# European legislation and live bivalve molluscs: Are the criteria for microbiological safety matching with the criteria for sanitary classification of harvesting areas?

Cesare Ciccarelli,[1] Melina Leinoudi,[2] Angela Marisa Semeraro,[1] Vittoria Di Trani,[1] Elena Ciccarelli,[3] Gaia Consorti[4]

[1]Regional Public Health Service Corporation of Marche – Extended Area 5, Italy; [2]Chemist, Greece; [3]Biologist, UK; [4]Veterinarian, Italy

## Abstract

The European Union (EU) established the criteria for the classification of shellfish harvesting areas, based on the results of monitoring *E. coli* in shellfish. The EU also defined *E. coli* as a microbiological criterion for end product safety, based on a three-class sampling plan. Both criteria are based on the MPN method, a test with a statistical approach and in which different factors contribute to the variability of the outcomes. This theoretical study, based on combinatorics and Bayes' theorem for conditional probability, investigates the consistency between these two criteria and aims to determine the probability of obtaining false compliant and non-compliant results when applying the safety criterion test to shellfish placed on the market. The results show that in the second case, the probability of non-compliant outcomes does not appear negligible within a range between 10% and 50% in the different hypothesized scenarios, with a probability of false non-compliant outcomes over 10%. In addition, the Bayes' Theorem shows that Class A, or Class B areas (as allowed), could be the origin of non-compliant shellfish, with a not negligible frequency. Therefore, within the limits of the assumed working hypotheses, the safety criterion for *E. coli*, as described in Regulation EC/2073/2005, does not appear to be consistent and coordinated with the classification criteria stated in the Regulation EU/2019/627 and it is not closely related to the sanitary status of shellfish harvesting area.

## Introduction

Consumption of raw fishery products involves non-negligible risks to human health, especially in the case of live bivalve molluscs (shellfish), as these filtering organisms can accumulate both nutrients and contaminants from the aquatic environment such as pathogenic bacteria, viruses, protozoa, algal biotoxins, heavy metals, etc.

To reduce the risk from faecal-oral transmitted microorganisms, the shellfish harvesting areas are classified according to the results of a monitoring plan based on faecal indicator bacteria. These organisms are used as indicators of faecal pollution and thus the possible presence of enteric pathogens such as bacteria (*Salmonella, Campylobacter*, pathogenic forms of *E. coli*), viruses (Norovirus, Hepatitis A virus) and parasites (*Giardia, Cryptosporidium*) (Potasman *et al.*, 2002). Since *Escherichia coli* is extensively found in the intestinal tract and in the faeces of mammals and birds, it is widely used as an indicator of faecal contamination (Younger *et al.*, 2002; Lee and Murray, 2010).

In fact, the European Union (EU) has established a series of rules, based on the monitoring of *E. coli* in shellfish flesh, aiming to protect shellfish consumers and to reduce the incidence of microbial diseases transmitted by this food category. With Regulation EU/2019/627 (European Commission, 2019), the EU laid down the criteria for the classification of shellfish harvesting areas and determined the level of post-harvest treatment required before shellfish can be considered suitable for human consumption. The harvesting areas are classified in three sanitary levels: shellfish coming from Class A areas can be directly placed on the market without further post-harvest treatments; those from Class B areas must undergo purification treatment and shellfish from Class C areas must be subjected to extended relaying treatment in areas with a sanitary level equivalent to Class A areas, before being placed on the market alive; shellfish from areas not classified cannot be placed on the market.

With Regulation EC/2073/2005 (European Commission, 2005), the EU established a safety criterion using *E.coli* as the microbiological parameter for assessing end-product safety. It is based on a three-class sampling plan and must be adhered to for live shellfish to be put on the market for direct human consumption.

Both regulations are based on the MPN Method described in EN ISO 16643-3: 2005: a harvesting area can be classified as "A" if 80% of the values obtained from the monitoring is ≤230 MPN/100g and no value is >700 MPN/100g; an outlier with a value of more than 700 MPN/100g can be accepted after risk assessment. The EU Guidelines (European Commission, 2018) later clarified that this assessment must take into account no less than 24 monthly values.

Likewise, shellfish placed on the market are considered compliant with the safety criterion if the value of at least 4 of the 5 sample units tested is ≤230 MPN/100g and none is >700 MPN/100g.

The MPN Method is a test based on a statistical approach, estimating the concentration of viable microorganisms in a given sample by assessing the presence/absence of microorganisms on replicated culture series with 10-fold diluted inocula, typically 5 series of cultures per three dilutions (Walker *et al.*, 2018).

However, many factors contribute to the variability of the MPN Method applied to shellfish (Ciccarelli *et al.*, 2017). Beside the uncertainty of the method, inherent in all microbiological tests and expressed in terms of repeatability and reproducibility of the test results, there is also the concern of sample-to-sample variability, one that is often raised with regards to the use of this method in environmental monitoring programs (van Elsas *et al.*, 2011). This factor is because the examination is inevitably performed on a pool of shellfish, of which each individual may contain a different number of *E. coli* and therefore the result only expresses the average of the entire sample (Walker *et al.*, 2018). Furthermore, the distribution of *E. coli* among the shellfish in any one bed will usually vary across the same bed and over time. This depends on a range of factors, such as growth and respiration of the individual animals, residence time,

bioaccumulation kinetics and decay of *E. coli* in the environment as well as the dispersal of *E.coli* in the environment related to oceanographic factors such as salinity, tides or streams (van Elsas *et al.*, 2011). Therefore, the result from a single sample cannot be expected to present an accurate estimate of the overall mean contamination of that bed. Long-term data provide the best overall indication of the sanitary status of a given harvesting area and are the rationale behind the classification system used in the EU (Walker *et al.*, 2018). Likewise, the three-class sampling plan, on which the *E. coli* safety criterion is based, may also be affected by this variability.

To minimize the effect of these factors on the classification of shellfish harvesting areas, the EU guidance (European Commission, 2018) introduced a number of criteria to avoid the related bias such as: restricting sampling to the points with the highest presumed level of faecal contamination; providing for monthly sampling over the entire year and on a randomised basis with respect to influencing variables such as rainfall, tide, stream, wind; and establishing a standard sampling protocol to minimise variation due to sampling and transport methods.

Based on these premises, the afore-mentioned points, namely i) non-uniform results from sample to sample, ii) the uncertainty of the method and iii) the selection of different sample sizes for classification criteria *vs* safety criteria (24 samples *vs* 5 samples), could have relevant effects on the expression of the results and, finally, on the consistency of the two criteria.

In fact, a recent study focused on the shortcomings of assessing the risk posed by human pathogenic bacteria and viruses, based on *E. coli* abundance in shellfish (Sharp *et al.*, 2021). The present study is based on a theoretical approach, and by using a probabilistic method and employing the Bayes' Theorem for the determination of conditional probability, aims to: i) investigate the consistency between the classification criteria for shellfish harvesting areas and the *E. coli* safety criterion for live shellfish, as established by the EU; ii) determine the probability of obtaining false compliant and non-compliant results when applying the safety criterion test to shellfish placed on the market.

## Materials and methods

*E.coli* content in shellfish samples is accepted, for monitoring purposes, as an indicator of the contamination level at the moment of sampling (Younger *et al.*, 2002). Temporal and spatial variation of contamination may be related to intra- and interspecific differences among shellfish (Lart and Hudson, 1993; Lees and Nicholson, 1995), proximity to polluting sources, environmental factors such as rainfall, tide, wind, and season (Younger *et al.*, 2002). That's why shellfish harvesting areas could be considered as a set of sample units with different *E. coli* content, with the extent of variability depending on the influences of the mentioned factors.

Assuming that, the study was organized into three steps:

*Step 1:* Definition of the working hypotheses. Two general hypotheses were formulated, one accepting high variability of contamination and one accepting low variability. For each hypothesis, different levels of contamination of the harvesting areas were allowed – both from a qualitative and quantitative point of view - and different levels of efficiency of post-harvest treatments were supposed

*Step 2:* Simulation of the results of MPN test: based on a probabilistic model, a simulation was carried out to determine the outputs generated by the MPN test. The Classification Process (CP) node was dedicated to obtaining results for the classification of harvesting areas, while the Safety Criterion (SC) node for the test on shellfish coming from these areas, subjected to post-harvest treatments if required, and put on the market.

*Step 3:* Conditional probability analysis: using the Bayes' Theorem separately on compliant and non-compliant outcomes generated for the safety criterion, the study determined the probability of shellfish originating from harvesting areas with a given sanitary status and having been subjected to post-harvest treatments.

Figure 1 shows the graphic representation of the core model used in this study, as described above, while Table 1 lists the parameters and variables used.
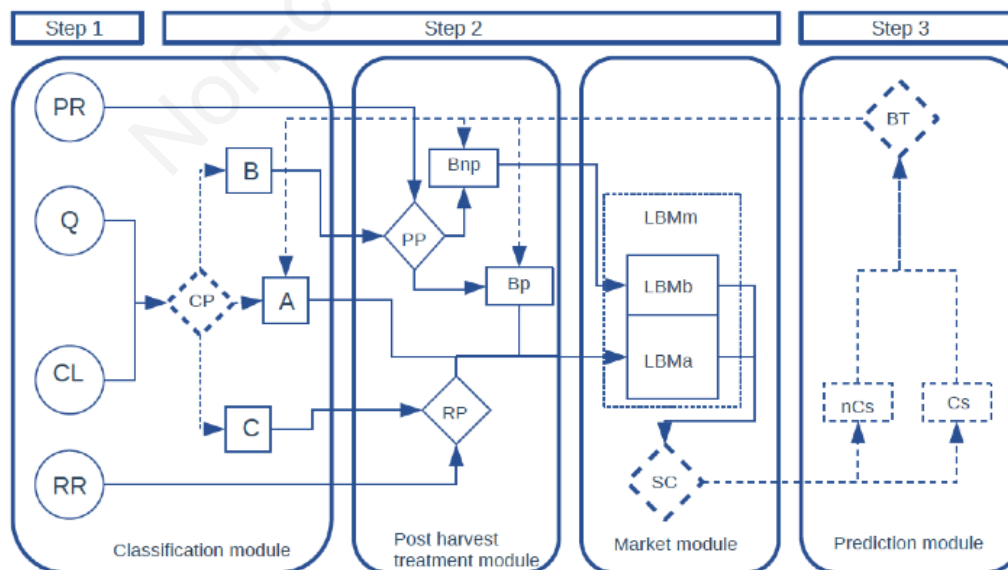


**Figure 1. Graphic description of the core model. Circles represent assumed input; rectangles are variables estimated by equations and diamonds represent calculation and stochastic nodes. Dash lines are predicted variables. See Table 1 for description of parameters and distribution of relationship.**

## Step 1

The structure of the model is as follows.

- For each hypothesis, four different scenarios of contamination were taken into account: very low level (VLL), medium level (ML), high level (HL) and very high level (VHL) into each scenario, four ranks of MPN test results were envisaged, corresponding to the ones stated in the Regulations under study. The ranks were named as:

- α when *E. coli* ≤230 MPN/100g,
- β if >230 MPN/100g and ≤700 MPN/100g,
- γ when >700 MPN/100g and ≤4.600 MPN/100 and
- δ if >4.600 MPN/100g.

- Since results >46.000 MPN/100g do not allow a sanitary classification of the harvesting area, this rank was excluded from calculations.

- For each rank, 6 distinct frequency levels were assumed and indicated with θ.

- The variability of the amount of contamination in the collected shellfish was accounted for by working three hypothesized quantities' ratios into the model; L>H = 80/20 (which means a low contamination level L is prevalent to a high contamination level H), L=H =50/50 and L<H = 20/80.

All the formulated hypotheses have been summarized in Supplementary Table S1.

Other factors comprised in the model were: i) the variability inherent in the purification treatment: the efficacy of the treatment depends on many factors, such as the contamination level, shellfish physiology and water quality; to incorporate the variability of this process into the model, four different possible levels of effectiveness were considered, namely 90%, 75%, 50% and 0%, the latter declaring an absence of purification; ii) the part of shellfish effectively purified assumes the values of the lowest hypothesized contamination level, namely α with θ=0.95, while the untreated part maintains the original contamination level. The postulation that the post-harvest relaying treatment is 100% effective (as it lasts a long time and takes place in areas with a very low level of contamination), and thus the shellfish in this scenario also assume the lowest level of contamination.

## Step 2

- Both stochastic nodes were carried out notionally and were based on the general assumptions that i) the use of MPN Method is carried out according to EN ISO 16643-3: 2005 and ii) the sampling protocol results in random and representative samples and avoids secondary contamination.

- CP node: the criteria to determine the health level during the recurring review of harvesting area classification have been summarized on Supplementary Table S4; the third column shows the corresponding criteria adopted from a mathematical point of view based on the last 24 results of monitoring plan.

- Assuming that all 24 results are independent and related to the frequency of the four contamination ranks, the probability that a given harvesting area is classified as "A" ($P_A$) is given by the sum of the probabilities of all combinations of results consistent with the classification's criteria, *i.e.* $P_A = P_1 + P_2 ... + P_x$, where $P_x$ is obtained with a probability mass function expressed as:

$$P_{(x|\theta,r,n)} = \sum^k \theta_\alpha{}^{r\alpha} \cdot \theta_\beta{}^{r\beta} \cdot \theta_\gamma{}^{r\gamma} \cdot \theta_\delta{}^{r\delta} \qquad [1]$$

where r (rα, rβ, rγ and rδ) is the number of results belonging to the four ranks and rα+rβ+rγ+rδ=24=n, while $\theta_\alpha$, $\theta_\beta$, $\theta_\gamma$ and $\theta_\delta$ are the frequencies of the ranks, with θ∈ [0 , 1] and $\theta_\alpha+\theta_\beta+\theta_\gamma+\theta_\delta$=1, and finally $k=P^r_n$ is the number of permutations for each combination.

- The probability that a given harvesting area is classified as "C" ($P_C$) is obtained in a similar way while, clearly, the probability to classify the area as "B" ($P_B$) is $P_B$=1-$P_A$-$P_C$.

- The obtained probability values, arranged according to hypothesized contamination scenarios, also represent the amount of shellfish coming from all classified harvesting areas. These values are resumed at Supplementary Table S2-1.

- After accounting for the formulated post-harvest efficiency hypotheses, the above values were "converted" and summarized at Supplementary Table S2-2a through Supplementary Table S2-2d. All these values, related to the unit for each hypothesis, also represent the number of shellfish placed on the market after being subjected to the safety criterion test.

- SC node: the three-class sampling plan conditions, as required by amended Regulation EC/2073/2005, are reported on Table 1, alongside the corresponding criteria selected for the statistical processing in this study. Assuming that the 5 outcomes are independent and related to the frequency of the contamination ranks, the probability of obtaining a compliant outcome ($P_{Cs}$) is given by the sum of the probabilities of all combinations of results that allow for the tested sample to be regarded as compliant, namely $P_{Cs}$=$P_1$+$P_2$...+$P_x$, and is obtained with a mass probability function, similar to the previous one, expressed as:

$$P_{(x|\theta,r,n)} = \sum^k \theta_\alpha{}^{r\alpha} \cdot \theta_\beta{}^{r\beta} \cdot \theta_{\gamma+\delta}{}^{r\gamma+\delta} \qquad [2]$$

where r (rα, rβ and rγ+δ) is the number of outcomes belonging to the four ranks with rα+rβ+rγ+δ=5=n, while $\theta_\alpha$, $\theta_\beta$ and $\theta_{\gamma+\delta}$ are the frequencies of the ranks with θ∈ [0, 1] and $\theta_\alpha+\theta_\beta,+\theta_{\gamma+\delta}$=1 and $k=P^r_n$ is the number of permutations for each combination. The values of $P_{Cs}$, calculated for the corresponding values of $\theta_\alpha$ and $\theta_{\gamma+\delta}$, are resumed in Supplementary Table S5.

- Likewise, the probability of obtaining a not compliant outcome ($P_{nCs}$) is given by $P_{nCs}$=1-$P_{Cs}$.

## Step 3

- Bayes Theorem is a simple mathematical formula used to calculate conditional probability (Joyce, 2019): it describes the probability of an event, based on a prior knowledge of the probability of another event related to the first one. The formula can be described by the following equation:

$$P_{(Ej|E)} = P_{(Ej)} \cdot P_{(E|Ej)} / \sum^k P_{(Ek)} \cdot P_{(E|Ek)} \qquad [3]$$

with

$$P_{(Ej)} \cap P_{(Ek)} = \emptyset \text{ and } P_{(Ek1)}+P_{(Ek2)}\ldots+P_{(Ekn)}=1$$

- Applying this equation in the case of an overall sampling plan to check the safety criterion - assuming that it is random, representative of all harvesting areas, and the corresponding outcomes are independent - the probability ($P_{(Ej|E)}$) that compliant shellfish originate from a Class A area is related to:

- the general probability ($P_{(Ej)}$) that marketed shellfish originate from a Class A area,
- the probability ($P_{(E|Ej)}$) of obtaining a compliant outcome with the safety criterion when shellfish originate from Class A areas,
- the probability ($\sum^k P_{(Ek)} \cdot P_{(E|Ek)}$) of obtaining a compliant outcome with the safety criterion regardless of the classification of the harvesting area of origin
- Likewise, for non-compliant outcomes ($\bar{E}$), the probability that the tested shellfish originate from a Class A area ($P_{(Ej|\bar{E})}$) can be determined with the following equation:

$$P_{(Ej|\bar{E})} = P_{(Ej)} \cdot P_{(\bar{E}|Ej)} / \sum^k P_{(Ek)} \cdot P_{(\bar{E}|Ek)} \qquad [4]$$

where $P_{(\bar{E}|Ej)}$ is the probability of obtaining an outcome not compliant with the safety criterion if shellfish originate from a Class A area, while $\sum^k P_{(Ek)} \cdot P_{(\bar{E}|Ek)}$ is the probability of obtaining outcomes not compliant, regardless of the classification of the

harvesting area of origin.

- The values of $P_{(Ej)}$, which represent the probability, namely the amount related to the unit, that shellfish are coming from an harvesting area with a given sanitary status, are those already determined in the step 2 and resumed in Supplementary Table S2-2a through Supplementary Table S2-2d.

- The values of $P_{(E|Ej)}$ and $P_{(\bar{E}|Ej)}$ are the product of $P_{(Ej)}$, mentioned above, and the corresponding value, of $P_{Cs}$ or $P_{nCs}$, for each contamination rank.

- The values of $\sum^{k} P_{(Ek)} \cdot P_{(E|Ek)}$ and $\sum^{k} P_{(Ek)} \cdot P_{(\bar{E}|Ek)}$ were determined from the corresponding values of $P_{(Ej)}$, $P_{(E|Ej)}$ and $P_{(\bar{E}|Ej)}$ described above. These values, summarized

on Table 2, also represent the general probability of obtaining outcomes compliant or not compliant to the safety criterion for *E. coli* with the MPN test.

- Table 3 summarizes the values of $P_{(Ej|E)}$, referred to the probability that shellfish compliant with the safety criterion are coming from Class B areas after not effective purification; Table 4, instead, summarizes the corresponding probability values of the not compliant outcomes, namely $P_{(Ej|\bar{E})}$, for shellfish coming from Class A areas or Class B areas after effective purification.

- Furthermore, these probability values, weighted by the corresponding values of $P_{Cs}$

or $P_{nCs}$, represent the overall frequencies of MPN test outcomes provided for a sampling plan to check the safety criterion, as defined above.

- Therefore, following Gerhardt and Keller (Gerhardt and Keller, 1986) and assuming as "true compliant" a compliant outcome when shellfish are coming from Class A or Class B after an effective post-harvest treatment, it is feasible to determine the probability of obtaining false compliant results and, likewise, the probability of obtaining false not compliant results.

- All calculations were performed using Libre Office Calc Version 6.4.4.2 spreadsheet.

**Table 1. Description of parameters and variables used on the core model.**

| Variable/ parameter | Definition | Description and output |
|---|---|---|
| CL | Contamination level | The supposed contamination level of shellfish from harvesting areas, corresponding to a specific frequency ($\theta$) of a contamination rank ($a$, $\beta$, $\gamma$ or $\delta$) of each of the four contamination scenarios (VLL, ML, HL, VHL) in which the two general hypotheses are divided into. Scenarios are resumed on Supplementary Table S1 |
| Q | Amount of production | Supposed amount of harvested shellfish for each rank of contamination level. We developed three different ratios (L>H, H=L and L<H) between "L" areas with low contamination level and "H" areas with high contamination level, where $Qn \in [0,1]$ and $Q=Q_1+Q_2 ...+Q_n=1$. Related values are shown on Supplementary Table S1. |
| RR | Relaying rate | Supposed rate of shellfish moving to the best contamination level rank. Relaying is a very efficient post-harvest treatment and a purification rate of 100% was supposed. |
| PR | Purification rate | Supposed rate of shellfish moving to contamination rank $\alpha$. Purification could prove to be not entirely effective, so we supposed four different rates: 0%, 50%, 75% and 90%, named, respectively, Pr0, Pr50, Pr75 and Pr90. |
| CP | Classification process | Probabilistic classification process following the classification criteria laid down by Regulation EU/2019/627 and based on mass probability function [1]; results are given at Supplementary Table S2-1. |
| A | Shellfish from Class A areas | Shellfish coming from harvesting areas classified as Class A, categorized according to contamination ranks and amount. Classification criteria based on 24 results are: $\alpha \geq 20$ and $\gamma+\delta \leq 1$. |
| B | Shellfish from Class B areas | Shellfish coming from harvesting areas classified as Class B, categorized according to contamination ranks and amount. . Classification criteria based on 24 results are: $\alpha<20$ or $\gamma+\delta>1$ or $\gamma\leq2$. |
| C | Shellfish from Class C areas | Shellfish coming from harvesting areas classified as Class C, categorized according to contamination ranks and amount. . Classification criteria based on 24 results are: $\gamma>2$. |
| PP | Purification process | Post-harvest treatment required for shellfish coming from Class B areas, The effectiveness is expressed as purification rate (PR): effectively purified shellfish acquire the contamination level best rank; those not effectively purified keep the original contamination level rank |
| RP | Relaying process | Post-harvest treatment required for shellfish coming from Class C areas. The very high effectiveness is expressed as relaying rate (RR) and harvested shellfish from C areas are included in the best contamination level rank of Class A areas. |
| Bp | Bivalve molluscs effectively purified | Amount of shellfish coming from Class B areas and effectively purified |
| Bnp | Bivalve molluscs not purified | Amount of ineffectively purified shellfish coming from Class B areas, categorized according to contamination ranks |
| LBMm | Live Bivalve Molluscs on the market | Total amount of shellfish put on the market for direct human consumption |
| LBMa | Live Bivalve Molluscs from Class A areas or equivalent | Part of shellfish put on the market for direct human consumption and originating from Class A areas or from Class B areas after effective purification. Amounts, categorized according to contamination level scenarios, purification rate and contamination ranks, are summarized on Supplementary Table S2-2a through Supplementary Table S2-2d. |
| LBMb | Live Bivalve Molluscs from Class B areas | Part of shellfish put on the market for direct human consumption, originating from Class B areas and ineffectively purified. Amounts, categorized according to contamination level scenarios, purification rate and contamination classes, are resumed on Supplementary Table S2-2a through Supplementary Table S2-2d. |
| SC | Safety criterion | Probabilistic process, based on mass probability function [2], used to determine the probability of obtaining compliant and non-compliant outcomes, following the three class sampling plan safety criterion (n=5, c=1, m=230 M=700) as laid down by Regulation CE/2073/2005. Criteria for compliance are: $\alpha \geq 4$ and $\gamma+\delta < 1$; criteria for non-compliance are: $\alpha < 4$ or $\gamma+\delta \geq 1$. |
| Cs | Compliant outcomes to safety criteria | Probability of obtaining samples compliant to the safety criterion. Data are resumed on Table 2. |
| nCs | Not compliant outcomes to safety criteria | Probability of obtaining samples not compliant to the safety criterion. Data are resumed on Table 2. |
| BT | Bayes' Theorem | Probability that shellfish, displaying either a compliant or non-compliant outcome, originate from a given classified harvesting area. Equation [3] was used to calculate the probability related to compliant outcomes and equation [4] was used for non-compliant outcomes. Data are resumed on Table 3 and Table 4. |

## Results and discussion

With step 1, resumed at Supplementary Table S1, a wide variety of scenarios of contamination levels in shellfish harvesting areas were expressed.

By step 2, the contamination level of shellfish coming from those areas, and regularly put on the market, was determined using a predictive module of MPN test; the simulation took into account the outcome of the sanitary classification process, according to the criteria defined by EU legislation, and different levels of effectiveness of shellfish purification process; the pre-purification contamination level and the post-purification contamination level are resumed at Supplementary Table S2-1 and Tables S2-2.

Furthermore, by applying the predictive module of the MPN test to the *E. coli* safety criterion, the probability of obtaining compliant (and non-compliant) results in shellfish put on the market was determined, according to the different contamination scenarios and purification effectiveness assumed: the results are summarized in Table 2.

As expected, the probability of obtaining compliant results is higher in scenarios with low variability and lower contamination levels while, in the reverse conditions, probability of obtaining not compliant results never reaches high values.

As shown, the probability that marketed shellfish, are compliant to the safety criterion never assumes very high values and, despite the best conditions – lowest level of contamination (low variability, very low level of contamination and ratio L> H) and high effectiveness of the purification process (Pr90) – the probability of obtaining shellfish compliant with the safety criterion does not reach 95%.

By step 3, with the Bayes' Theorem for conditional probability, two different kinds of data were obtained regarding:

I- The origin of shellfish tested according to the safety criterion. Table 3 summarizes the probability values ($P_{(Ej|E)}$) obtained when the test result is compliant. These values are directly proportional to the effectiveness of the purification process, while the initial contamination level is less relevant. When the effectiveness is very high (Pr90), the probability that the shellfish originate from Class A areas or equivalent (Class B areas after purification), assumes very high values >95%. However, this probability decreases when the effectiveness of the purification process is limited and, at the same time, the probability of compliant results in shellfish subjected to a not effective purification treatments arises to relevant levels.

Table 3 also summarizes the probability values ($P_{(Ej|Ē)}$) when the test result is not compliant. We can recognize how the probability that the not compliant outcome matches with shellfish coming from Class A areas, or equivalent, always reaches high levels, mainly when the purification's efficacy rate is higher. This fact is only seemingly unexpected because can be explained by the very high values that the factor $P_{(Ē|Ej)}$, namely the probability of not compliant outcomes in shellfish from Class A areas, it assumes in such conditions. At the same time, the competence of the safety criterion in identifying shellfish not effectively purified, is weak and becomes significant only when the purification efficacy rate is very low.

II - The ability of the safety criterion test to double as a test for the actual origin of marketed shellfish. The data obtained so far allow us to easily derive the false compliant and false non-compliant frequencies: results are also summarized on Table 4. It can be observed that the frequency of false not compliant outcomes always remains at a not negligible level, even if values reach around 10% in scenarios with a low contamination level and high effectiveness of purification treatment.

In a symmetrical way, the incidence of false compliant outcomes can reach more than 20% in scenarios with a high contamination level and limited purification's effectiveness.

**Table 2. PCs Probability of obtaining outcomes compliant to the three class sampling plan safety criterion for shellfish put on the market for direct human consumption; consequently the probability of non-compliant outcomes is PnCs= 1- PCs.**

| | | Purification rate | Probability of compliant outcomes | | | | Probability of not compliant outcomes | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | PCs_VLL | PCs_ML | PCs_HL | PCs_VHL | PnCs_VLL | PnCs_ML | PnCs_HL | PnCs_VHL |
| High variability | L>H | Pr90 | 0,88 | 0,87 | 0,84 | 0,82 | 0,12 | 0,13 | 0,16 | 0,18 |
| | | Pr75 | 0,86 | 0,84 | 0,81 | 0,78 | 0,14 | 0,16 | 0,19 | 0,22 |
| | | Pr50 | 0,83 | 0,81 | 0,75 | 0,72 | 0,17 | 0,19 | 0,25 | 0,28 |
| | | Pr0 | 0,76 | 0,73 | 0,64 | 0,60 | 0,24 | 0,27 | 0,36 | 0,40 |
| | L=H | Pr90 | 0,87 | 0,86 | 0,85 | 0,83 | 0,13 | 0,14 | 0,15 | 0,17 |
| | | Pr75 | 0,83 | 0,82 | 0,79 | 0,78 | 0,17 | 0,18 | 0,21 | 0,22 |
| | | Pr50 | 0,78 | 0,75 | 0,70 | 0,69 | 0,22 | 0,25 | 0,30 | 0,31 |
| | | Pr0 | 0,67 | 0,60 | 0,50 | 0,52 | 0,33 | 0,40 | 0,50 | 0,48 |
| | L<H | Pr90 | 0,85 | 0,85 | 0,85 | 0,77 | 0,15 | 0,15 | 0,15 | 0,23 |
| | | Pr75 | 0,81 | 0,79 | 0,76 | 0,69 | 0,19 | 0,21 | 0,24 | 0,31 |
| | | Pr50 | 0,73 | 0,68 | 0,61 | 0,57 | 0,27 | 0,32 | 0,39 | 0,43 |
| | | Pr0 | 0,57 | 0,46 | 0,30 | 0,33 | 0,43 | 0,54 | 0,70 | 0,67 |
| Low variability | L>H | Pr90 | 0,94 | 0,86 | 0,85 | 0,81 | 0,06 | 0,14 | 0,15 | 0,19 |
| | | Pr75 | 0,92 | 0,84 | 0,82 | 0,77 | 0,08 | 0,16 | 0,18 | 0,23 |
| | | Pr50 | 0,89 | 0,81 | 0,78 | 0,71 | 0,11 | 0,19 | 0,22 | 0,29 |
| | | Pr0 | 0,82 | 0,74 | 0,71 | 0,59 | 0,18 | 0,26 | 0,29 | 0,41 |
| | L=H | Pr90 | 0,92 | 0,87 | 0,85 | 0,82 | 0,08 | 0,13 | 0,15 | 0,18 |
| | | Pr75 | 0,87 | 0,82 | 0,80 | 0,77 | 0,13 | 0,18 | 0,20 | 0,23 |
| | | Pr50 | 0,78 | 0,73 | 0,71 | 0,69 | 0,22 | 0,27 | 0,29 | 0,31 |
| | | Pr0 | 0,61 | 0,56 | 0,53 | 0,51 | 0,39 | 0,44 | 0,47 | 0,49 |
| | L<H | Pr90 | 0,90 | 0,88 | 0,86 | 0,79 | 0,10 | 0,12 | 0,14 | 0,21 |
| | | Pr75 | 0,81 | 0,80 | 0,77 | 0,73 | 0,19 | 0,20 | 0,23 | 0,27 |
| | | Pr50 | 0,68 | 0,66 | 0,64 | 0,62 | 0,32 | 0,34 | 0,36 | 0,38 |
| | | Pr0 | 0,41 | 0,39 | 0,36 | 0,40 | 0,59 | 0,61 | 0,64 | 0,60 |

**Table 3. Probability (P_{(Ej|E)}) that marketed shellfish, compliant to the safety criterion, originate from Class B areas after ineffective purification versus Probability (P_{(Ej|Ê)}) that, marketed shellfish not compliant, originate from Class A areas or Class B areas after effective purification.**

| Contamination levels | Purification rate | Compliant outcomes P_{(Ej|E)} - Probability of origin from Class B areas after not effective purification | | | | Non-Compliant outcomes P_{(Ej|Ê)} - Probability of origin from Class A areas or Class B areas after effective purification | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | VLL | ML | HL | VHL | VLL | ML | HL | VHL |
| High variability | L>H Pr90 | 0,02 | 0,03 | 0,03 | 0,03 | 0,88 | 0,88 | 0,84 | 0,84 |
| | Pr75 | 0,05 | 0,07 | 0,07 | 0,08 | 0,74 | 0,73 | 0,67 | 0,67 |
| | Pr50 | 0,11 | 0,14 | 0,16 | 0,17 | 0,58 | 0,57 | 0,49 | 0,49 |
| | Pr0 | 0,24 | 0,30 | 0,37 | 0,41 | 0,39 | 0,38 | 0,29 | 0,30 |
| | L=H Pr90 | 0,03 | 0,03 | 0,02 | 0,02 | 0,82 | 0,78 | 0,70 | 0,77 |
| | Pr75 | 0,07 | 0,08 | 0,06 | 0,05 | 0,64 | 0,58 | 0,47 | 0,57 |
| | Pr50 | 0,15 | 0,18 | 0,14 | 0,10 | 0,46 | 0,39 | 0,28 | 0,38 |
| | Pr0 | 0,35 | 0,43 | 0,39 | 0,28 | 0,28 | 0,22 | 0,13 | 0,21 |
| | L<H Pr90 | 0,03 | 0,03 | 0,02 | 0,01 | 0,77 | 0,68 | 0,55 | 0,78 |
| | Pr75 | 0,08 | 0,08 | 0,05 | 0,03 | 0,56 | 0,45 | 0,30 | 0,58 |
| | Pr50 | 0,18 | 0,19 | 0,12 | 0,07 | 0,37 | 0,27 | 0,15 | 0,39 |
| | Pr0 | 0,47 | 0,57 | 0,47 | 0,23 | 0,21 | 0,13 | 0,05 | 0,22 |
| Low variability | L>H Pr90 | 0,01 | 0,01 | 0,02 | 0,04 | 0,75 | 0,89 | 0,89 | 0,85 |
| | Pr75 | 0,02 | 0,03 | 0,05 | 0,11 | 0,54 | 0,75 | 0,75 | 0,68 |
| | Pr50 | 0,03 | 0,07 | 0,11 | 0,24 | 0,36 | 0,60 | 0,59 | 0,50 |
| | Pr0 | 0,07 | 0,15 | 0,25 | 0,57 | 0,20 | 0,42 | 0,40 | 0,30 |
| | L=H Pr90 | 0,01 | 0,02 | 0,02 | 0,03 | 0,56 | 0,72 | 0,74 | 0,78 |
| | Pr75 | 0,04 | 0,05 | 0,06 | 0,09 | 0,32 | 0,49 | 0,52 | 0,57 |
| | Pr50 | 0,09 | 0,12 | 0,14 | 0,19 | 0,17 | 0,31 | 0,33 | 0,38 |
| | Pr0 | 0,22 | 0,30 | 0,37 | 0,52 | 0,07 | 0,16 | 0,16 | 0,20 |
| | L<H Pr90 | 0,02 | 0,03 | 0,03 | 0,02 | 0,45 | 0,53 | 0,58 | 0,77 |
| | Pr75 | 0,07 | 0,07 | 0,07 | 0,07 | 0,22 | 0,29 | 0,33 | 0,56 |
| | Pr50 | 0,16 | 0,17 | 0,17 | 0,16 | 0,10 | 0,15 | 0,17 | 0,37 |
| | Pr0 | 0,53 | 0,59 | 0,61 | 0,49 | 0,03 | 0,06 | 0,06 | 0,20 |

**Table 4. Outcomes of three class sampling plan safety criterion for shellfish put on the market for direct human consumption: the overall probability of False Compliant versus the overall probability of False Non-Compliant.**

| Contamination levels | Purification rate | Overall probability of False Compliant outcomes | | | | Overall probability of False Non-Compliant outcomes | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | VLL | ML | HL | VHL | VLL | ML | HL | VHL |
| High variability | L>H Pr90 | 0,02 | 0,02 | 0,02 | 0,02 | 0,11 | 0,12 | 0,13 | 0,15 |
| | Pr75 | 0,05 | 0,06 | 0,06 | 0,06 | 0,10 | 0,12 | 0,13 | 0,15 |
| | Pr50 | 0,09 | 0,11 | 0,12 | 0,12 | 0,10 | 0,11 | 0,12 | 0,14 |
| | Pr0 | 0,18 | 0,22 | 0,23 | 0,24 | 0,09 | 0,10 | 0,11 | 0,12 |
| | L=H Pr90 | 0,02 | 0,03 | 0,02 | 0,01 | 0,11 | 0,11 | 0,10 | 0,13 |
| | Pr75 | 0,06 | 0,07 | 0,05 | 0,04 | 0,11 | 0,11 | 0,10 | 0,12 |
| | Pr50 | 0,12 | 0,13 | 0,10 | 0,07 | 0,10 | 0,10 | 0,09 | 0,12 |
| | Pr0 | 0,24 | 0,26 | 0,19 | 0,14 | 0,09 | 0,09 | 0,07 | 0,10 |
| | L<H Pr90 | 0,03 | 0,03 | 0,01 | 0,01 | 0,11 | 0,10 | 0,08 | 0,18 |
| | Pr75 | 0,07 | 0,06 | 0,04 | 0,02 | 0,11 | 0,10 | 0,07 | 0,18 |
| | Pr50 | 0,13 | 0,13 | 0,07 | 0,04 | 0,10 | 0,09 | 0,06 | 0,17 |
| | Pr0 | 0,27 | 0,26 | 0,14 | 0,08 | 0,09 | 0,07 | 0,03 | 0,15 |
| Low variability | L>H Pr90 | 0,01 | 0,01 | 0,02 | 0,03 | 0,04 | 0,12 | 0,14 | 0,16 |
| | Pr75 | 0,01 | 0,03 | 0,04 | 0,08 | 0,04 | 0,12 | 0,13 | 0,15 |
| | Pr50 | 0,03 | 0,06 | 0,09 | 0,17 | 0,04 | 0,12 | 0,13 | 0,14 |
| | Pr0 | 0,06 | 0,11 | 0,18 | 0,34 | 0,04 | 0,11 | 0,12 | 0,12 |
| | L=H Pr90 | 0,01 | 0,02 | 0,02 | 0,03 | 0,05 | 0,09 | 0,11 | 0,14 |
| | Pr75 | 0,03 | 0,04 | 0,05 | 0,07 | 0,04 | 0,09 | 0,10 | 0,13 |
| | Pr50 | 0,07 | 0,09 | 0,10 | 0,13 | 0,04 | 0,08 | 0,09 | 0,12 |
| | Pr0 | 0,14 | 0,17 | 0,20 | 0,27 | 0,03 | 0,07 | 0,08 | 0,10 |
| | L<H Pr90 | 0,02 | 0,02 | 0,02 | 0,02 | 0,05 | 0,07 | 0,08 | 0,16 |
| | Pr75 | 0,05 | 0,06 | 0,06 | 0,05 | 0,04 | 0,06 | 0,07 | 0,15 |
| | Pr50 | 0,11 | 0,11 | 0,11 | 0,10 | 0,03 | 0,05 | 0,06 | 0,14 |
| | Pr0 | 0,22 | 0,23 | 0,22 | 0,19 | 0,02 | 0,03 | 0,04 | 0,12 |

## Conclusions

This study, based on probabilistic methods and on Bayes' Theorem for conditional probability, made it possible to obtain information on the degree of consistency between the criteria for determining the sanitary status of shellfish harvesting areas and the microbiological safety of live shellfish put on the market, adopted in the EU.

The results are subject to the formulated hypotheses and the stated study conditions, which are: the samples are randomized, representative of the sampled population and protected from secondary contamination; the outcomes of MPN method virtual tests are independent of each other and related to the prevalence of the hypothesised contamination levels; in the sanitary classification process, the sampling points represent the worst case scenario for a given harvesting area, however it is assumed that the contamination level of the sampling point mirrors that of the whole area; the safety criterion virtual tests are performed on shellfish not exposed to further contamination once put on the market and representative of all classified harvesting areas.

Taking all these conditions into account, the results showcase a limited consistency between the two criteria. In fact:

1) Assuming a high purification's efficacy rate, the safety criterion would be compliant in high percentage only with low contamination levels; in the other cases, which could be more frequent in real conditions, the rate of compliancy would be significantly lower. Therefore, even though the shellfish are legitimately put on the market, a non-negligible number of non-compliant test outcomes could be expected when the safety criterion is applied.

2) The probability that compliant shellfish are coming from Class A areas, or equivalent, is directly related to the effectiveness of the purification treatment. Nevertheless, it seems that shellfish from class "A" areas, or equivalent, are most probable to be the origin of non-compliant results, while shellfish not effectively purified would account for only a limited fraction of these results. This goes to show that a non-compliant outcome is not, *ipso facto*, indicative of the original sanitary state of the shellfish.

3) When the contamination level is low, the frequency of false-non-compliant outcomes arises to 10%. In opposite conditions, when contamination level is higher, the frequency of false-compliant outcomes could exceed even 20%: even in this case, the safety criterion is not closely related to the sanitary status of shellfish harvesting area and it is able to give us a false image of compliance.

This framework of limited consistency could be explained by the variability of MPN method and by the changeability of *E. coli* distribution among the shellfish in the same natural bed. Nevertheless, in the opinion of the authors, the main reason could be attributed to the different sample size implemented by the two criteria.

Finally, in the opinion of the authors, while the criterion used to determine the sanitary status of harvesting areas is effective and, as recently highlighted by EFSA Scientific Report (EFSA, 2019), has contributed to a reduction of Noroviruses in shellfish, this study showed that the *E.coli* safety criterion is not an adequate test to properly determine the origin of shellfish and the limited consistency with the classification criteria raises doubts about its usefulness and would require a serious consideration in evaluating its results.

## References

Belliaeff B, Cochard M-L, 1995. Applying geo-statistics to identification of spatial patterns of faecal contamination in a mussel farming area (Havre de la Vanlée, France). Wat. Res. 29:1541-1548.

Ciccarelli C, Semeraro AM, Di Trani V, Leinoudi M, Ciccarelli E, 2017. Statistical assessment of MPN test method and three class sampling plan used to determine Escherichia coli contamination level in shellfish. In AAVV. Book Abstract of 12th International Conference of Molluscan Shellfish Safety. Ensenada (Mexico). pp 58.

EFSA, 2019. Scientific report on analysis of the European baseline survey of Norovirus in oyster. EFSA J.

European Commission, 2015. Regulation of the European Commission of 8 December 2015 amending Annex II to Regulation (EC) No 854/2004 of the European Parliament and of the Council laying down specific rules for the organisation of official controls on products of animal origin intended for human consumption as regards certain requirements for live bivalve molluscs, echinoderms, tunicates and marine gastropods and Annex I to Regulation (EC) No 2073/2005 on microbiological criteria for foodstuffs. In: Official Journal, L 323/2, 9/12/2015.

European Commission "Community Guide to Principles of Good Practice for the Microbiological Classification and Monitoring of Bivalve Mollusc Production and Relaying Areas with regard to Regulation 854/2004". 2017:3 URL= <https://ec.europa.eu/food/system/files/2018-12/biosafety_fh_guidance_community_guide_bivalve_mollusc_monitoring_en.pdf>.

European Commission, 2019. Regulation of the European Commission of 15 March 2019 laying down uniform practical arrangements for the performance of official controls on products of animal origin intended for human consumption in accordance with Regulation (EU) 2017/625 of the European Parliament and of the Council and amending Commission Regulation (EC) No 2074/2005 as regards official controls. In: Official Journal, L 131/51, 17/05/2019

Gerhardt W, Keller H, 1986. Evaluation of Test Data from Clinical Studies. Scandinavian J Clinical Lab Investigation. 181:1-74.

Joyce J, 2019. Bayes' Theorem In: Zalta EN, Spring 2019 Edition. The Stanford Encyclopaedia of Philosophy. URL = <https://plato.stanford.edu/archives/spr2019/entries/bayes-theorem/>.

Lart WJ, Hudson SA, 1993. Factors affecting Escherichia coli levels in shellfish with reference to E.E.C. Directive 91/492. In: AAVV, 1993 Seafish Report No 419. Sea Fish Industry Authority, St Andrews Dock, Hull HU3 4QE, England.

Lee R, Murray L, 2010. Components of microbiological monitoring programs. In: AAVV, 2010. Safe Management of Shellfish and Harvest Waters: Minimizing Health Risks from Sewage-Contaminated Shellfish. World Health Organisation. JWA Publishing, London. pp 91-108.

Lees DN, Nicholson M, 1995. Faecal coliform in shellfish. In: D. Kay & C. Fricker editors. Coliforms and E. coli – Problems or solutions?. The Royal Society of Chemistry. pp 66-77.

Potasman I, Paz A, Odeh M, 2002. Infectious Outbreaks Associated with Bivalve Shellfish Consumption: A Worldwide Perspective. Clin. Infect. Dis. 35:921-928.

Sharp JH, Clements K, Diggens M, McDonald JE, Malham SH, Jones DL, 2021. E. coli is a poor end-product criterion for assessing the general microbial risk posed from consuming norovirus contaminated shellfish. Frontiers in Microbiology. 12:608888. <doi:10.3389/fmicb.2021.608888>

Younger AD, Lee RJ, Lees DN, 2002. Microbiological monitoring of bivalve

mollusc harvesting areas in England and Wales - Rationale and Approach. In: AAVV, 2003. Proceedings of International Conference for Molluscan Shellfish Safety, Santiago de Compostela, Spain, June 2002. pp. 265-277.

van Elsas JD, Semenov AV, Costa R, Trevors JT, 2011. Survival of E. coli in the environment: fundamental and public health aspects. In: AAVV, 2011. The ISME Journal. Nature Publishing Group. 5:173-183.

Walker DJ, Younger A, Stockley L, Baker-Austin C, 2018. Escherichia coli testing and enumeration in live bivalve shellfish – Present methods and future directions. Food Microbiology. 73:29-38.