J. Hyg., Camb. (1985), **94**, 193-200
Printed in Great Britain

Implications of coliform variability in the assessment of the sanitary quality of recreational waters

BY JAY M. FLEISHER

Department of Preventive Medicine and Community Health, Downstate Medical Center 450 *Clarkson Avenue, Box* 43, *Brooklyn, New York* 11203, *U.S.A.*

(*Received* 23 *February* 1984; *accepted* 20 *August* 1984)

SUMMARY

The most widely used indicator of the sanitary quality of recreational waters is the coliform group of bacteria. Present techniques of coliform enumeration are imprecise, and this fact is too often overlooked in routine water quality surveys as well as in research efforts seeking quantitative relationships between coliform density and the health effects of recreational waters. To illustrate this point, three years of data gathered by the New York City Department of Health as part of their routine beach water sampling programme were re-analysed, taking the limited precision of each coliform estimate into account. Re-analysis showed 56·6 % of the data were not significantly different ($P > 0.05$) from the standard being used. This large percentage of the data was of little value in determining the acceptability of the waters being sampled relative to the standard being used and thus represented a substantial waste of time and expense. Of the remaining data, half indicated acceptable water quality and half indicated unacceptable water quality relative to the standard. These three years of data, therefore, gave little information on the acceptability of the water quality at this location with respect to the standard being used. The data further suggest significant differences in coliform density within sample dates. It is recommended that in future water quality surveys, or in studies of the health effects of recreational waters as related to coliform density, emphasis should shift from maximizing the number of sample dates to maximizing the number of replicate determinations made per sample date.

INTRODUCTION

Recreational water quality assessment utilizes the coliform group of bacteria as an indicator organism to establish the sanitary quality of a particular body of water. The relationship between coliform density and increased risk of disease to the public health remains unclear (e.g. Stevenson, 1953; Moore, 1959; Henderson, 1968; Geldreich, 1972; Smith, Twedt & Flannigan, 1973; Carney, Carty & Colwell, 1975; Cabelli, 1979; Gerba *et al.* 1980; Brenniman, Rosenberg & Northrop, 1981). This has led to large variations in water quality standards governing recreational waters where primary contact takes place (Foster, Hanes & Lord, 1971). Regardless of the number of coliform organisms that comprise a particular standard, or the validity of the particular standard with respect to the public health, it becomes the task of local health departments or other governmental agencies to monitor the recreational waters under their jurisdiction. In this regard the limited precision associated with current techniques of coliform enumeration has largely been ignored. The failure to account for this limited precision is also true for those engaged in research activities seeking to establish quantitative relationships between coliform density and health risk. Unless particular attention is paid to this lack of precision, data generated in research efforts or routine water quality surveys can be inadequate and even misleading. This paper will show that unless strict attention is paid to the limited precision inherent in current techniques of coliform enumeration, information obtained in sanitary water quality surveys or in research studies seeking to establish a quantitative relationship between coliform densities and health can be severely limited.

METHODS

Water quality data for the years 1979 to 1981 were obtained for a 'marginally polluted' beach located within the city of New York (New York City Department of Health, 1982). The term marginally polluted reflects the fact that recorded coliform densities usually stayed within the same order of magnitude as the standard being used to determine suitability of these bathing waters with respect to public access. Taking into account the limited precision inherent in current methods of coliform enumeration, the sanitary quality of such waters becomes difficult to assess. The sampling scheme followed by the New York City Health Department in determining the sanitary quality of its bathing waters is as follows. One sample of water is taken each hour for a period of seven hours on each sample date for each beach sampled. Each sample is iced and transported back to the laboratory where estimates of coliform density are obtained using the MPN procedures (APHA Standard Methods, 1971). The particular MPN procedure used is five tubes in each of three decimal dilutions. The $\log_{10}$ average of the seven determinations is then calculated and reported as the coliform density for that sample date. Sampling is conducted several times each month during the months of May to September.

New York City used the following criteria to determine suitability for public access to its beaches.

(*a*) A total coliform logarithmic average of less than 2400 organisms per 100 ml of sample. These averages are computed using the seven determinations taken on a particular sample date.

(*b*) Not more than 20% of the total samples taken may exceed 5000 organisms per 100 ml of sample nor more than 50% of the total samples exceed 2400 organisms per 100 ml of sample.

Classification is based upon an evaluation of at least three logarithmic averages from the most current sampling dates or from an evaluation of the logarithmic average of a series of five or more individual samples collected within any 30-day period. The more restrictive data are used in determining the coliform level at any particular location.

Analysis

A single-classification ANOVA procedure was used to determine whether the mean water quality was significantly different over the years 1979 to 1981 (Sokal & Rohlf, 1969). Prior to analysis the data were transformed ($\log_{10}$) and tested for normality. The 95 % confidence interval was then calculated for each sample date using the standard deviation of the $\log_{10}$ MPN as reported by Velz (1970). The standard of 2400 coliform organisms per 100 ml of sample was then compared to the 95 % confidence interval calculated for each sample date. This was done to ascertain whether the standard fell within the confidence interval generated for that particular sample date. If this occurred it would not be possible to determine whether coliform density on that particular sample date differed significantly from the standard. The percentage of the total sample dates was then calculated for each of the following situations: the number of sample dates in which the estimated coliform density was significantly less than the standard, the number of sample dates in which the estimated coliform density was significantly greater than the standard, the number of sample dates in which the standard fell within the 95 % confidence interval of the coliform estimate and thus the number of sample dates in which the coliform estimate was not significantly different from the standard. This procedure was done for each year individually and for a composite comprising all three years of data. Since sampling was conducted such that samples were collected hourly from 07.00 h to 13.00 h on each sample date, a single classification ANOVA procedure using time of day as the main effect was carried out on the composite of all three years of data. Significant differences among these means would indicate differences in coliform density by time of day. The Student–Newman–Keuls (SNK) test (Sokal & Rohlf, 1969) was then applied in an attempt to identify possible trends in coliform density within sample days. In order to assess the variability of coliform densities with respect to the theoretical variance within a particular sampling date, the variance of each sample date was compared to the theoretical variance of the MPN procedure. This was done by calculating the shortest unbiased confidence interval about the variance for each sample date (Tate & Klett, 1959). The theoretical variance was then compared to each confidence interval to ascertain significant differences.

RESULTS

Table 1 shows the results of the 95 % confidence interval comparisons for the years 1979 to 1981. Inspection of Table 1 indicates that 33·3 % of the samples taken in 1979 failed to show a significant difference between the particular coliform estimate and the standard. Similarly, 64·3 % of the sample taken in 1980 and 70·6 % of the samples taken in 1981 did not differ significantly from the standard at $P > 0{\cdot}05$. As shown in Table 1, the overall mean coliform levels for each year were 3081, 1767 and 2796 respectively for the years 1979 to 1981. These means, however, incorporate within them this large proportion of the data that were not significantly different from the standard. The use of the mean as an overall indicator of water quality is therefore questionable. Assuming that these data are representative of the entire bathing season, acceptable water quality would be

https://doi.org/10.1017/S0022172400061398 Published online by Cambridge University Press

Table 1. *Comparisons of* 95 % *confidence intervals* 1979–1981.

(Data shown as percentage of total samples taken for that year. Numbers in parentheses indicate actual number of sample dates.)

Year	*Acceptable water quality	†Questionable water quality	‡Unacceptable water quality	§Mean water quality
1979	20 % (3)	33·3 % (5)	46·7 % (7)	3081/100 ml
1980	28·6 % (4)	64·3 % (9)	7·1 % (1)	1767/100 ml
1981	17·6 % (3)	70·6 % (12)	11·8 % (2)	2796/100 ml

* Acceptable = coliform densities significantly less than the standard.
† Questionable = coliform densities not significantly different than the standard.
‡ Unacceptable = coliform densities significantly greater than the standard.
§ Mean densities based on all data taken for that year.

Table 2. *Single-classification ANOVA of logarithmic means of coliform density using year in which data were taken as main effect*

(Each mean represents seven replicate determinations made on each sample date. Data transformed ($\log_{10}$) prior to analysis. See text for details of sampling procedure.)

Source	DF	SS	MS	F
Year	2	0·4817	0·2408	2·345 (n.s.)
Error	44	4·4157	0·1027	

n.s., $P > 0{\cdot}05$.

expected 20 % of the time while substandard water quality would be expected 46·7 % of the time for the 1979 bathing season. Similarly, one would expect acceptable water quality 28·6 % and substandard water quality only 7·1 % of the time for the year 1980. For the year 1981 one would expect acceptable water quality 17·6 % of the time and substandard water quality 11·8 % of the time.

The results of the single classification ANOVA procedure using the year in which the particular sample was taken as the main effect are shown in Table 2. These results show no significant difference in water quality over the years 1979 to 1981 at $P > 0{\cdot}05$. Since no significant difference in water quality was found the data were pooled, and the results of the 95 % confidence interval comparisons on the pooled data are shown in Table 3. Inspection of Table 3 shows that 56·6 % of the data taken over the three-year period fail to show coliform densities significantly different from the standard being used. Assuming that the data taken are representative of water quality of the entire three years under study, a person using this beach could expect acceptable water quality 21·7 % of the time and unacceptable water quality 21·7 % of the time.

The results of the single-classification ANOVA procedure using time of day as the main effect are shown in Table 4. Inspection of Table 4 indicates significant differences in coliform density by time of day. The results of the SNK procedure are shown in Table 5. Inspection of Table 5 indicates that the significant difference shown by the ANOVA procedure was due to high coliform density in the first 2 h of sampling and significantly lower coliform levels during the last 2 h of sampling.

Table 3. *Comparisons of 95 % confidence intervals*

(Data shown as percentage of all samples taken over three year period. Numbers in parentheses indicate actual number of sample dates.)

*Acceptable water quality	†Questionable water quality	‡Unacceptable water quality	§Mean water quality
21·7 % (10)	56·6 % (26)	21·7 % (10)	2510/100 ml

* Acceptable = coliform densities significantly less than the standard.
† Questionable = coliform densities not significantly different from the standard.
‡ Unacceptable = coliform densities significantly greater than the standard.
§ Mean density based on all data taken.

Table 4. *Single-classification ANOVA procedure using time of day at which individual sample was taken as main effect*

(Data transformed ($\log_{10}$) prior to analysis. See text for details of sampling procedure.)

Source	d.f.	SS	MS	F
Time of day	6	3·5824	0·5971	3·07 ($P < 0·01$)
Error	305	59·3074	0·1944	

Table 5. *Results of SNK procedure on means comprising single-classification ANOVA procedure using time individual sample was taken as main effect* (*Table* 4)

(Means included in the overlap of the two lines shown are not significantly different ($P > 0·05$). Means are arranged by descending order of magnitude. Data transformed ($\log_{10}$) prior to analysis. See text for further details.)

Mean	N	Time of day
3·5208	44	08·00
3·4557	44	07.00
3·4268	45	10.00
3·4172	44	09.00
3·2702	45	11.00
3·2484	45	13.00
3·2280	45	12.00

The results of the comparisons of the theoretical variance with the confidence intervals about the variance calculated for each sample date indicate the following: 6·7 % of the sample dates showed variances significantly lower than the theoretical variance; 60 % of the sample dates were not significantly different from the theoretical variance; 33·3 % of the sample dates contained variances significantly higher than the theoretical variance of the MPN procedure utilized.

DISCUSSION

The limited precision of the technique used to enumerate coliform density coupled with the problems inherent in the small sample sizes used cast doubt on the appropriateness of the criteria used by the City of New York to determine water

https://doi.org/10.1017/S0022172400061398 Published online by Cambridge University Press

quality at this location. Re-analysis of the original data taking the limited precision of the coliform estimate into account shows that no definitive conclusion can be reached with respect to the acceptability of the water quality at this location. More than half (56·5 %) of the data taken over the three-year period failed to show coliform densities significantly different from the standard and thus could not be used to ascertain water quality with respect to the standard. Of the remaining data, one-half (21·7 %) showed coliform densities significantly greater than the standard while one-half (21·7 %) showed coliform densities significantly lower than the standard. Assuming these data to be representative of water quality at this location, all that can properly be surmised from the three years of data is that on the average a person using these waters has an equal chance of encountering acceptable or unacceptable water quality. Moreover, more than half of all the data taken over the three-year period could not be used to determine water quality at this location due to the limited precision of the enumeration technique used.

Analysis of the coliform densities at this location reveals significant differences in mean density by time of day. The data show a significantly higher ($P < 0·05$) mean density in the early morning hours, the range being from 3317 to 1690 organisms per 100 ml of sample. Calculation of the 95 % confidence interval about these means shows early morning densities significantly greater than the standard of 2400 organisms, while afternoon densities appear to be significantly lower than the standard. These differences are likely due to the effect of tides, wind, currents or other environmental factors at the location sampled. These differences strongly argue for a sampling strategy that includes a greater number of replicate determinations per sample date.

Comparison of the theoretical variance with the variances obtained for each sample date shows evidence of agreement. Caution must be used in interpreting these results owing to the small sample sizes upon which these calculations are based. The fact that only 6·7 % of the samples showed a variance significantly lower than the theoretical justifies the use of the theoretical standard deviation in the calculation of confidence intervals. Use of the theoretical standard deviation on the 33·3 % of the sample dates that showed a significantly higher variance would comprise a conservative approach to the estimation of coliform density. If the actual variances obtained did not differ significantly from the theoretical variance this would give evidence of uniform variability of coliform density at this particular beach. Although 60 % of the calculated variances were not significantly different from the theoretical the small sample sizes used in these calculations prohibit any definite conclusions. Further replication will be required in any attempt to relate coliform variability to environmental factors at this beach. If variability with respect to environmental factors is large than replication within sample days would need to be emphasized. If it is shown that variability with respect to environmental factors is not significant further replication will still be required due to the differences in mean coliform densities with respect to time of day as well as the limited precision inherent in the method of coliform enumeration.

Although the data presented in this paper were taken as part of a routine sanitary water quality survey, the lack of attention paid to the limited precision inherent in techniques of coliform enumeration as well as subsequent limitations of the data derived from such sampling regimes apply equally to the research situation. There is much discussion in the literature regarding the validity of the sanitary water quality standards in use today. Studies which attempt to establish quantitative relationships between coliform densities and increased risk of disease must be based on both sound epidemiological principles and reliable coliform estimates. The limited precision inherent in current techniques of coliform enumeration has not received enough emphasis in the experimental design of such studies. The sampling regimes used in such studies tend to maximize the number of sample dates taken. The same is true for routine water quality surveys. This practice too often does not adequately address the limited precision of the resulting coliform estimates. This in turn contributes to the confusing and often contradictory literature dealing with the use of the coliform organism as an indicator of recreational water quality. The only way to improve the precision of coliform estimates is through the use of replicate determinations. In fact the confidence interval about the estimate will be reduced by a factor related to the square root of the number of replicates used. The confidence interval can be used as an estimate of the precision of the method of enumeration, and as such the precision can be effectively increased through replicate determinations. Once a baseline study of coliform density has been completed emphasis should shift from maximizing the number of sample dates taken to maximizing the number of replicate determinations taken per sample date. This would ensure more precise estimates of coliform density. This principle applies equally to the assessment of marginally polluted waters as well as in research studies seeking to establish a quantitative relationship between coliform densities and increased risk of disease to those using marginally polluted waters.

Since the introduction of the coliform organism as an indicator of sanitary water quality in the early 1900s, there has been little improvement in the experimental design used in routine water quality assessment or in research situations. There remains in the literature a large quantity of inaccurate, contradictory, and sometimes confusing studies that all have a common error; the failure to deal adequately with the limited precision inherent in techniques of coliform enumeration. The coliform organism, with all of its associated problems, still remains an acceptable indicator of sanitary water quality. Through the use of proper experimental design, the effectiveness of the use of coliform bacteria as an indicator organism will be vastly improved.

REFERENCES

American Public Health Association (1971). *Standard Methods for the Examination of Water and Wastewater*. Washington, D.C.: American Public Health Association.

Brenniman, G. R., Rosenberg, S. H. & Northrop, R. L. (1981). Microbial sampling variables and recreational water quality standards. *American Journal of Public Health* **71**, 283–289.

Cabelli, V. J. (1979). Relationship of microbial indicators to health effects at marine bathing beaches. *American Journal of Public Health* **69**, 690–696.

Carney, J. F. Carty, C. E. & Colwell, R. R. (1975). Seasonal occurrence and distribution of microbial indicators and pathogens in the Rhode River of Chesapeake Bay. *Applied Microbiology* **30**, 771–780.

Foster, D. H., Hanes, N. B. & Lord, S. M. (1971). A critical examination of bathing water quality standards. *Journal of Water Pollution Control Federation* **43**, 2229.

Geldreich, E. E. (1972). Buffalo Lake recreational water quality: a study in bacteriological data interpretation. *Water Research* **6**, 913–924.

GERBA, C. P., LABELLE, R. L., GOYAL, S. M., MELNICK, S. M., CECH, I. & BOGDAN, G. F. (1980). Relationships between environmental factors, bacterial indicators, and the occurrence of enteric viruses in estuarine sediments. *Applied and Environmental Microbiology* **39**, 588–596.

HENDERSON, J. J. (1968). Enteric disease criteria for recreational waters. *Journal of the Sanitary Engineering Division* **94**, 1253–1275.

MOORE, B. (1959). Sewage contamination of coastal bathing waters in England and Wales. A bacteriological and epidemiological study. *Journal of Hygiene* **57**, 435–472.

NEW YORK CITY DEPARTMENT OF HEALTH (1982). *Beach and Harbor Water Sampling Program 1979–1981*. New York.

SMITH, R. J., TWEDT & FLANNIGAN, L. K. (1973). Relationships of indicator and pathogenic bacteria in stream water. *Journal of Water Pollution Control Federation* **45**, 1736–1745.

SOKAL, R. R. & ROHLF, F. J. (1969). *Biometry*. San Francisco: Freeman.

STEVENSON, A. J. (1953). Studies of bathing water quality and health. *American Journal of Public Health* **43**, 529–538.

TATE, R. F. & KLETT, F. W. (1959). Optimal confidence intervals for the variance of a normal distribution. *Journal of the American Statistical Association* **54**, 674–682.

VELZ, C. J. (1970). *Applied Steam Sanitation*. New York: Wiley.