Revista Visa em debate
sociedade, ciência & tecnologia

EXPERIENCE REPORT
https://doi.org/10.22239/2317-269x.01297

Data quality of water fluoridation surveillance: proposal for a checking protocol

Qualidade dos dados de vigilância da fluoretação de sistemas de abastecimento de água: proposta de um protocolo de crítica dos dados

Isaque Mendes Prado

Paulo Frazão*[I]

ABSTRACT

Introduction: Fluoride concentration values in water must be inserted in an information system; however, no protocol was proposed to perform the verification of the inserted data. **Objective:** To present a proposal to critically review fluoride concentration data in order to provide reliability of information production on water fluoridation surveillance. **Method:** The fluoride concentration monitoring data from the water supply systems of the State of São Paulo registered in the year 2015 in the Surveillance Information System of the Water Quality for Human Consumption were used. A proposal was applied using Excel® software. **Results:** The database was reduced from 23,840 registers of fluoride concentration distributed in 586 municipalities to 22,807 distributed in 543 municipalities. Notation errors, underreporting during the year, municipalities with 50% or more of the samples equal to 0.000 mg F/L or with standard deviation equal to 0.000 mg F/L, and outliers were the main critical factors. **Conclusions:** The application was simple, feasible and could be routinely adopted by the surveillance agencies.

KEYWORDS: Public Health; Surveillance; Drinking Water; Fluoridation; Data Quality

RESUMO

Introdução: Valores de concentração do fluoreto na água devem ser inseridos em um sistema de informação, entretanto nenhum protocolo foi proposto para efetuar a verificação dos dados inseridos. **Objetivo:** Apresentar uma proposta de crítica dos dados de concentração do fluoreto a fim de propiciar confiabilidade na produção de informações sobre a vigilância da fluoretação da água de abastecimento público. **Método:** Foram utilizados os dados de vigilância da concentração de fluoreto dos sistemas de abastecimento de água do estado de São Paulo registrados no ano de 2015 no Sistema de Informação de Vigilância da Qualidade da Água de Consumo Humano. Uma proposta foi testada empregando-se recursos do aplicativo Excel®. **Resultados:** A base de dados foi reduzida de 23.840 registros de concentração de fluoreto distribuídos em 586 municípios para 22.807 distribuídos em 543 municípios. Erros de notação, subalimentação durante o ano, municípios com 50% ou mais das amostras iguais a 0,000 mg F/L ou com desvio-padrão igual a 0,000 mg F/L e amostras com valores anormais foram os principais fatores críticos. **Conclusões:** A aplicação mostrou-se simples e viável, podendo ser adotada rotineiramente pelos órgãos de vigilância.

PALAVRAS-CHAVE: Saúde Pública; Vigilância; Água de Beber; Fluoretação; Qualidade dos Dados

[I] Faculdade de Saúde Pública, Universidade de São Paulo, São Paulo, SP, Brasil

* **E-mail:** pafrazao@usp.br

Received: Apr 4, 2019
Approved: Jun 26, 2019

INTRODUCTION

Public supply water surveillance can be defined as an articulated system of actions that ensure the collection, analysis, interpretation of data and dissemination of information to ensure safety and quality standards that are acceptable to human consumption and in accordance with predetermined health goals[1]. It does not exempt water companies and utilities from the responsibility of carrying out their own operational controls[2]. When this type of monitoring is part of ongoing State action, it is called surveillance[1]. Surveillance can be done by auditing data produced by water utilities, or by direct observation and examination of water samples from the distribution network[2].

In 1992, for the first time in Brazil, the state of São Paulo began to have a series of systematic actions of health surveillance of water quality, which gave rise to the procedures of the Quality Surveillance Program of Water for Human Consumption in the state of São Paulo[3]. Despite having a potability standard since 1977, the quality monitoring of water for human consumption was only implemented in Brazil as a program after the creation of the National System of Environmental Surveillance in Health in 2002[4].

The Information System on Quality Surveillance of Water for Human Consumption (Sisagua) was created under this national program with the objective of producing, analyzing and disseminating data on the quality of water for human consumption according to potability standards, creating conditions for the enforcement of water quality surveillance by municipal and state health departments. It was designed to: systematize record information about the various forms of water supply (public, private, and collective and individual alternative solutions); foster the quality surveillance of water for human consumption by municipal managers, helping them identify risk situations and make decisions about preventive and corrective actions; and share information with public agencies and civil society[5].

Along with other parameters like turbidity, residual chlorine, colimetry, pesticides and mercury, fluoride is one of the parameters used to evaluate water quality. It works as an indicator because adjusting its concentration in the water is a mandatory step for public supply systems in Brazil[6]. The fluoridation of public supply water is part of the National Oral Health Policy, an important cross-sector measure of public health. The monitoring of fluoride concentration in drinking water is usually done to ensure the quality of the contents, aiming at the maximum benefit of preventing dental caries with the minimum risk of fluorosis or stains on tooth enamel. The exceptions are those situations in which the water has naturally-occurring fluoride above the values recommended for the prevention of dental caries. In these cases, the monitoring of fluoride concentration in drinking water is done only to ensure compliance with potability standards, while resources to reduce concentration to recommended levels for the prevention of dental caries are not available.

Considering that the magnitude of dental caries and dental fluorosis resulting from exposure to fluorides in water can be measured only after a few years, Brazilian experts have recommended that fluoridation surveillance be performed by bodies not directly responsible for water treatment (external control principle) through direct evaluation of water samples taken from the distribution network[7,8,9] to ensure process quality, information validity and reliability to achieve oral health goals[10].

In Brazil, the first fluoride content monitoring systems in public supply water were implemented at the municipal level in the late 1980s, supporting the relationship between government and supply companies[7,9]. The responsibility of the municipal health authorities for the surveillance and implementation of their own sampling plan was established in 2000[11], and data on coverage and surveillance of public water fluoridation began to be entered in Sisagua, maintained by the National Program of Surveillance in Environmental Health, related to the quality of water for human consumption[12].

However, no protocol was proposed to check the data entered into the system. Nationwide research with scientific information from Sisagua data reports did not provide a detailed description of the procedures adopted to remove data logging problems[13,14].

Similarly to Brazil, many countries do not have a regular plan to monitor and assess the data quality of their health information systems, and are limited to non-systematic and standalone initiatives. Establishing procedures to detect and remove errors and data inconsistencies is key for improving information quality, which, in turn, is essential for objective analysis of the health situation, evidence-based decision-making, and health action programming[15].

The objective of this report was to present a proposal to critically review water fluoridation surveillance data in order to provide alternatives to ensure consistency in the production of information.

METHOD

Data source

The data came from Sisagua, maintained by the Brazilian Ministry of Health (MS) in cooperation with state and municipal health departments. They were made available by the Environmental Service of the Health Surveillance Center of the São Paulo State Department of Health. The Service is responsible for the surveillance of about 80 potability parameters established by Ordinance MS n. 2.914 of December 12, 2011, which must be monitored by sanitation companies at intervals ranging from daily to semi-annual[16]. The State Program for Quality Surveillance of Water for Human Consumption (Proagua) was created in 1992. It is grounded on an integrated and hierarchical system of health surveillance that ensures actions to control the health risk associated with water in all municipalities of

São Paulo, with regional and central technical support[3]. After the extensive remodeling of Sisagua that ensured more sophisticated resources for data managing and reporting to support risk control strategies, the state of São Paulo took full control of the system, in what can be considered an essential step toward improving water quality surveillance[17].

With about 43 million inhabitants in 2015, the state of São Paulo has 645 municipalities and is the most populous state in Brazil, with almost 22% of the Brazilian population. According to the population census, in 2010, public water supply reached 97.9% of urban permanent private households. Water fluoridation began in the state in 1956, in the municipality of Marília. In 2009, after more than half a century of implementation, coverage comprised 546 (84.7%) of the 645 municipalities of São Paulo, reaching 85.1% of the total population and 93.5% of the population with access to the water distribution network[18].

An Excel file with four columns was obtained with records corresponding to the Municipality; Code of the Brazilian Institute of Geography and Statistics (IBGE); Date of collection; Fluoride concentration (mg F/L), sorted alphabetically by municipality name.

Checking procedures

As described below, five check steps were proposed for each municipality to identify entry errors; frequency of months per year with recorded collection samples; distributions with excess values of fluoride concentration equal to zero; distributions with standard deviation and coefficient of variation values equal to 0.000 mg F/L; and aberrant values. Microsoft Excel® was chosen because we consider it to be a platform widely used in strategic level government agencies, thus enabling the easy sharing of the proposition.

1. Identification of entry errors

The obtained spreadsheet was copied. With the FIND command, all records with dots instead of commas (used in Portuguese to write decimal numbers) were identified and deleted.

2. Identification of the number of months per year with records

a. In this worksheet, a column called "Month" was created with the help of the following function of the application: CAPITAL(TEXT(cell with date of collection;"MMMM")). With this, based on the collection dates, codes were created for the months of collection. For example, if the collection date was 1/04/2015, in the new column the created code was "January".

b. To identify the number of months corresponding to each municipality, the data were transcribed into a new spreadsheet. With the help of the REMOVE DUPLICATES command, the MUNICIPALITY and MONTH columns were selected to delete the repeated months in each municipality. Then, using a pivot table, the months recorded by municipality were counted. The values obtained were arranged in alphabetical order according to the name of the municipality and classified in ascending order to enable the identification of cities with insufficient number of recorded months. Based on previous studies[1,14], only municipalities with more than three months of sample recording were included. Four months or more was considered – at this stage of public policy implementation – as the minimum acceptable value for estimating the annual situation in each municipality.

3. Identification of municipalities with excess zero values

a. After excluding the municipalities with insufficient number of months in the worksheet, a column was created to classify the fluoride concentration values of each sample. For this classification we used the classification intervals proposed by a technical consensus of the Collaborating Center of the Brazilian Ministry of Health for Oral Health Surveillance maintained by the University of São Paulo (CECOL/USP)[19], which takes into account the benefit/risk binomial in relation to fluoride content in water according to the average of the maximum annual temperatures. In the present case, i.e. the state of São Paulo, the proposed ranges were adopted for temperatures between 26.3°C and 32.5°C, as shown in the Chart. For this, the "SE" logical function of Excel was used with the following formula: = SE (cell concentration value = 0.000; "zero"; SE (cell concentration value <0.445; "0.000-0.444"; SE (cell concentration value <0.555; "0.445-0.544"; SE (cell concentration value <0.845; "0.555-0.844"; SE (cell concentration value <1.145; "0.845-1.144"; SE (cell concentration value <1.445; "1.145-1.444";"> = 1.445")))))).

b. By means of a pivot table created in a specific spreadsheet, the total number of samples in each classification interval of fluoride concentration values was determined for each municipality. Using the ACTIONS command, the entire pivot table was selected, copied and a new spreadsheet was created to calculate the proportions in each concentration range. Data from the ZERO column were sorted in ascending order to identify municipalities with 50% or more of null values, assuming that the distribution of values with excess zeros may indicate nonconformities in the technical procedure of collection, laboratory analysis or sample recording. Thus, only municipalities that did not have excess zeros were included.

4. Identification of distributions with null variation

From the worksheet mentioned in items 2.a and 3.a, a specific spreadsheet pivot table was created to obtain the mean value and standard deviation of fluoride concentration of the sample distribution of each municipality. For this, the "Municipality" field was placed in line label and the fluoride concentrations were placed in the Σ values field, which was set to calculate the mean and standard deviation. Using the ACTIONS command, the entire pivot table was selected, copied, and a new worksheet was created to sort the standard deviation values in ascending order and to identify the values equal

Chart. Fluoride concentration ranges in water according to caries prevention benefit and dental fluorosis risk levels at locations with annual maximum temperature means between 26.3°C and 32.5°C. São Paulo, 2011.

Fluoride content in the water (in ppm or mg F/L)	Benefit (preventing caries)	Risk (causing dental fluorosis)
0.00 to 0.44	Insignificant	Insignificant
0.45 to 0.54	Minimum	Low
0.55 to 0.84*	Maximum	Low
0.85 to 1.14	Maximum	Moderate
1.15 to 1.44	Questionable	High
1.45 or more	Harmful	Very high

Best risk-benefit combination occurs within this range.
Source: CECOL/USP[19].

to 0.000 mg F/L that result from a unimodal distribution of the fluoride concentration values and may indicate non-compliance with technical procedures for sample collection, laboratory analysis or recording.

5. Identification of aberrant values

a. Then, in the worksheet of the previous item, the coefficient of variation (CV) of fluoride concentration resulting from dividing the value of the standard deviation by the mean value was calculated. Data from the CV column were sorted in ascending order to identify municipalities with coefficients of variation equal to 50% or more. Data from these municipalities were transcribed from the worksheet to a specific spreadsheet, and 88 municipalities were identified. To explore the presence of aberrant values, the Tukey's test was applied.

b. Tukey's test (Box-plot):

   1. Quartile statistical function to identify the values of Q1 and Q3 (quartile 1 and quartile 3) corresponding to the distribution in each municipality;

   2. Calculation of interquartile range (IQR): Q3-Q1 = IQR;

   3. Calculation of lower (LL) and upper (UL) limits by the following formulas:

$$LI = Q1 - (1{,}5 * IIQ)$$

$$LS = Q3 + (1{,}5 * IIQ)$$

   4. Logical function or (cell_with_concentration<LL; cell_with_concentration> UL);

   5. Identification and disposal of aberrant values.

471 aberrant values were found in 67 municipalities.

c. For municipalities with IQR = 0 (two municipalities) an alternative method corresponding to the modified Z Score was applied:

   1. The median of the concentration values was calculated by the "MEDIAN" Excel function;

   2. Calculation of absolute deviations in a column by the formula: cell with concentration value minus mean;

   3. The arithmetic mean of the absolute deviations is obtained;

   4. In a new column, the modified z values (z*i) were calculated for each cell with a concentration value by the formula:

      z*i = 0.6745 * (absolute deviation/arithmetic mean of absolute deviations); and

   5. Aberrant values were considered to be those with |z*i| > 3.5.

   In both municipalities, 18 records were discarded.

RESULTS

The results from the application of the protocol are shown in the Table. Regarding the number of municipalities, it can be noted that the highest percentages of loss refer to municipalities that did not feed the system (9.1%) and municipalities that fed it less than four months a year (5.8%). Regarding the records, the highest percentage of record loss was in municipalities that had 50% or more of samples with a value equal to 0.000 mg F/L and samples with abnormal values. The adopted criteria are discussed below.

DISCUSSION

This study presented a five-step protocol proposition to critically review the monitoring data of fluoride concentration in public supply water. For this, we used an Excel spreadsheet® and the surveillance database of the state of São Paulo for the year 2015, in which more than 90% of the municipalities had records in the system.

The first step related to entry errors was the removal of values expressed with a dot. The number of errors found was negligible, suggesting that, in this respect, the system has no problem in the state of São Paulo. In the second step, the frequency of months with records during the year was determined. This enabled us

Table. Result of the critical protocol of surveillance data.

	Records		Municipalities	
	N	%	N	%
Total municipalities			645	
Total data recorded	23,840		586	9.1
Value entry errors (dot instead of comma = 5)	23,835	0.0	586	
Municipalities (= 34) with number of months of records (= 153) smaller than four	23,682	0.7	552	5.8
Municipalities (= 7) with 50% or more samples equal to 0.000 mg F/L (= 345)	23,337	1.5	545	1.3
Municipalities (= 2) with standard deviations and coefficients of variation equal to 0.0 (= 41)	23,296	0.2	543	0.4
Municipalities (= 69) with abnormal values (= 489)	22,807	2.1	543	0.0
Total	22,807	4.3	543	7.3

[a] Percentage is a result of absolute values and does not match the sum of column percentages.
Source: Prepared by the authors from the database of the Health Surveillance Center of the São Paulo State Department of Health.

to assess the degree to which the sampling guidelines determined by the national surveillance program were followed by the municipalities. Given the current state of public policy implementation, experts have recommended at least four months of records to estimate the annual situation in each territory[1,14]. The application of this amount led to the loss of 5.8% of municipalities that did not have more than three months of records. By improving the sample collection and system feeding process, it is estimated that the minimum value can be increased to six or more months, ensuring a bigger number of samples to represent the annual situation.

The third checking procedure was the identification of distributions with zero values. Since any water, either surface or groundwater, contains some minerals, including fluoride, a sample of 0.000 mg F/L is very unlikely. In general, water with low fluoride concentration has between 0.01 and 0.24 mg F/L[20,21]. The value of 0.01 mg F/L corresponds to the limit of detection by ion chromatography. The most common measurement techniques such as ion-selective electrode and SPADNS have as detection limits the values of 0.05 mg F/L and 0.02 mg F/L, respectively[22]. Therefore, there is no reason to record values smaller than 0.01 mg F/L. The presence of records below this value is a clear nonconformity situation, and it should prompt the technician and the system administrator to identify possible causes and implement corrective actions. A point to be questioned is that in this protocol proposition only municipalities with 50% or more samples with null values were discarded. Future protocols could evaluate the possibility of discarding all null values in the first step of verification.

Another anomaly concerns the detection of unimodal distributions of fluoride concentrations or with standard deviation and coefficient of variation values of 0.000 mg F/L. This type of distribution certainly does not reflect the distribution of values resulting from the chemical analysis of water samples, performed in accredited laboratories, and is indicative of nonconformities, which can be associated with some abnormality in the technical collection procedure, laboratory analysis or in the step of recording the value of the sample concentration in the information system.

The last procedure we adopted was the detection of aberrant values. These atypical values at the edges of the distributions may hinder the production of estimates of the fluoride concentration level in a given locality and distort the interpretation of the annual situation. Two methods were proposed, according to the interquartile range value. Improvement of the sampling and feeding process should lead to the recording of values according to the measurement technique to at least two decimal places. At this stage, a distribution in which the interquartile range is equal to zero is unlikely, which is why the modified Z-score method tends to have limited application.

This protocol proposition was applied with the help of the most common Excel®commands, from an inventory of about 24,000 records, distributed in 586 territories. Considering an inventory with ten times more records and territories, a situation that does not apply individually to the states of the Brazilian federation, researchers would need tools with more control and editing options for spreadsheets. *Visual Basic for Application* could be used as a user-friendly programming language to enable the creation of macros and automate various processes within spreadsheets and charts prepared in Excel®.

CONCLUSIONS

This protocol was built as part of a scientific initiation program for undergraduate students, and its application proved feasible and could be routinely adopted by state and regional surveillance agencies across Brazil. In addition to encouraging the improvement of water fluoridation surveillance actions, by applying the error and inconsistency detection procedures presented, it is possible to improve the quality of information on fluoride concentration in public supply water. That is an essential condition to subsidize decision-making by health authorities at different management levels.

REFERENCES

1. Venturini CQ, Narvai PC, Manfredini MA, Frazao P. Vigilância e monitoramento de fluoretos em águas de abastecimento público: uma revisão sistemática. Rev Ambient Agua. 2016;11(4):972-88. https://doi.org/10.4136/ambi-agua.1929
2. World Health Organization - WHO. Guidelines for drinking-water quality. 4a ed. Geneva: World Health Organization; 2011.
3. Valentim LSO, Elmec AM, Mario Júnior RJ, Bataiero MO. Novos cenários de produção e de vigilância da qualidade da água para consumo humano: 20 anos de Proagua no estado de São Paulo parte I. Bol Epidemiol Paul. 2012;9(100):29-39.
4. Freitas MB, Freitas CM. A vigilância da qualidade da água para consumo humano: desafios e perspectivas para o sistema único de saúde. Cienc Saude Colet. 2005;10(4):993-1004. https://doi.org/10.1590/S1413-81232005000400022
5. Ministério da Saúde (BR). Manual de procedimentos do sistema de informação da vigilância da qualidade da água para consumo Humano (Sisagua). Brasília: Ministério da Saúde; 2007.
6. Frazão P, Peres M, Cury J. Qualidade da água para consumo humano e concentração de fluoreto. Rev Saude Publ. 2011;45(5):964-73. https://doi.org/10.1590/S0034-89102011005000046
7. Schneider Filho DA, Prado IT, Narvai PC, Barbosa SR. Fluoretação da água: como fazer a vigilância sanitária. Rio de Janeiro: Rede Cedros; 1992.
8. Anais do 14º Encontro Nacional de Administradores e Técnicos do Serviço Público Odontológico; 9-13 ago 1998;Fortaleza, CE. Fortaleza: 1998.
9. Narvai PC. Cárie dentária e flúor: uma relação do século XX. Cienc Saude Colet. 2000;5(2):381-92. https://doi.org/10.1590/S1413-81232000000200011
10. Frazão P, Ely HC, Noro LRA, Pinheiro HHC, Cury JA. O modelo de vigilância da água e a divulgação de indicadores de concentração de fluoreto. Saude Deb. 2018;42(116):274-86. https://doi.org/10.1590/0103-1104201811622
11. Ministério da Saúde (BR). Portaria Nº 1.469, de 29 dezembro 2000. Estabelece os procedimentos e responsabilidades relativos ao controle e vigilância da qualidade da água para consumo humano e seu padrão de potabilidade, e dá outras providências. Diário Oficial União. 19 jan 2001.
12. Ministério da Saúde (BR). Programa nacional de vigilância em saúde ambiental relacionada à qualidade da água para consumo humano. Brasília: Ministério da Saúde; 2005[acesso 21 jun 2014]. Disponível em http://bvsms.saude.gov.br/bvs/publicacoes/programa_agua_consumo_humano.pdf
13. Cesa K, Abegg C, Aerts D. A vigilância da fluoretação de águas nas capitais brasileiras. Epidemiol Serv Saude. 2011;20(4):547-55. https://doi.org/10.5123/S1679-49742011000400014
14. Frazão P, Soares CCS, Fernandes GF, Marques RAA, Narvai PC. Fluoretação da água e insuficiências no sistema de informação da política de vigilância à saúde. Rev Assoc Paul Cir Dent. 2013;67(2):94-100.
15. Lima CRA, Schramm JMA, Coeli CM, Silva MEM. Revisão das dimensões de qualidade dos dados e métodos aplicados na avaliação dos sistemas de informação em saúde. Cad Saude Publ. 2009;25(10):2095-109. https://doi.org/10.1590/S0102-311X2009001000002
16. Ministério da Saúde (BR). Portaria Nº 2.914, de 12 de dezembro de 2011. Dispõe sobre os procedimentos de controle e de vigilância da qualidade da água para consumo humano e seu padrão de potabilidade. Diário Oficial União. 14 dez 2011.
17. Valentim LSO. Potabilidade da água e controle do risco sanitário: uma visão a partir do contexto paulista. In: Frazão P, Narvai PC, organizadores. Cobertura e vigilância da fluoretação da água de abastecimento público. São Paulo: Universidade de São Paulo; 2017. p. 33-40.
18. Alves RX, Fernandes GF, Razzolini MTP, Frazão P, Marques RAA, Narvai PC. Evolução do acesso à água fluoretada no Estado de São Paulo, Brasil: dos anos 1950 à primeira década do século XXI. Cad Saude Publ. 2012;28(supl):s69-s80. https://doi.org/10.1590/S0102-311X2012001300008
19. Universidade de São Paulo – USP. Consenso técnico sobre classificação das águas de abastecimento segundo o teor de flúor. São Paulo: USP; 2011[acesso 21 jun 2014]. Disponível em http://www.cecol.fsp.usp.br/dcms/uploads/arquivos/1398177715_CECOL-USP-ClassificacaoAguasSegundoTeorFluor-DocumentoConsensoTecnico-2011.pdf
20. Akpata ES, Danfillo IS, Otoh EC, Mafeni JO. Geographical mapping of fluoride levels in drinking water sources in Nigeria. Afr Health Sci. 2009;9(4):227-33.
21. Silva JS, Moreno WG, Forte FDS, Sampaio FC. Natural fluoride levels from public water supplies in Piauí State, Brazil. Cienc Saude Colet. 2009;14(6):2215-20. https://doi.org/10.1590/S1413-81232009000600030
22. Fawell J, Bailey K, Chilton J, Dahi E, Fewtrell L, Magara Y. Fluoride in drinking-water. London: World Health Organization; 2006.

Conflict of Interest

Authors have no potential conflict of interest to declare, related to this study's political or financial peers and institutions.