Italian Journal of Food Safety 2017; volume 6:6826

Assessment of relationship between rainfall and *Escherichia coli* in clams (*Chamelea gallina*) using the Bayes Factor

Cesare Ciccarelli,[1] Angela Marisa Semeraro,[1] Melina Leinoudi,[2] Vittoria Di Trani,[1] Sandra Murru,[1] Piero Capocasa,[1] Elena Ciccarelli,[3] Luca Sacchini[4]

[1]Regional Public Health Service Corporation of Marche – Extended Area N° 5, San Benedetto del Tronto (AP), Italy; [2]General Chemical State Laboratory, Thessaloniki, Greece; [3]Biologist; [4]Veterinarian, Italy

Abstract

Consumption of bivalve shellfish harvested from water contaminated with sewage pollution presents a risk of human infections and targeting control measures require a good understanding of environmental factors influencing the transport and the fate of faecal contaminants within the hydrological catchments. Although there has been extensive development of regression models, the point of this paper, focused on the relationship between rainfall events and concentrations of *Escherichia coli* monitored in clams, was the use of a Bayesian approach, by the Bayes Factor. The study was conducted on clams harvested from the south coast of Marche Region (Italy), a coastal area impacted by continuous treated effluents, intermittent rainfall-dependent untreated sewage spillage - as a consequence of stormwater overflowing - and rivers with an ephemeral flow regime. The work compared the different interpretation criteria of Bayes Factor, confirmed that *E. coli* concentrations in clams from the studied area varied in correlation with rainfall events, and demonstrated the effectiveness of Bayes Factor in the assessment of shellfish quality in coastal marine waters. However, it suggested that further investigations would be warranted to determine which environmental factors provide the better basis for accurate and timely predictions. Furthermore the gathered data could be useful, to the local authorities of Marche Region, in the definition of flexible monitoring programmes, taking into account the atmospheric events that could affect the correct functioning of sewage managing systems and the flow of tributary rivers.

Introduction

Consumption of bivalve shellfish harvested from water contaminated with sewage pollution presents a risk of human infections, mainly with species that are usually consumed raw or lightly cooked. A primary cause of bivalve shellfish related outbreaks worldwide is contamination, mainly by norovirus, during primary production associated with events of sewage pollution due to sewerage system failures and malfunctioning, extreme rainfall events overloading the treatment capacity of sewerage system and overboard disposal of faeces from boating activity (Campos *et al.,* 2015, 2017). Shellfish post-harvest purification treatments have a limited effectiveness as far as viruses are concerned, consequently the best control measures are the production of shellfish in waters that are not faecally contaminated and the restriction of commercial harvesting from contaminated waters (European Food Safety Authority, 2011). The European Union Regulation n. 854/2004 requires an evaluation of the sources and types of faecal contamination impacting shellfish harvesting areas combined with the monitoring of *Escherichia coli* in shellfish flesh to be undertaken in order for an indication of the risk of contamination with bacterial and viral pathogens to be provided. Targeting these control measures requires a good understanding of environmental factors influencing the transport of faecal contaminants within the hydrological catchments, particularly from inputs from human sewage pollution to commercial sewage beds (Cheng *et al.,* 2013; Ciccarelli *et al.,* 2014). So far a number of studies have confirmed the link between precipitations and increased microbial pollution which can result in reduced water quality in coastal environments (Pommepuy *et al.,* 2004; Strubbia *et al.,* 2016; Tilburg *et al.,* 2015).

Currently reductions in water quality after high precipitation events and the subsequent increases in river discharge lead local authorities to close shellfish harvesting areas after large events (Pommepuy *et al.,* 2004). But the inability of local authorities to accurately predict these events or to immediately assess the water quality exacerbates the losses to fishing economies (Tilburg *et al.,* 2015). So the need of accurate and timely predictions of water quality becomes acute. Mathematical models offer great potential in the delineation of shellfish harvesting exclusion zones (Ciccarelli *et al,* 2014; Mok *et al.,* 2016; Pommepuy *et al.,* 2004; Strubbia *et al.,* 2016; Suffredini *et al.,* 2008), especially where contamination arises from point sources discharges as per this study. However we believe more work is needed to validate and improve these models from a risk assessment perspective. Following a recent work (Ciccarelli *et al,* 2017), that study investigates the relationship between *Escherichia coli* contamination levels in clams (*Chamelea gallina)* harvested from the coast of Marche Region (Italy), from October 2002 to December 2016, and a potential predictor variable, such as rainfall events, using the Bayes Factor, a Bayes Theorem's application.

Despite the fact that the impact of faecal contaminants discharge into the environment on shellfish quality is difficult to evaluate, it is crucial to understand the way in which shellfish become contaminated if we want to improve it (Strubbia *et al.,* 2016).

Correspondence: Cesare Ciccarelli, Regional Public Health Service Corporation of Marche – Extended Area N° 5, 63074 San Benedetto del Tronto (AP), Italy.
Tel: +39.0735.7937474. Fax: +39.0735 793529.
E-mail: cesare.ciccarelli@sanita.marche.it

Key words: Bayes Factor; Clams; *Escherichia coli*; Faecal contamination; Norovirus.

Received for publication: 31 May 2017.
Revision received: 26 June 2017.
Accepted for publication: 27 June 2017.

Italian Journal of Food Safety 2017; 6:6826
doi:10.4081/ijfs.2017.6826

Materials and Methods

A database was created by the authors of this study, containing the results on *E.coli* levels as obtained from the official monitoring plan executed by the Local Competent Authority (LCA) according to EU Regulation n. 854/2004 (European Commission, 2004b). Results were reported as most probable number (MPN/100g), quantified in wild clams (*Chamelea gallina)* from ten sampling points within six harvesting areas, on the south coast of Marche Region On the whole the database contains 1032 values as shown on Table 1. The periodical Sanitary Survey, executed by the LCA and aimed to evaluate the impact of faecal contamination on shellfish harvesting areas, has suggested these coastal areas are impacted by continuous treated effluent, intermittent, rainfall-dependent, untreated sewage spillages, as consequence of stormwater overflowing, and rivers with an ephemeral flow regime. The monitoring plan was based on sampling points representative of the areas, following the directions of the Sanitary Survey, whereas samples were collected following a protocol, aimed at minimizing the secondary contamination of shellfish, based on the Guidelines of the EU Commission (European Commission, 2014). The reference laboratory was the IZSUM of Fermo and analysis method was the official reference method in force (now ISO TS 16649-3). The database also contains the daily rainfall levels, of the seven days before the sampling, recorded from two gauging stations managed by the Regional Authority for Civil Protection and representative of the studied area (Figure 1 shows the map of harvesting areas, sampling points and gauging stations).

Two different time windows were chosen for the analysis of the effects of rainfall on *E. coli* concentration; one concerning rainfall levels four days before sampling and one concerning rainfall levels seven days before sampling. For each window the rainfall levels were further categorized as cumulative and maximum and a criterion (R) was appointed. For the cumulative rainfall levels the criterion (R) was classified as R_{cum} >5 mm, >10 mm, >15 mm, >20 mm, >30 mm, >40 mm and >50 mm; for the maximum rainfall levels the criterion (R) was classified as R_{max} >5 mm, >10 mm, >15 mm, >20 mm and >25 mm. Taking into account the requirements provided by EU Regulation n. 853/2004 (European Commission, 2004a), for each sampling point *E. coli* results were classified as compliant (C) when ≤230 MPN/100g or non compliant (nC) when >230 MPN/100g. Following the Bayesian approach (Kaas and Raftery, 1995), the probability of non compliant *E. coli* levels, given a rainfall level occurs (R>) is:

$$P_{(nC|R>)} = P_{(nC)} * P_{(R>|nC)} / P_{(nC)} * (P_{(R>|nC)} + P_{(C)} * P_{(R>|C)}) \qquad (1)$$

and the related Bayes Factor (BF) is:

$$(BF>)\ B>_{nC,C} = P_{(R>|nC)} / P_{(R>|C)} \qquad (2)$$

By similar way the probability of non compliant *E. coli* levels, given a rainfall level doesn't occur (R<) is:

$$P_{(nC|R<)} = P_{(nC)} * P_{(R<|nC)} / P_{(nC)} * (P_{(R<|nC)} + P_{(C)} * P_{(R<|C)}) \qquad (3)$$

and the related BF is:

$$(BF<)\ B<_{nC,C} = P_{(R<|nC)} / P_{(R<|C)} \quad (4)$$

The BF is a summary of the evidence provided by the data in favour of one scientific theory, represented by a statistical model, as opposed to another (Scaranaro, 2005), and, with the BF> we compared the two hypothesis:

H1 : nC|R> (E. coli results are non compliant if rainfall level is > R*)*
H0 : C|R> (E. coli results are compliant if rainfall level is > R)

Whereas with BF< we compared the hypothesis:

H1 : nC|R< (E. coli results are non compliant if rainfall level is < R*)*
H0 : C|R< (E. coli results are compliant if rainfall level is < R*)*

We expressed BF> and BF< as *log10* and *loge*, respectively, for each classification of the rainfall criterion R and we verified the statistical significance using the criteria suggested by Jeffreys (J) and Kass and Raftery (K&R) (Kaas and Raftery, 1995): these criteria, the first as log_{10} and the second as log_e, are shown in Table 2. When the results were statistically significant, considering the occurrence of rainfall as a test for *E. coli* contamination, we calculated the respective sensitivity (sens), specificity (spec), positive (PPV) and negative (NPV) predictive values, based on the running prevalence. We used LibreOffice Calc software Version: 5.0.3.2.

Results

The calculations based on formulas (2) and (4) generated 240 values expressed as log_{10} *and* likewise as log_e. Table 3 summarises all obtained statistical significant values at the lower rainfall level for all clam

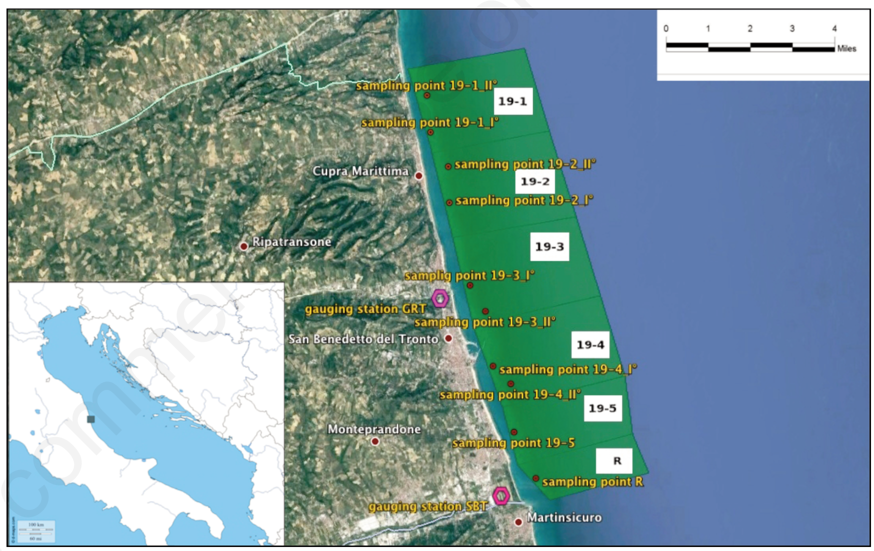

Figure 1. Map of harvesting areas, sampling points and gauging stations.

Table 1. Database description: summary of the results of the official monitoring plan, from October 2002 to December 2016, in wild clams, from ten sampling points on the south coast of Marche Region.

Sampling point	Period	*E. coli* results (MPN/100g)		
		Samples	≤230	>230
19-1_I	Oct-2002 Dec-2016	119	105	14
19-1_II	Oct-2002 Dec-2016	105	98	7
19-2_I	Oct-2002 Dec-2016	114	107	7
19-2_II	Jun-2003 Dec-2016	104	89	15
19-3_I	Oct-2002 Dec-2016	98	81	17
19-3_II	Oct-2002 Dec-2016	110	94	16
19-4_I	Oct-2002 Dec-2016	99	80	19
19-4_II	Oct-2002 Dec-2016	106	94	12
19-5	Oct-2002 Dec-2016	92	81	11
R	Oct-2002 Dec-2016	85	68	17

MPN, most probable number.

sampling points. The significant results are identified with an asterisk (*) if the evidence is substantial/positive and a double asterisk (**) if the evidence is strong, according to Table 2. Moreover, Table 3 shows the related sensitivity (sens), specificity (spec), positive predictive value (PPV) and negative predictive value (NPV) obtained from the statistical analysis of the results.

Discussion

A first point that can be made is that there is no evidence of a correlation between non compliant *E. coli* levels and the absence of rainfall incidents (expressed as BF<), whereas in many cases there is a substantial or strong evidence of a correlation between non compliant *E. coli* levels and given rainfall levels (expressed as BF>): that highlights the relevance of rainfall dependent contamination sources in this coastal area. Furthermore we can distinguish: no large differences are recognisable between the two described temporal windows, however this observation based on four days generally shows higher positive predictive values; the finding needs more investigation; cumulative level and maximum level are both predictive but the first one at a lower rainfall level; most significant results are at the lower rainfall levels; this observation could be related to the weakness of sewerage system to manage the water input from stormwater runoff in the described costal area, resulting in higher contamination levels.

Taking into account the rainfall level as a test to identify non compliant *E. coli* values, the related sensitivity is restricted and sometimes poor, the specificity is generally more high instead. The overall low prevalence of non compliant *E. coli* values keeps the positive predictive value at a low level, except for two cases for 4 days time window, whereas the negative predictive value is often very high: this could allow the authorities to predict when the clams are compliant with the *E. coli* criterion. However the relationship with environmental factors needs more investigation because the obtained rainfall predictive capacity is not strong enough.

As a last point we compared the differ-

Table 2. Bayes Factor interpretation criteria.

Jeffreys (log_{10})	Kass and Raftery (loge)	Evidence against the H_0 Hypothesis
0 to 1/2	0 to 2	Not worth more than a bare mention
½ to 1	2 to 6	Substantial/positive
1 to 2	6 to 10	Strong
>2	>10	Very strong

Table 3. Significant Bayes Factor of clams sampling points for both temporal windows.

Temporal window	Sampling point	Rainfall level	Loge BF>	Log_{10} BF>	Sens	Spec	PPV	NPV
7 days	19-1 I°	R_{cum}>5	1.17	0.51*	0.71	0.60	0.19	0.94
	19-1 I°	R_{max}>15	1.20	0.52*	0.36	0.89	0.29	0.91
	19-1 II°	R_{cum}>20	1.29	0.56*	0.43	0.85	0.17	0.95
	19-2 I°	R_{cum}>10	2.61*	1.13**	0.86	0.73	0.17	0.99
	19-2 I°	R_{max}>5	2.37*	1.03**	0.86	0.67	0.15	0.99
	19-2 II°	R_{cum}>5	1.36	0.59*	0.73	0.64	0.26	0.93
	19-2 II°	R_{max}>5	1.48	0.64*	0.73	0.67	0.28	0.94
	19-3 I°	R_{cum}>5	1.51	0.65*	0.76	0.65	0.32	0.93
	19-3 I°	R_{max}>5	1.59	0.69*	0.76	0.68	0.33	0.93
	19-4 I°	R_{cum}>5	1.29	0.56*	0.74	0.64	0.33	0.91
	19-4 I°	R_{max}>5	1.16	0.50*	0.68	0.66	0.33	0.90
	19-4 II°	R_{cum}>5	1.33	0.58*	0.75	0.60	0.19	0.95
	19-5	R_{cum}>5	2.48*	1.08**	0.91	0.60	0.24	0.98
	19-5	R_{max}>5	1.42	0.62*	0.73	0.65	0.22	0.95
	R	R_{cum}>5	1.66	0.72*	0.82	0.62	0.35	0.93
	R	R_{max}>5	1.28	0.56*	0.71	0.68	0.35	0.90
4 days	19-1 I°	R_{cum}>15	1.34	0.58*	0.14	0.97	0.40	0.89
	19-1 I°	R_{max}>15	1.57	0.68*	0.14	0.98	0.50	0.90
	19-2 I°	R_{cum}>40	2.94*	1.27**	0.14	1.00	1.00	0.95
	19-2 I°	R_{max}>15	1.82	0.79*	0.14	0.98	0.33	0.95
	19-2 II°	R_{cum}>20	1.35	0.59*	0.13	0.98	0.50	0.87
	19-2 II°	R_{max}>15	2.06*	0.89*	0.13	1.00	1.00	0.87
	19-3 I°	R_{cum}>5	1.52	0.66*	0.65	0.79	0.39	0.91
	19-3 I°	R_{max}>10	1.36	0.59*	0.35	0.93	0.50	0.87
	19-3 II°	R_{cum}>5	1.45	0.63*	0.44	0.89	0.41	0.90
	19-4 I°	R_{cum}>5	1.47	0.64*	0.61	0.81	0.42	0.90
	19-4 I°	R_{max}>5	1.30	0.56*	0.53	0.84	0.43	0.88
	19-4 II°	R_{cum}>15	1.55	0.67*	0.96	0.88	0.43	0.91
	19-4 II°	R_{max}>10	1.36	0.59*	0.17	0.97	0.40	0.90
	19-5	R_{cum}>5	1.97	0.85*	0.73	0.79	0.32	0.96
	19-5	R_{max}>5	1.40	0.61*	0.55	0.81	0.29	0.93
	R	R_{cum}>5	1.54	0.67*	0.65	0.81	0.46	0.90
	R	R_{max}>5	1.54	0.67*	0.59	0.85	0.50	0.89

sens, sensitivity; spec, specificity; PPV, positive predictive values; NPV, negative predictive values. *Substantial/positive evidence against *H0* hypotesis when log_{10} >0,5 or log_e>2; **strong evidence against *H0* hypotesis when log_{10} >1.

ences between interpretation criteria (J) based on log_{10}, and (K&R) based on log_e shown in the Table 3.

As expected (K&R) is more conservative: in fact, by using (J) we get 29 cases in which the evidence of a correlation between rainfall and *E. coli* contamination is substantial and 4 cases where it is strong. By comparison the use of the (K&R) criterion gives evidence only of a substantial correlation, and only in five cases. It can be seen in Table 3 through that four of these coincide with the four cases where the use of (J) shows strong evidence of correlation, the fifth also coinciding with a case in which the use of (J) shows evidence.

Conclusions

The Bayes Factor, as a summary of the evidence provided by the data in favour of one scientific theory, represented by a statistical model as opposed to another, could be a useful tool in the studies on relationships between environmental factors and shellfish contamination. Since the method is not dependant on the faecal indicator used it should keep its effectiveness with viruses or protozoa contamination too. In fact in this study, using Bayes Factor, we obtained substantial or strong evidence of a relationship between rainfall level and *E. coli* contamination in clams, in a well delimited South Marche coast on Adriatic Sea, as shown on the Table 3. Nevertheless, the relationship with environmental factors needs more investigation because the rainfall predictive capacity is not as strong as needed. Finally the data gathered in this study could be useful to the LCA for the periodical review of the Sanitary Survey, for the definition of targeted prevention strategies and for the modulation of monitoring plans taking into account the atmospheric events that could affect the correct functioning of sewage managing systems and the flow of the tributary rivers.

References

Campos CJA, Avant J, Gustar N, Lowther J, Stockley L, Lees DN, 2015. Fate of human noroviruses in shellfish and water impacted by frequent sewage pollution events. Environ Sci Technol 49:8377-85.

Campos CJA, Kershaw S, Morgan OC, Lees DN, 2017. Risk factors for norovirus contamination of shellfish water catchments in England and Wales. Int J Food Microbiol 241:318-24.

Cheng J, Niu S, Kim Y, 2013. Relation between water quality parameters and the survival of indicator microorganisms, Escherichia coli-in a stormwater wetland. Water Sci Technol 68:1650-6.

Ciccarelli C, Semeraro AM, Aliventi A, Di Trani V, Capocasa P, 2014. Seasonal variations of Escherichia coli contamination in clams (Chamelea gallina) harvested in the Adriatic Sea (San Benedetto del Tronto district, Italy). It J Food Safety 3:1645-7.

Ciccarelli C, Semeraro AM, Leinoudi M, Di Trani V, Murru S, Capocasa P, Ciccarelli E, 2017. Assessment of relationship between rainfall and Escherichia coli in bivalves using the Bayes Factor. 11th International Conference on Molluscan Shellfish Safety Book of Abstracts 1:24.

European Commission, 2004a. Regulation of the European Parliament and of the Council of 29 April 2004 laying down specific hygiene rules for on the hygiene of foodstuffs, 853/2004/CE. In: Official Journal, L 139/55, 30/04/2004.

European Commission, 2004b. Regulation of the European Parliament and of the Council of 29 April 2004 laying down specific rules for the organisation of official controls on products of animal origin intended for human consumption, 854/2004/CE. In: Official Journal, L 139/206, 30/04/2004.

European Commission, 2014. Community guide to the principles of good practice for the microbiological classification and monitoring of bivalve mollusc production and relaying areas with regard to Regulation 854/2004. Available from: http://ec.europa.eu/food/safety/biosafety/food_hygiene/guidance_en

European Food Safety Authority, 2011. Scientific opinion on an update on the present knowledge on the occurrence and control of foodborne viruses. EFSA J 9:2190-288.

Kaas RE, Raftery AE, 1995. Bayes Factors. J Am Stat Assoc 90;773-95.

Mok JS, Lee TS, Kim PO, Lee HJ, Kwang SA, Shim KB, Lee KJ, Yung YJ, Kim JH, 2016. Bacteriological quality evaluation of seawater and oysters from the Hansan-Geojeman area in Korea, 2011-2013: impact of inland pollution sources. SpringerPlus 5:1412.

Pommepuy M, Dumas F, Caprais MP, Camus P, Le Mennec C, Parneadeau S, Haugarreau L, Sarrette B, Vilagines P, Pothier P, Kholl E, Le Guyadier F, 2004. Sewage impact on shellfish microbial contamination. Water Sci Technol 50:117-24.

Scaranaro L, 2005. Alcune considerazioni sul fattore di Bayes nei problemi di verifica d'ipotesi. Padova University Thesis, Padova, Italy.

Strubbia S, Lyons BP, Lee RJ, 2016. Geographical and temporal variation of E. coli and norovirus in mussels. Mar Pollut Bull 107:66-70.

Suffredini E, Corrain C, Arcangeli G, Fasolato L, Manfrin A, Rossetti E, Biazzi E, Mioni R, Pavoni E, Losio MN, Sanavio G, Croci L, 2008. Occurrence of enteric viruses in shellfish and relation to climatic-environmental factors. Lett Appl Microbiol 47:467-74.

Tilburg CE, Jordan LM, Carlson AE, Zeeman SI, Yund PO, 2015. The effects of precipitation, river discharge, land use and coastal circulation on water quality in coastal Maine. Roy Soc Open Sci 2:140429.