorination site is site 1, and the postchlorination site is site 2. Most of the results can be obtained by using any convenient statistical package available from computer services.

Single-analyst precision. The 12 samples with multiple determinations are selected, and the estimates of precision are calculated for each method by using a one-way ANOVA. The data, which are multiplied by 10 to increase resolution, are presented in Table 7.

Because the data are multiplied by 10, the variances are multiplied by 100. Thus, the variance estimates must be divided by 100 to obtain the original scale:

$$\sigma^2_{SM} = 0.00625$$
$$\sigma^2_{AM} = 0.00528$$

The precisions are then the square roots of the variances:

$$\sigma_{SM} = 0.079 \text{ mg/L}$$
$$\sigma_{AM} = 0.073 \text{ mg/L}$$

To test for equality, the ratio of the largest variance to the smallest is computed by using the F_{max} test:

$$F_{max} = 0.00625/0.00528 = 1.18$$

This value is referred to the appropriate F_{max} statistical table and declared not to be significant at the 99 percent level of confidence.

The relative precisions may also be calculated by computing the single-analyst precisions as percents of the means (Table 8) for prechlorination (site 1) and postchlorination (site 2).

In this particular case, the prechlorination means are much lower than the postchlorination means, resulting in different patterns of relative precision. Nevertheless, the methods are comparable within each site.

Equivalency testing. The results of the ANOVA used for comparison (equivalency) testing are shown in Table 9.

The analyst, site, and analyst × site factors are tested against samples (analyst × site), a between-sample error term that is less precise than the within-sample comparisons. Methods, and all interactions with method, are tested against experimental error. Thus, the method comparisons are within-sample comparisons and are more precisely determined than site and analyst comparisons.

The estimate of experimental error (S^2) is 0.01274. This estimate includes variation resulting from single-analyst precision as well as other factors not accounted for in the model.

Ninety-five percent confidence intervals on the actual differences for both sites can now be calculated by using the experimental error estimate:

$$D \pm 1.96 \sqrt{2S^2/30}$$

where D = difference in site means, S^2 = experimental error, and 30 = number of pairs in the test. The results of these calculations are given in Table 10.

A pronounced site effect is evident in the magnitude of the site means and the interaction of site and method as indicated by differences in the effect of methods between sites. The method difference is due to the prechlorination site alone for which the difference is 0.14 mg/L. For the postchlorination site, the observed difference (0.03 mg/L) has 95 percent confidence intervals that include 0; hence the hypothesis of equality of means cannot be rejected. In this case, the test of methods is inconclusive despite the strong method effect in the ANOVA, owing to the interaction of method and site.

The estimated experimental error of 0.01274 would suggest a minimum detectable difference of about 0.07 mg/L with 80 percent power. This indicates that the test is sensitive to small differences in the method means.

Summary

A statistically designed experiment for comparison (equivalency) testing of analytical methods used at water treatment plants has been developed. The design takes into account the variables most often encountered in routine water quality determinations. This design should allow for the testing of numerous water quality parameters and for realistic estimates of single-analyst precision obtained with minimum involvement of any one treatment plant. It is possible that the estimates obtained by using this experimental design are the best approximation of what may be encountered in any water treatment plant.

Acknowledgments

This research was supported by the U.S. Army Medical Research and Development Command, Ft. Detrick, Md., under contract DAMD17-80-C-0164 and by the Drinking Water Research Center, Florida International University, Miami, Fla. The opinions contained in this article are those of the authors and are not to be construed as an official position of the Department of the Army unless so designated by other authorized documents.

The complete computer program, example data sheet, and instructions to operators are available from the Drinking Water Research Center, Florida International University.

References

1. Guidelines Establishing Test Procedures for the Analysis of Pollutants. Fed. Reg. 40:24535 (June 9, 1975).
2. Guidelines Establishing Test Procedures for the Analysis of Pollutants. Fed. Reg. 41:232 52780 (Dec. 1, 1976).
3. Alternative Test Procedure Equivalency Testing System. Envir. Monitoring and Support Lab., USEPA, Cincinnati, Ohio.
4. Cooper, W.J. et al. Equivalency Testing of Test Procedures for Free Available Chlorine, Amperometric Titration, DPD and FACTS. Accepted for publication in Jour. AWWA.
5. Winer, B.J. *Statistical Principles in Experimental Design*. McGraw-Hill Book Co., New York (1962).

Paul H. Gibbs is a statistician, US Army Medical Bioengineering Research and Development Laboratory, Ft. Detrick, MD 21701. William J. Cooper is an associate professor, Drinking Water Research Center, Florida International University, Miami, FL 33199. Edward M. Ott is a systems analyst, Gould Inc., SEL Computer Systems Division, Ft. Lauderdale, FL 33313.

APPENDIX C*

William J. Cooper
Paul H. Gibbs
Edward M. Ott
Perin Patel

EQUIVALENCY TESTING OF PROCEDURES FOR MEASURING FREE AVAILABLE CHLORINE: AMPEROMETRIC TITRATION, DPD, AND FACTS

Reprinted from JOURNAL AWWA, Vol. 75, No.12
(December 1983), by permission.
Copyright© 1983, The American Water Works Association.

* Note: This Appendix C includes Figure 1 and Tables 1 - 13.

Equivalency testing of procedures for measuring free available chlorine: amperometric titration, DPD, and FACTS

William J. Cooper, Paul H. Gibbs, Edward M. Ott, and Perin Patel

The results of 192 analyses for free available chlorine were obtained from each of 16 participating laboratories. These data were statistically analyzed to determine whether the DPD, FACTS, and amperometric titration methods were equivalent. This analysis indicated that FACTS is equivalent to both DPD and amperometric titration.

Chlorine is used in the water treatment process for various purposes. Of these, the most common applications are taste and odor control, color removal, and disinfection.

The numerous methods available for determining chlorine residuals in aqueous solutions[1,2] include iodometric titration, amperometric titration, and several colorimetric procedures. Iodometric titration may be used to measure total residuals >1 mg Cl_2/L. Amperometric titration is capable of differentiating among free chlorine (hypochlorous acid-hypochlorite ion), monochloramine (NH_2Cl), and dichloramine ($NHCl_2$) and is generally the method of choice in the laboratory. Field measurements are limited by the complexity of the instrumentation. Several colorimetric determinations are used in both the laboratory and the field: the DPD (N,N-diethyl-p-phenylenediamine), LCV (leuco crystal violet), and FACTS (free available chlorine test with syringaldazine) procedures.

All test procedures used for monitoring in compliance with the National Interim Primary Drinking Water Regulations (NIPDWR) must be approved by the US Environmental Protection Agency.[3] This article reports the results of (1) a comparison of the approved standard test procedure (DPD) with the FACTS test procedure, (2) a comparison of the FACTS test procedure with amperometric titration, and (3) a limited three-way comparison of DPD, FACTS, and amperometric titration.

Methods and materials

The experimental design and proposed statistical analysis for this study have been described in detail previously.[4] Briefly, two analysts at each water plant ran a paired sample comparison for 30 days, during which 60 samples were analyzed—30 by one analyst and 30 by the other. If prechlorination (before treatment) was practiced, half of the samples were obtained after prechlorination and half after postchlorination. Each sample was analyzed by two methods; this resulted in 60 analyses per analyst or a total of 120 analyses. In addition, for six samples, each analyst was required to make four measurements with the standard test procedure used at that particular water treatment plant and four measurements with the alternative test procedure (FACTS). This involved a total of 72 additional tests (36 per analyst). The repeated observations were divided between the prechlorination and postchlo-

rination sites if prechlorination was practiced. Thus, a grand total of 192 analyses per plant was obtained.

For the alternative test procedure, a FACTS test kit* was supplied to each water treatment plant participating in this study. The DPD or amperometric procedure normally used at each plant served as the standard test procedure.

Results and discussion

General participants. A minimum of six participating laboratories is necessary to conduct this test, and these must be representative and geographically dispersed water treatment plants. Sixteen plants cooperated in this particular study. Ten treatment plants provided results obtained with the DPD and FACTS methods, and six plants provided results obtained with the amperometric titration and FACTS procedures.

One plant provided results obtained with the DPD and FACTS procedures as well as a single amperometric titration result for each test sample, thus allowing a limited three-way comparison.

Statistical methods. Although the design used for this experiment has been described elsewhere,[4] the main points of the analysis are described here for clarity. The single-analyst precision was estimated from the within-sample variation (36 degrees of freedom) for each location. Only those samples with multiple determinations (12 samples with four measurements by each method) were used to determine the single-analyst precision. The relative magnitudes of precision were expressed as percentages (coefficients of variation) of the site means for each location. Finally, the variances of the two methods for each location were compared by the F_{max} statistical test.[5] For this experiment, an F-value of 2.49 or greater was required to reject the hypothesis of equality of variances at the 99 percent level of confidence. This analysis compared only the precision of the methods.

To test for equivalency of methods, an analysis of variance (ANOVA) procedure was used.[6] All 60 samples were used in order to have a representative range of samples. Multiple determinations were truncated by omitting the extra observations and retaining only the first determination by each method. This was done because multiple determinations do not represent true replicates of the methods, owing to carry-over effects from previous observations in the sequence. Thus variation among repeated observations was not considered a viable estimate of experimental error. The ANOVA table for those locations with no missing data is given in Table 1.

The error mean square, i.e., the experimental error sum of squares divided by the degrees of freedom, was used to test for method effect and to estimate the experimental error for use in confidence limits. The method comparison was a within-sample comparison and, thus, not affected by sample-to-sample variability. The sample means could therefore vary considerably from day to day without disturbing the comparison of methods. The effects of analyst, site, and analyst × site were between-sample effects and therefore not precisely determined by this experiment because variation in sample means tends to inflate the error term used to test these effects. However, these effects were less important than the method effect. In addition, the interactions of method with other effects were also within-sample comparisons and therefore also tested in this model against the error mean square. The analysis was analogous to a paired t-test, in that sample variation was removed prior to comparison of the methods. A balanced design with respect to analyst and site provided additional information beyond a simple paired t-test through use of the ANOVA procedure to analyze the other factors and interactions in the model.

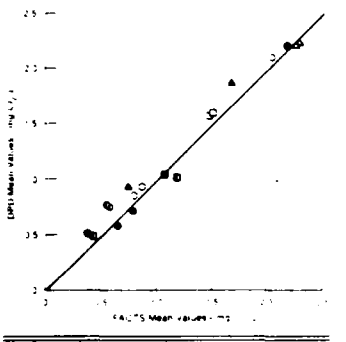
Power of the method comparisons. Power refers to the probability of rejecting the hypothesis of equality of population means or to the sensitivity of the statistical test. Power depends in part on the magnitude of the experimental error relative to the true difference in means and on the sample size. In general, the larger the number of samples, the smaller the difference that can be detected. In this experiment 60 samples were taken, 30 from each site in most cases. Power calculations have been performed for a range of experimental errors.[4] From these calculations it was observed that the experiment was able to detect small differences in the method means (from ±0.06 mg/L to ±0.10 mg/L or greater) when testing at a 0.05 significance level with 80 percent confidence. Therefore, in this study, if the F-test indicated a difference in the method means, the actual magnitude of the mean differences and the 95 percent confidence limits on the differences were examined. This step involved scientific judgment as to what level of difference was operationally important. If such a difference could be determined before the experiment, then the statistical tests could be formulated to test for departures from that specified value rather than from the zero difference value generally used. It is important that the definition of operationally important differences be determined prior to conducting the experiment to avoid biasing the definition by the experimental outcome.

DPD and FACTS

Single-analyst precision results. Single-analyst precision is the standard deviation of repeated measurements made on a single sample by a single analyst under uniform conditions.[6] Table 2 lists estimates of single-analyst precision within each location for the comparison of DPD and FACTS. In only one case, location 4, was a difference in precision detected between the two methods at the 99 percent

TABLE 1
The ANOVA table for testing method effects within location

Source	Degrees of Freedom
Total	120
Mean	1
Between samples	59
Analyst	1
Site	1
Analyst × site	1
Samples (analyst × site)	56
Within samples	60
Method	1
Method × analyst	1
Method × site	1
Method × analyst × site	1
Experimental error	56



Location	Symbol	Mean Values DPD	Mean Values FACTS
1	▲	1.88	1.86
2	●	1.98	1.94
3	◐	2.75	2.58
4	◧	1.49	1.81
5	□	1.58	1.47
6	△	2.23	2.28
7	●	2.53	2.53
8	○	2.10	2.17
9	■	1.95	1.86
10	○	1.85	1.86

Figure 1. Comparison of means for free available chlorine obtained with FACTS and DPD test procedures at 10 water treatment plants

*The FACTS® kit was manufactured by Palintest Ltd., Gateshead, England.

TABLE 2
Single-analyst precision for the DPD and FACTS test procedures

Location	Single-Analyst Precision, mg L*		Single-Analyst Precision				Fmax Test of Variances
			As Percent of Prechlorination Mean		As Percent of Postchlorination Mean		
	DPD†	FACTS	DPD†	FACTS	DPD†	FACTS	
1	0.08	0.10	4.3	6.0	8.6	13.7	1.49
2	0.03	0.02	5.2	3.1	4.2	2.6	2.31
3	0.03	0.00	4.0	0.0	3.9	0.0	‡
4	0.10	0.17	9.8	14.4	2.0	4.1	3.13§
5	0.23	0.21	14.6	14.3	14.6	14.0	1.16
6	0.07	0.07	3.1	3.1	3.1	3.1	1.04
7	0.08	0.07	15.7	18.9	3.6	3.2	1.16
8	0.08	0.06			3.8	3.0	1.86
9	0.08	0.11			7.6	10.4	2.02
10	0.04	0.04	4.2	4.7	4.7	5.1	1.10

*Based on 36 degrees of freedom after adjusting for between-sample variation for each method
†DPD-colorimetric procedure
‡Not applicable
§$p < 0.01$

TABLE 3
ANOVA F-values for experimental factors in the comparison of the DPD and FACTS test procedures

Location	Degrees of Freedom	Analyst	Site	Method	Analyst × Site	Method × Analyst	Method × Site	Method × Site × Analyst	Error Mean Square (× 10^3)
1	1, 54	2.19	**53.66***	**55.66***	3.27	0.57	0.04	0.08	22.4
2	1, 56	1.16	**57.98***	**57.62***	2.32	2.58	0.84	0.21	1.58
3	1, 56	0.53	0.00	**58.00***	0.00	**4.58***	1.52	0.25	22.30
4	1, 56	**11.00***	**188.8***	**4.30***	1.13	1.95	**38.48***	0.49	10.95
5	1, 56	2.21	0.08	**23.72***	1.35	1.17	0.13	0.03	10.24
6	1, 56	0.44	0.12	2.82	0.08	0.00	0.52	0.16	13.02
7	1, 56	0.64	**575.0***	**18.38***	0.38	0.06	**6.29***	0.53	12.74
8	1, 58	1.28		**18.27***		0.16			8.20
9	1, 59			0.04					8.13
10	1, 56	0.15	2.89	**16.58***	0.10	5.07	0.52	1.01	9.08

*$p < 0.05$

TABLE 4
*Paired mean differences and 95 percent confidence intervals for the comparison of the DPD and FACTS test procedures**

Location	Prechlorination Site					Postchlorination Site				
	Sample Size	Means		Difference (DPD - FACTS)	95 Percent Confidence Interval	Sample Size	Means		Difference (DPD - FACTS)	95 Percent Confidence Interval
		DPD	FACTS				DPD	FACTS		
1	30	1.88	1.86	0.22	0.14 0.29	28	0.93	0.73	0.20	0.12 0.28
2	30	0.58	0.64	-0.06	-0.08 -0.04	30	0.72	0.77	-0.05	-0.07 -0.03
3	30	0.75	0.58	0.17	0.09 0.25	30	0.77	0.54	0.23	0.16 0.31
4	30	1.02	1.18	-0.16	-0.21 -0.11	30	0.49	0.41	0.08	0.03 0.13
5	30	1.57	1.47	0.10	0.05 0.15	30	1.58	1.50	0.08	0.03 0.13
6	30	2.23	2.28	-0.05	-0.11 0.01	30	2.23	2.25	-0.02	-0.08 0.04
7	30	0.51	0.37	0.14	0.08 0.20	30	2.20	2.17	0.03	-0.02 0.08
8						60	2.10	2.03	0.07	0.03 0.10
9						60	1.05	1.06	-0.01	-0.04 0.03
10	30	0.94	0.86	0.08	0.04 0.13	30	0.85	0.78	0.08	0.01 0.11

*Confidence intervals based on mean square error from ANOVA, i.e. $\Delta \pm 1.98\sqrt{2S^2/m}$ where Δ = observed difference in means, S^2 = mean square error from ANOVA, and m = number of sample pairs

TABLE 5
Paired t-test results for the comparison of DPD and FACTS means at each location by site

Site	Means		Difference (DPD - FACTS	Number of Pairs	Standard Error of Difference	95 Percent Confidence Limits on Difference	Paired t Test	p Value
	DPD	FACTS						
Prechlorination	1.185	1.128	0.056	8	0.047	-0.08 0.19	1.20	0.27
Postchlorination	1.291	1.223	0.068	10	0.029	0.00 0.13	2.37	0.042

level of confidence. In all other cases, the methods had comparable precision, varying from 2.0 percent to 18.9 percent of the respective site means. In one case (location 3) the estimate of precision was 0.0 because all repeated observations were the same for the given method. For this location, no comparison of precision could be made.

Overall, no systematic effect was observed between methods: sometimes FACTS had smaller relative percent precision than DPD and sometimes it did not.

Equivalency test results. The ANOVA F-values and estimated error mean squares (× 10^3) for the comparison of the DPD and FACTS test procedures are reported in Table 3. For all results, $\alpha = 0.05$ was used as the level of significance. A consistent method effect was detected in all but two cases (locations 6 and 9). Sporadic interactions of method with other factors were detected, but in general, the data show few interactions. A pronounced site effect was detected in four of the eight locations that had more than one sampling site. Only one location had an analyst effect (location 4); although measurements by analyst 2 tended to be higher than those of analyst 1, no effect of analyst × method was detected for this location, indicating that the method differences were comparable for each analyst. A puzzling reversal of the relative magnitudes of the method means across the two sites was detected for location 4 (Table 4), making the F-test for method meaningless. Another case of interaction (method × site for location 7) also resulted in a meaningless F-test for method, as did the interaction of method × analyst at location 3. Thus, three of the eight cases of statistically significant F-tests for method were statistically meaningless because of interactions, leaving five cases of method differences. Owing to previously discussed considerations of the power or sensitivity of this test, however, it was necessary to evaluate the actual magnitude of the mean difference for each location and site.

Table 4 shows that the absolute value of the method differences varied from 0.01 to 0.23 mg Cl$_2$ L. In no case did the upper or lower 95 percent confidence limit on the difference in means exceed 0.31 mg L in absolute value. To illustrate the spread of these differences, a graph of DPD versus FACTS mean values is shown in Figure 1. The 45° line represents no difference in means. This graph shows the wide range of sample values in this experiment. The dispersion of points above and below the line indicates no pronounced bias toward either method. Because the original F-test is extremely sensitive to departures from this 45° line, a further analysis was performed. For each site, the individual pairs of mean values were subjected to a paired *t*-test. In this case the data for each location were reduced to a mean value for each method at each site. Systematic departures from the 45° line were then tested across locations. The results in Table 5 indicate no statistical difference between the methods at the prechlorination site ($p = 0.27$) and marginal difference ($p = 0.042$) at the postchlorination site. Further, the 95 percent confidence limits on the difference (DPD - FACTS) for the postchlorination site were from near 0.0 to 0.13 mg L, indicating a marginal difference in the method means. This analysis supports the hypothesis that the departures

of the points from the 45° line in Figure 1 are due to chance sample deviations of the means rather than to a systematic difference between methods.

Amperometric titration and FACTS

Single-analyst precision results. Estimates of single-analyst precision within each location for the comparison of amperometric titration and FACTS are listed in Table 6. In three cases (locations 12, 14, and 15), a difference in precision was detected at the 99 percent level of confidence. At location 12 the FACTS test had a single-analyst precision twice that of the amperometric titration. At location 14 precision with the FACTS test was four times higher than that with the amperometric titration. The greatest difference was observed at location 15, with the single-analyst precision of the FACTS procedure being five times that of amperometric titration. In all cases the magnitude of precision was ≤10.9 percent of the mean, and the absolute magnitudes of single-analyst precision were all ≤0.15 mg Cl_2 L.

Equivalency test results. The ANOVA F-values and estimated error mean squares [× 10^3] for the comparison of the amperometric titration and FACTS test procedures are reported in Table 7. At the 0.05 level of significance, a method effect was detected in all but two cases (locations 11 and 16). Several interactions of method with other factors were detected, but, in general, the data show few interactions, a pattern similar to that found in the comparison of the DPD and FACTS test procedures. Of the four locations at which the procedures were tested at two sites, three showed significant site effects. In only one of these cases was an interaction of method × site detected (location 13). Table 8 shows that at this location, there was nearly a fourfold difference in the method differences between the two sites. Thus, no clear evaluation of methods was possible for this location.

At locations 15 and 16 variation between analysts and interactions of method × analyst were found. For location 15, the method effect was entirely due to the second analyst (Table 9). Thus, the method effects could not be evaluated. An interaction of method × analyst was also found at location 16. In neither case, however, was the difference due to methods significant at the 0.05 significance level.

Thus, a method difference was detected for four of six locations. Results at two of these locations have ambiguities related either to site or to analyst interactions with method. For neither of the remaining two locations (locations 12 and 14) at which method effects were clearly indicated did the upper or lower 95 percent confidence bound on the difference exceed 0.31 mg L in absolute value (Table 8).

To illustrate the spread of these differences, a graph of amperometric titration versus FACTS mean values is shown in Figure 2. Throughout the range of 0.55 to 2.7 mg Cl_2/L it appears that the results obtained with the FACTS procedure were slightly higher than those obtained by amperometric titration; i.e., the points fall below the 45° line. The significance of this observation was tested by using a paired t-test for each individual pair of mean values at each site. The results in Table 10 indicate no statistical difference between the methods at the prechlorination site ($p = 0.21$) or at the postchlorination site ($p = 0.12$). The 95 percent confidence limits on the differences for the prechlorination and postchlorination sites include zero; therefore, no statistically significant difference was detected. As was the case for the comparison of the DPD and FACTS, this analysis supports the hypothesis that the departures of the points from the 45° line in Figure 2 are due to chance sample deviations of the means rather than to a systematic difference between methods.

TABLE 6
Single-analyst precision for the amperometric titration and FACTS test procedures

Location	Single-Analyst Precision mg L* AMP*	FACTS	Single-Analyst Precision As Percent of Prechlorination Mean AMP	FACTS	As Percent of Postchlorination Mean AMP	FACTS	F_{max} Test of Variances
11	0.06	0.05	4.1	3.4	3.4	2.7	1.09
12	0.03	0.06			2.8	4.9	2.66‡
13	0.04	0.06	1.9	2.2	5.1	6.3	2.23
14	0.02	0.08	1.5	5.0	1.3	4.6	10.42‡
15	0.03	0.15			1.8	9.7	28.08‡
16	0.06	0.06	10.9	10.7	3.0	3.0	1.08

*Based on 36 degrees of freedom after adjusting for between-sample variation for each method
*Amperometric titration
‡$p = 0.01$

TABLE 7
ANOVA F-values for experimental factors in the comparison of the amperometric titration and FACTS test procedures

Location	Degrees of Freedom	Analyst	Site	Method	Analyst × Site	Method × Analyst	Method × Site	Method × Site × Analyst	Error Mean Square × 10^3
11	1, 58	0.06	11.76*	1.17	6.74*	2.28	0.22	2.79	29.08
12	1, 58	0.20		53.06*		0.00			13.73
13	1, 58	0.25	617.7*	214.6*	0.20	0.02	66.66*	0.00	21.72
14	1, 58	0.26	2.32	101.4*	0.40	0.03	0.78	2.48	15.54
15	1, 58	4.24*		6.79*		6.47*			8.11
16	1, 58	5.66*	388.1*	3.00	0.00	5.63*	1.89	1.57	5.84

*$p < 0.05$

TABLE 8
Paired mean differences and 95 percent confidence intervals for the comparison of the amperometric titration and FACTS test procedures

Location	Prechlorination Site Sample Size	Means AMP*	Means FACTS	Difference (AMP - FACTS)	95 Percent Confidence Interval		Postchlorination Site Sample Size	Means AMP*	Means FACTS	Difference (AMP - FACTS)	95 Percent Confidence Interval	
11	30	1.46	1.46	-0.02	-0.11	0.07	30	1.77	1.82	-0.05	-0.13	0.04
12							60	1.07	1.22	-0.15	-0.29	-0.01
13	30	2.08	2.89	-0.81	-0.88	-0.54	30	0.79	0.96	-0.17	-0.25	-0.10
14	30	1.38	1.61	-0.25	-0.31	-0.19	30	1.54	1.75	-0.21	-0.27	-0.15
15							60	1.64	1.54	0.10	0.01	0.20
16	30	0.55	0.56	-0.01	-0.04	0.03	30	1.97	2.01	-0.04	-0.08	0.00

*Amperometric titration

Amperometric titration, DPD, and FACTS

At location 6, measurements were obtained with the DPD and FACTS test procedures, and, in addition, a single measurement of each sample was obtained by using amperometric titration. Single-analyst precision estimates were not possible because multiple observations were not conducted for the amperometric titration method. The ANOVA F-values for the three-way comparison are summarized in Table 11. The only effect detected was a difference in methods. The observed means and paired mean differences are shown in Table 12.

There was no difference detected between the DPD and FACTS test procedures. The amperometric titration differed from both DPD and FACTS test procedures at the 0.05 level of significance. The

TABLE 9
Analyst and method effects at locations 15 and 16

Analyst	Location 15 Means*		Location 16 Means*	
	AMP*	FACTS	AMP	FACTS
1	1.65	1.64	1.15	1.20
2	1.63	1.63	1.37	1.36

*n = 30
*Amperometric titration

TABLE 10
Paired t-test results for the comparison of amperometric titration and FACTS means at each location by site

Site	Means		Difference (AMP - FACTS)	Number of Pairs	Standard Error of Difference	95 Percent Confidence Limits on Difference	Paired t-Test	p Value
	AMP*	FACTS						
Prechlorination	1.38	1.58	-0.22	4	0.14	-0.67, 0.22	1.58	0.21
Postchlorination	1.46	1.55	-0.08	6	0.05	-0.20, 0.03	1.87	0.12

*Amperometric titration

TABLE 11
ANOVA F-values for experimental factors in the comparison of the amperometric titration, DPD, and FACTS test procedures

Factors	Degrees of Freedom	F-Value
Analyst	1	0.83
Site	1	0.07
Analyst * site	1	0.05
Methods	2	36.87*
Methods * analyst	2	0.2*
Methods * site	2	0.49
Methods * site * analyst	2	0.23
Error mean square*	112	11.58

*$F_{.05}$

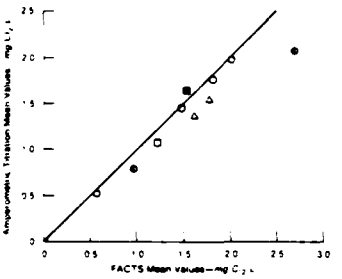
TABLE 12
Paired mean difference and 95 percent confidence intervals for the three-way comparison of amperometric titration, DPD, and FACTS

Sites	Sample Size	Means AMP*	DPD	FACTS
Prechlorination	30	2.13	2.23	2.28
Postchlorination	30	2.13	2.23	2.25
Overall	60	2.13	2.23	2.27

Overall Comparison

Comparison	Mean Difference	95 Percent Confidence Limit	p-Value
DPD-AMP	0.09*	(0.058, 0.135)	0.000;
DPD-FACTS	-0.035	(-0.074, 0.004)	0.078
FACTS-AMP	0.132	(0.093, 0.171)	0.000;

*Amperometric titration



Location	Symbol	Mean Values Amperometric Titration	FACTS
11	○	1.46	1.48
		1.**	1.82
12	□	1.07	1.22
13	●	2.08	2.69
		0.79	0.86
14	△	1.36	1.61
		1.54	1.75
15	■	1.64	1.54
16	○	0.55	0.56
		1.9*	2.01

Figure 2. Comparison of means for free available chlorine obtained with FACTS and amperometric titration at six water treatment plants

TABLE 13
Single-analyst precision as percent of the mean at the postchlorination site

Test Procedure	Range of Single-Analyst Precision	Number of Water Treatment Plants
Amperometric titration	1.3-5.1	6
DPD	2.0-14.6	10
FACTS	0.0-14.0	16

absolute magnitude of the differences was ≤ 0.17 mg Cl_2 L in both cases.

These data allowed a three-way comparison of the amperometric titration, DPD and FACTS test procedures. However, because this comparison was tested at only one location without replicate measurements, no generalized conclusions could be drawn from this analysis.

Conclusions

1. The FACTS and DPD test procedures for measuring free chlorine are equivalent.
2. The FACTS and amperometric test procedures for measuring free chlorine are equivalent.
3. Single-analyst precision, expressed as a percent of the mean at the postchlorination site for each test procedure, is summarized in Table 13.

Acknowledgments

This research was supported by the U.S. Army Medical Research and Development Command, Ft. Detrick, Md., under contract DAMD 17-80-C-0164 and by the Drinking Water Research Center, Florida International University, Miami, Fla. The opinions contained in this article are those of the authors and are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

References

1. *Standard Methods for the Examination of Water and Wastewater*. APHA, AWWA, and WPCF. Washington, D.C. (14th ed., 1976).
2. *Annual Book of ASTM Standards, Part 31—Water*. ASTM, Philadelphia, Pa. (1980).
3. Guidelines Establishing Test Procedures for the Analysis of Pollutants. *Fed. Reg.*, 41, 232, 52780 (Dec. 1, 1976).
4. Gibbs, P.H., Cooper, W.J., & Ott, E.M. A Statistical Experimental Design for Comparison Testing of Analytical Procedures. *Jour. AWWA*, 75, 11, 578 (Nov. 1983).
5. Winer, B.J. *Statistical Principles in Experimental Design*. McGraw-Hill Book Co., New York (1962).
6. Youden, W.J. & Steiner, E.H. *Statistical Manual of the Association of Official Analytical Chemists*. AOAC, Arlington, Va. (1975).
7. Kimm, V.J. Nationwide Approval of Alternate Test Procedure for Free Chlorine Residual. Memorandum to USEPA Regional Administrators (Sept. 7, 1982).

William J. Cooper is an associate professor, Drinking Water Research Center, Florida International University, Miami, FL 33199. Paul H. Gibbs is a statistician, US Army Medical Bioengineering Research and Development Laboratory, Ft. Detrick, MD 21701. Edward M. Ott is a systems analyst, Gould, Inc., SEL Computer Systems Division, Ft. Lauderdale, FL 33313. Perin Patel is a computer programmer, Academic Computer Services, Florida International University, Miami, FL 33199.

APPENDIX D

Nationwide Approval of Alternate Test Procedure for Free Chlorine Residual

FACTS

(Free Available Chlorine Test with Syringaldazine)

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

DATE SEP 7 1982

SUBJECT Nationwide Approval of Alternate Test Procedure for Free Chlorine Residual

FROM Victor J. Kimm, Director
Office of Drinking Water (WH-550)

TO Regional Administrators

Listed below is an alternate test procedure for determining free chlorine residual by a colorimetric method which I have approved for nationwide use for "National Interim Primary Drinking Water Regulation" (NIPDWR) compliance monitoring.

The principle of this method is the oxidation of syringaldazine (3,5-dimethoxy-4-hydroxy-benzaldazine) by free available chlorine on a 1:1 molar basis to produce a colored complex which is then measured at 530 nm. The complete method write-up is provided in the 15th edition of "Standard Methods for the Examination of Water and Wastewater", pages 298-301. A report entitled, "Equivalency Testing of the Free Available Chlorine Test with Syringaldazine, FACTS" indicates that this method is comparable to the USEPA approved DPD method for NIPDWR compliance monitoring.

Measurement	Method
Free Chlorine Residual	Method 408G - "Syringaldazine (FACTS) Method"

cc: Robert L. Booth, Acting Director, EMSL
Regional Water Supply Representatives
Regional Quality Assurance Officers

pm

SEP 1 3 1982
052-0014
58238 2

EPA Form 1320-6 (Rev. 3-76)

DISTRIBUTION LIST

No. of Copies	
5	US Army Medical Research and Development Command ATTN: SGRD-RMS Fort Detrick, Frederick, MD 21701
12	Defense Technical Information Center (DTIC) ATTN: DTIC-DDA Cameron Station Alexandria, VA 22314
1	Commandant Academy of Health Sciences, US Army ATTN: HSHA-DCD Fort Sam Houston, TX 78234
2	Library Technician US Army Medical Bioengineering Research and Development Laboratory ATTN: SGRD-UBZ-T Fort Detrick, Frederick, MD 21701
1	Chief USAEHA Regional Division, West Fitzsimmons AMC Aurora, CO 80045
1	Chief USAEHA Regional Division, North Fort George G. Meade, MD 20755
1	Chief USAEHA Regional Division, South Bldg 180 Fort McPherson, GA 30330
1	Commander USA Health Services Command ATTN: HSPA-P Ft. Sam Houston, TX 78234
1	Commander US Army Mobility Equipment Research & Development Command ATTN: DRDME-GS
1	Commander USA Const. Engr. Research Laboratory ATTN: Dr. Ed Smith/ENS P.O. Box 4005 Champaign, IL 61820

1 Commandant
Academy of Health Sciences, US Army
ATTN: HSHA-IPM
Ft. Sam Houston, TX 78234

1 Commander
US Army Material Development & Readiness Command
ATTN: DRCSG
5001 Eisenhower Avenue
Alexandria, VA 22333

1 HQDA (DASG-PSP-E)
WASH DC 20310

1 Commander
US Army Environmental Hygiene Agency
ATTN: HSHB-E
Aberdeen Proving Ground, MD 21010

1 Commander
US Army Environmental Hygiene Agency
ATTN: HSHB-L
Aberdeen Proving Ground, MD 21010

1 Commander
US Army Environmental Hygiene Agency
ATTN: HSHB-O
Aberdeen Proving Ground, MD 21010

1 Commander
US Army Environmental Hygiene Agency
ATTN: HSHB-OI
Aberdeen Proving Ground, MD 21010

1 Commander
US Army Environmental Hygiene Agency
ATTN: HSHB-R
Aberdeen Proving Ground, MD 21010

1 Commander
US Army Environmental Hygiene Agency
ATTN: HSHB-RE
Aberdeen Proving Ground, MD 21010

1 Commander
US Army Environmental Hygiene Agency
ATTN: Librarian, HSHD-AD-L
Aberdeen Proving Ground, MD 21010

1 Commander
US Army Natick Research and Development Command
ATTN: DRDNA-YE/Mr. Leo A. Spano
Natick, MA 01760

1 Commander
Chemical Research and Development Center
US Army Armament, Munitions and Chemical Command
ATTN: DRSMC-CLB(A)
Aberdeen Proving Ground, MD 21010

1 Commander
10th Medical Laboratory
ATTN: AEMML-PM
APO New York, 09180

1 Commander
US Army Medical Command-Korea
ATTN: EAMC-H-L
APO SF 96301

1 Commander
7th Medical Command
ATTN: AEMPS-PM
APO New York 09102

1 Dean
School of Medicine
Uniformed Services University of the Health Sciences
4301 Jones Bridge Road
Bethesda, MD 20014

2 Commander
U.S. Army Medical Bioengineering Research and
Development Laboratory
ATTN: SGRD-UBZ-C
Fort Detrick, Frederick, MD 21701-5010

END

DTIC

8-86