Grasas y Aceites 76 (2)
April-June 2025,2195
ISSN-L: 0017-3495, eISSN: 1988-4214
https://doi.org/10.3989/gya.0647241.2195

Inter-laboratory validation of newly formulated reference materials for olive oil sensory analysis

A. Bendini[a], S. Barbieri[a,✉], R. Aparicio-Ruiz[b], A. Koidis[c], K. Brkić Bubola[d], M. Bučar-Miklavčič[e], F. Lacoste[f], U. Tibet[g], O. Winkelmann[h], D.L. García-González[i] and T. Gallina Toschi[a]

[a]Department of Agricultural and Food Sciences, Alma Mater Studiorum - Università di Bologna, Bologna, Italy
[b]Department of Analytical Chemistry, Universidad de Sevilla, Sevilla, Spain
[c]Institute for Global Food Security, Queen's University Belfast, Belfast, UK
[d]Institute of Agriculture and Tourism, Poreč, Croatia
[e]Science and Research Centre Koper, Koper, Slovenia
[f]Institut des Corps Gras, Pessac, France
[g]Ulusal Zeytin ve Zeytinyăgı Konseyi, Izmir, Turkey
[h]Eurofins Analytik GmbH, Hamburg, Germany
[i]Instituto de la Grasa (IG), CSIC, Sevilla, Spain
✉Corresponding author: sara.barbieri@unibo.it

Submitted: 19 June 2024; Accepted: 09 June 2025; Published: 03 February 2026

SUMMARY: Among the available reference materials (RMs) for virgin olive oil (VOO) sensory analysis, it could be very useful to include: i) a suitable training of sensory assessors and ii) the determination of the validity of the sensory evaluation. In this study, an inter-laboratory validation of RMs for the winey-vinegary and rancid defects in VOOs, developed in the framework of the OLEUM project is carried out. RMs were firstly olfactory certified by six panels, subsequently assessed during the pre-trial phase and finally, re-evaluated in twelve sensory laboratories. The RMs were defined as appropriate by the majority of panellists interviewed and the participant panels reached a satisfactory agreement on the intensity of the main perceived defect. The results confirmed that the adoption of the proposed RMs could enhance the tasters' skills in recognizing and quantifying sensory attributes in VOOs by comparing their judgment to the "assigned values" of RMs.

KEY-WORDS: *Rancid defect; Reference materials; Sensory analysis; Taster's training; Virgin olive oil; Winey-vinegary defect.*

RESUMEN: ***Validación interlaboratorio de una nueva formulación de materiales de referencia para el análisis sensorial del aceite de oliva.*** La disponibilidad de materiales de referencia (MRs) para el análisis sensorial del aceite de oliva virgen (AOV) podría ser muy útil para asegurar: i) una formación adecuada de los evaluadores sensoriales, y ii) la determinación de la veracidad de la evaluación sensorial. En este trabajo se ha llevado a cabo una validación interlaboratorios de los MR para los defectos avinado-avinagrado y rancio en los AOVs, desarrollada en el marco del proyecto OLEUM. Los MRs fueron en primer lugar certificados olfativamente por los seis paneles, posteriormente evaluados durante la fase previa al ensayo y, por último, reevaluados por doce laboratorios sensoriales. Los MRs fueron considerados como apropiados por la mayoría de los paneles entrevistados y los paneles participantes alcanzaron un acuerdo satisfactorio sobre la intensidad del principal defecto percibido (coeficiente de variación robusto % de 13,2 para el MR1 y 6,6 para el MR2). Los resultados confirmaron que su adopción podía mejorar la destreza del catador a la hora de reconocer y cuantificar los atributos sensoriales en los VOOs comparando su juicio con los "valores asignados" de los MRs.

PALABRAS CLAVE: *Aceite de oliva virgen; Análisis sensorial; Defecto avinado-avinagrado, Defecto rancio; Entrenamiento de panelistas; Materiales de referencia.*

Citation/Cómo citar este artículo: Bendini A, Barbieri S, Aparicio-Ruiz R, Koidis A, Brkić Bubola K, Bučar-Miklavčič M, Lacoste F, Tibet U, Winkelmann O, García-González DL and Gallina Toschi T. 2025. Inter-laboratory validation of newly formulated reference materials for olive oil sensory analysis. *Grasas Aceites* **76** (2), 2195. https://doi.org/10.3989/gya.0647241.2195

1. INTRODUCTION

In the case of virgin olive oil (VOO), sensory assessment is a standard method and, according to the International Olive Council (IOC) standard (IOC, 2024) and other national and international regulations (EU Regulations, Codex Alimentarius) (EU, 2022a; EU, 2022b; CODEX STAN 33, 1981), its application is mandatory for ensuring that the marketed VOO is consistent with the specific characteristics applicable to declare quality grades (e.g. for EU: extra virgin, virgin, lampante) (EU, 2022a; EU, 2022b).

The method for the sensory analysis of VOOs, commonly called "panel test" by a group of selected, trained and monitored tasters who work as an analytical instrument, was initially developed in 1987 by the IOC and was then adopted by the European Commission in 1991 (EEC, 1991) with a Regulation that has currently been repealed and included in the recently updated EU regulation (EU, 2022a; EU, 2022b). Organoleptic characteristics of VOOs must be certified by tasting panels following numerous procedures (IOC and EU Regulations) necessary to make the methodology standardized and therefore reliable. Although it has been the only recognized standard method that allows the evaluation of sensory attributes (olfactory/retro-olfactory/gustatory), positive and negative (defects) at the same time, it is sometimes criticized for being affected by certain subjectivity as it uses persons rather than equipment.

As discussed in a recent EU report (EC, 2020), the panel test should not be questioned but can be improved in certain aspects. The establishment of reference samples for the different VOOs with which national tasting panels can be trained, and the organization of regular ring tests involving tasting panels from different Member States, are among the proposed solutions aimed at improving the performance of organoleptic assessment (EC, 2020).

The use of RMs is mentioned as necessary for sensory analysis in current IOC regulations (IOC, 2023a; IOC, 2023b). The RMs provide numerical values for specific sensory characteristics and therefore are very useful to train assessors, help them in memorizing sensory stimuli at given intensities and calibrating their judgments. For these reasons, the RMs should be applied for continuous monitoring of panels over time as well as for intra-panel and inter-panel control studies. Moreover, the RMs could be very useful for the selection of new assessors recruited during numerous courses organized at national and international levels and addressed at the specific formation of sensory panelists of VOOs.

At the EU level, organoleptic characteristics of VOOs must be characterized by tasting panels approved by Member States. Usually, each Member State has two types of panels: professional, for the classification and description of the characteristics of VOOs (e.g., PDO certification) and official, for quality control (conformity checks on VOOs, at EU level). Most tasting panels approved by Member States for the purpose of olive oil quality control are also recognized by IOC. IOC recognition is based on the successful results of panel participation in two proficiency tests organized every year, and the IOC supports panels through natural RMs (real olive oil samples), which can be used for panel training.

RMs for VOOs can be samples from IOC proficiency tests (with a clear indication of the type and the intensity of fruitiness and the main defect) or pre-tested (characterized) samples that have been evaluated by comparison with at least three accredited and/or IOC recognized sensory panels (IOC, 2023a; IOC, 2023b). These samples are effective for training regarding defects, although they have some limitations in their use over time since they are obtained from a natural matrix (e.g., limited availability, difficult to obtain, low homogeneity year after year) (Barbieri *et al.*, 2021). Using certified reference materials for taster training as a tool for diminishing discrepancies among panels or taster results was also suggested by Ortega-Gavilán *et al*., (2022), who proposed a protocol for the certification of real VOO samples as RM for sensory analysis. Homogeneity assessment of real VOO samples used as certified reference materials was also studied by instrumental analysis application (i.e., fingerprint methodology) (Ortega-Gavilán *et al*., 2020). An alternative to the RMs produced from real virgin olive oil samples are chemical ones; in fact, RM can be any chemical or natural material that adequately represents a particular sensory characteristic (Krasner, 1995).

In this context, part of the activities carried out within the OLEUM project (http://www.oleumproject.eu/) aimed at the improvement of the method for VOO organoleptic assessment. In fact, one of the main objectives of the project was to develop new artificial reference materials to be shared among the panels (IOC and EU) to improve their performance and to reinforce the panel test method.

Therefore, in the framework of the OLEUM project, an analytical procedure for implementing olfactory formulations that emulated rancid and winey-vinegary defects found in VOOs was developed. The aim was to provide reproducible RMs that can be prepared on demand. These two RMs were obtained by diluting compounds responsible for the winey-vinegary defect (e.g., acetic acid and ethanol) and rancid defect (e.g., hexanal) in combination with other molecules in order to avoid non-natural sensory notes (Aparicio-Ruiz *et al*., 2020).

Subsequently, a study on the practical application and test certification of the two new formulated aroma RMs was performed by sensory analysis (Barbieri *et al*., 2021). Considering the required properties (ISO, 2015a; ISO, 2015b; ISO, 2017) and some other key issues (appropriateness of selected materials, homogeneity in testing, preparation, and stability during distribution and storage) for the preparation of RMs (EA, 2003; Karambelkar, 2018), the authors (Barbieri *et al*., 2021) developed a protocol for the suitability of RMs.

The present study represents the conclusion of the activities carried out on RMs to validate their performance in the OLEUM project; specifically, the two artificial olfactory RMs, for rancid and for winey-vinegary defects, were selected from those previously evaluated (Barbieri *et al*., 2021) and assessed in an inter-laboratory study performed with the collaboration of twelve internationally recognized panels for validation of their appropriateness in resembling the natural defects (compared to authentic VOO samples with corresponding defects) and their intensity values.

2. MATERIALS AND METHODS

2.1. Sensory panels

Twelve public/private laboratories from Europe and the USA (coded as LAB1-LAB9, and LAB11-LAB13), with trained sensory panellists for virgin olive oil analysis (the number ranged from 8 to 12) recognized by national/international control bodies, participated in this international study. The participating laboratories were comprised of 4 from Italy, 2 from Germany, and 1 each from Croatia, France, Greece, Portugal, Slovenia, and the USA. Eleven panels were recognised by the IOC and most of them were also accredited according to ISO/IEC 17025 (2017), nine of them were official panels at the national level, and one was an internal panel from a private laboratory.

The DISTAL Italian panel (OLEUM partner and method developer, coded as LAB 10) assessed RMs and samples but, being responsible for selecting samples, organization of trial tests and elaboration of data, did not contribute with its data to the results of interlaboratory validation.

Each participating panel evaluated 2 reference materials (RM1, Reference Material 1 resembling the winey-vinegary defect; RM2, Reference Material 2 resembling the rancid defect) and 10 test materials (2 training VOO samples and 8 validation VOO samples, described below in details).

2.2. Reference materials

Reference materials, RM1 and RM2, properly formulated by adding specific chemical compounds to a refined olive oil as described by Aparicio-Ruiz *et al*. (2020), have already been olfactory tested at different dilutions by the six sensory panels involved in the OLEUM project. Specifically, a protocol for the suitability of RMs was defined, shared, and applied by sensory panels; it consisted of three steps: i) evaluation of representativeness of RMs (compared to actual samples) and of its intensity values; ii) determination of the panel's detection threshold and; iii) evaluation of long-term stability (Barbieri *et al*., 2021).

Taking into consideration the results from Barbieri *et al*., (2021) and the information obtained by a pre-trial, a specific dilution using the same methodology applied by panels for the detection threshold of the group of candidates for characteristic attributes (IOC, 2023a) was prepared of each RM stock solution and then selected for the trial (dilution 5 and 6 for RM1 and RM2, respectively, characterized by intervals of intensities close to the border line between the VOO and LOO quality grades).

Moreover, during the pre-trial (data not shown), different oil samples (already certified in the previous work involving 6 panels) were proposed to the 12 panels to select two authentic, i.e. "natural", virgin olive oils (coded SA-6 and SA-13) with intensities of winey-vinegary and rancid defects similar to the proposed RMs (dilution 5, for RM1 and dilution 6, for RM2).

2.3. Methodology applied and the VOO samples

As training samples, two additional authentic VOOs (SA-6, and SA-13), previously evaluated by the developer and organizer of the trial (panel DISTAL), were provided to the participants. Their reference intensity values for classifying sensory attributes (fruitiness and main perceived defect, m.p.d., as median values) and quality grade according to EU Regulation (EU, 2022a; EU, 2022b) were also supplied (SA-6: VOO, fruitiness 1.7, m.p.d. 2.6 winey-vinegary; SA-13: LOO, fruitiness 1.5, m.p.d. 3.6 rancid).

Finally, four authentic VOO samples (bulk test materials coded K, F, V, J) were supplied to panels as they contained defects linked to fermentation (as winey-vinegary) or oxidative (as rancid) degradations. Each one of these four bulk test materials was split into two aliquots to obtain replicates of samples. Therefore, each participating laboratory received 8 validation samples (blind pairs coded as 1=8, 2=4, 3=11,10=12).

The evaluation of the representativeness of reference materials (RM1 and RM2), in terms of an assessment by the panellists of their ‘usefulness in resembling natural sensory defects of winey-vinegary and rancid’ (i.e., each assessor providing a binary answer useful/not useful), was the first objective of the study. Each assessor was asked to indicate on a non-structured 10-cm scale the intensity of the defect (perceived only by smelling) related to the RMs, using, as a reference, the known intensity of winey-vinegary (2.6) of the sample SA-6 and the known intensity of rancid (3.6) of the SA-13, thus confirming or not the perceived intensity of the defect in both training samples. The same procedure was also applied for the sensory evaluation of the validation samples.

RMs and test materials were stored at around 4 °C until the day of analysis, and prior to analysis, they were kept at room temperature for at least 1 h (or for the time needed for assuring a complete solubilization of possible glycerides and crystallized waxes) and homogenized by shaking carefully.

3. RESULTS AND DISCUSSION

3.1. Usefulness of the RMs

The first step was the collection of feedback on the feasibility of the use of these two RMs as a support in the identification and evaluation of winey-vinegary and rancid defects in VOOs, as well as on the highlighting of eventual problems and critical issues that can help to further improve the RMs. These RMs have already been olfactory tested with seven different dilutions by the six sensory panels involved in the OLEUM project (Barbieri *et al*., 2021), during the pre-trial phase thus, in the context of the full international study, a possible confirmation of their appropriateness was requested.

The usefulness of the RMs was “measured” as a percentage based on responses of panellists for both RMs: specifically, the numerical values were calculated considering the answers of assessors participating in each panel and values were split into four intervals: 0-25, 26-50, 51-75, 76-100%. Finally, the frequency of values in each interval was determined.

For RM1, eleven panels were within the interval 76-100% in terms of “usefulness” and one was in 0-25%. Ten out of twelve panels were within the interval 76-100% and two in the interval 51-75% for RM2.

These percentages demonstrated that the reference materials were valued as appropriate by a high proportion of end-users (83 and 93 positive answers out of a total of 106 tasters interviewed for RM1 and RM2, respectively) (Table 1).

3.2 Perceived intensities of RMs

Another important outcome was the determination of perceived intensities of the winey-vinegary and rancid defects for RM1 and RM2, respectively. This determination was carried out because the definition of the intensity of each sensory RM is fundamental information to permit its possible future use in the training of assessors. Intensities for RM1 and RM2, expressed as “median of medians” (GM, great median), were 3.9 and 5.8 (Table 1), and calculated as “mean of medians” were 3.8 and 5.6.

TABLE 1. Usefulness of the RMs: Positive answers from 106 assessors of 12 participating sensory panels.

RMs	Usefulness	Intensity (GM)	s*	CV_r%	C.I upper C.I. lower	Reference values
RM1 Winey-vinegary	83	3.9	0.5	13.2	4.9 2.9	2.5
RM2 Rancid	93	5.8	0.4	6.6	6.6 5.0	3.0

Data of intensities expressed as great median (GM) of 12 sensory panels (total of 106 assessors), robust standard deviation (s*), robust coefficient of variation % (CVr%), confidence intervals (C.I.) of the great median at 95%. Reference values (dilutions number 5 and 6 for RM1 and RM2, respectively) obtained from previous sensory evaluations within the OLEUM project.

Regarding the lower and upper intervals of confidence, the intensities can be considered between 2.9 and 4.9 for RM1 and from 5.0 and 6.6 for RM2. As shown in Table 1, the statistical indices of data dispersion were satisfactory (robust coefficient of variation lower than 20%, low values of robust standard deviation). The median values measured for the same dilutions of RM1 and RM2 by the six OLEUM panels (reference values) were 2.5 and 3.0, respectively (Barbieri *et al*., 2021). In the case of the RM1 (winey-vinegary defect) the intensity previously established intra the OLEUM laboratory trial was closer to the results of the same trial than those obtained for RM2 (rancid defect) for which the reference value was significantly lower. A possible incremental effect of the natural rancidity of the refined olive oil used in the RMs to dissolve the pure molecules due to the storage and shipment steps (despite the application/recommendation of proper storage/shipping conditions) cannot be excluded.

The same method adopted by the IOC for its proficiency test (IOC z-score) was applied for checking the alignment among panels with respect to the intensity of defect in each RM. It was calculated using: i) the median (Me) of the predominant defect (winey-vinegary in RM1 and rancid in RM2) ii) the great median (assigned value, GM) calculated as median of the medians (detected by all panels as consensus value) iii) the standard deviation (σ obj) of the scores calculated from IOC historical data (± 0.7) (IOC, 2023a).

The results of the z-score estimation are illustrated by the quality control charts shown in Figure 1. The vertical axis represents the z-score and the horizontal one identifies the sample codes; the z-score can be positive or negative values and was calculated for the intensity of the main defect (winey-vinegary for RM1 and rancid for RM2).

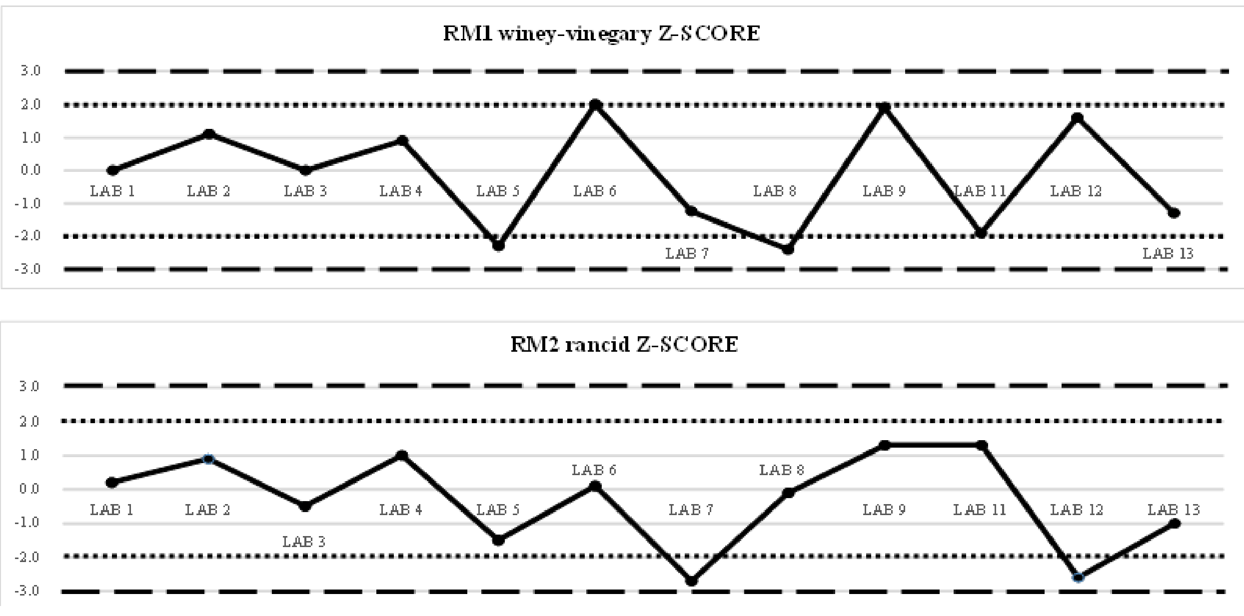

FIGURE 1. Z-score values calculated for each participant laboratory and for each one of the two RMs, applying the IOC formula. The interpretation is: |z| ≤ 2, satisfactory performance; 2 < |z| ≤ 3, questionable performance; |z| > 3, unsatisfactory performance.

For both RM1 and RM2, 10 panels out of 12 evidenced a satisfactory performance (|z| ≤ 2), only 2 questionable (2< |z| ≤3) and not one was unsatisfactory. Therefore, it is possible to extrapolate that the intensity ranges and the GM values for the two RMs can be considered reliable (Figure 1).

In order to establish whether the reference standards proposed for winey-vinegary and rancid defects (RM1 and RM2, respectively) can aid in the recognition of the negative attributes in VOO, as well as to quantify their intensities, the two authentic VOO samples (training samples) characterized in the pre-trial were also evaluated by assessors in the initial trial (coded as SA-6 and SA-13). Thus, these samples were now re-analyzed in this study in parallel with the two RMs under evaluation (RM1 and RM2).

Based on the intensity data provided by the twelve participating sensory laboratories, sample SA-6 was characterised by 2.5 of winey-vinegary and SA-13 by 3.2 of rancid, substantially confirming the results obtained in the pre-trial (Barbieri *et al*., 2021) (Table 2).

TABLE 2. Sensory characteristics of the two training samples (SA-6 and SA-13) analyzed in parallel with RMs for winey-vinegary and rancid as references of kind and intensity of the main perceived defect (m.p.d.) determined by 12 sensory panels in the same trail.

Sample code	Intensity	Quality grade*	m.p.d.	Secondary perceived defects
SA-6	**2.5**	11 V; 1 L	Winey-vinegary	3 Fusty-muddy sediment; 2 Frostbitten; 1 Musty
SA-13	**3.2**	8 V; 4 L	Rancid	

EV: Extra Virgin olive oil; V: virgin olive oil; L: Lampante olive oil (EU, 2022a; EU, 2022b). *the number of panels providing the same quality grade or secondary perceived defect are specified.

Four additional test materials, divided into 2 aliquots (blindly coded replicates) were finally evaluated by the 12 panels as validation samples (blind pair codes: 1 = 8, 2 = 4, 3 = 11, 10 = 12). The twelve participant panels reached a satisfactory agreement on quality grades of samples as well as on the intensity of the m.p.d. (the most complicated case, in terms of the kind of m.p.d. was related to sample number 11, see Table 3), also considering the similarity of the four samples in duplicate. This satisfactory performance is probably due to the fact that the majority of the participating panels were recognized as official panels at national and international level. Therefore, the assessors had a good level of training in the identification of sensory defects and evaluation of their intensities.

TABLE 3 Sensory characteristics (quality grade, kind and intensity of the main perceived defect – m.p.d. expressed as median and mean values) of the validation samples analyzed blindly and in duplicate (1 = 8, 2 = 4, 3 = 11, 10 = 12).

Sample code	Quality grade	Kind of m.p.d.	Intensity of m.p.d. (median)	Intensity of m.p.d. (mean)
1	**9 V**; 2 L; 1 EV	9 **Rancid**; 2 Winey-vinegary	2.0	2.2
2	**8 V**; 4 EV	5 **Winey-vinegary**; 2 Rancid; 1 Fusty/muddy sediment	1.4	1.7
3	**10 V**; 2 EV	8 **Rancid**; 2 Fusty/muddy sediment	1.6	1.8
4	**9 V**; 3 EV	7 **Winey-vinegary**; 1 Rancid; 1 Musty	1.6	1.6
8	**10 V**; 1 L; 1 EV	10 **Rancid**; 1 Musty	2.1	1.7
10	**9 V**; 3 EV	8 **Rancid**; 1 Others	2.0	2.0
11	**10 V**; 2 EV	5 **Rancid**; 3 Musty; 1 Fusty/muddy sediment; 1 Winey-vinegary	2.1	1.9
12	**9 V**; 3 EV	9 **Rancid**	2.3	2.3

EV: Extra virgin olive oil; V: Virgin olive oil; L: Lampante olive oil (EU, 2022a; EU, 2022b).

The control of the panel's precision was performed by using replicate analyses. The repeatability of panels was controlled by comparing the medians obtained from validation samples in duplicate (1 = 8, 2 = 4, 3 = 11, 10 = 12) and determining whether the results were homogenous and, therefore, statistically acceptable. Specifically, the level of agreement between intensity values expressed for same sample during independent evaluations was estimated by calculating the precision number (PN) and normalized error (En), whose acceptability limits are ≤ 2 and ≤ 1, respectively (IOC, 2023a) (Table 4).

TABLE 4. Values of normalized error (En) and precision number (PN) of each panel for the main perceived defect calculated on the four pairs of samples (1 = 8, 2 = 4, 3 = 11, 10 = 12) valued in the replicate analysis (blind conditions).

	1 = 8		*2 = 4*		*3 = 11*		*10 = 12*	
	En	**PN**	**En**	**PN**	**En**	**PN**	**En**	**PN**
LAB 1	0.4	0.6	0.1	0.0	0.4	0.6	0.7	2.1
LAB 2	0.5	0.8	0.1	0.0	0.1	0.0	0.6	1.4
LAB 3	0.2	0.1	0.2	0.2	0.3	0.3	0.9	3.2
LAB 4	1.2	5.3	0.0	0.0	0.1	0.0	0.2	0.1
LAB 5	0.4	0.5	0.2	0.1	0.1	0.0	0.2	0.1
LAB 6	0.4	0.6	0.0	0.0	0.5	1.0	0.3	0.3
LAB 7	0.9	2.9	0.0	0.0	0.0	0.0	0.0	0.0
LAB 8	1.0	3.8	0.7	2.0	0.0	0.0	0.0	0.0
LAB 9	0.8	2.3	0.1	0.0	0.5	0.8	0.1	0.0
LAB 11	0.6	1.2	0.0	0.0	0.3	0.4	0.3	0.4
LAB 12	0.1	0.0	0.2	0.1	0.1	0.0	0.0	0.0
LAB 13	1.0	3.6	0.1	0.0	0.0	0.0	0.3	0.4

These indices are based on the evaluation of the correct intensity of the mean perceived defect (and therefore the identification of the quality grade) by each panel and do not consider the type of defect. En evidenced statistically homogenous results whereas PN values were more frequently (5 out of 12) high in the case of sample pairs 1-8 (Table 4).

In general, the results of the panels were consistent (same classification of samples in quality grade), although the pairs of samples 1 & 8 and 10 & 12 presented errors (values higher than the limits of acceptability), in five (LAB 4, 7, 8, 9, 13) and two (LAB 1 and 3) panels.

4. CONCLUSIONS

The panel test method has effectively contributed to the improvement of the quality of VOOs on the market in the last 30 years and to a reduction in the prevalence of fraudulent practices but there is still a need for panel standardization tools, such as sufficiently stable and validated reference materials. In this context, the OLEUM project "Advanced Solutions for Assuring Authenticity and Quality of Olive Oil at Global Scale", funded by the European Commission within the Horizon 2020 Programme, has developed two artificial olfactory reference materials formulated to resemble specific sensory defects (winey-vinegary and rancid) that are here rigorously validated to determine their applicability and to certify them for their use.

The results illustrated herein highlight that the two newly formulated olfactory RMs were effective in simulating the winey-vinegary and rancid defects found in VOOs and therefore they can be adopted to improve the training of panellists. In fact, the proposed RMs were defined as appropriate by the majority of the sensory panels interviewed (eleven out of twelve panels for RM1 and ten for RM2 were within the interval 76-100%, in terms of "usefulness"). Moreover, the participant panels reached a great agreement on quality grades of samples, as well as on the intensity of the main perceived defect (robust coefficient of variation % of 13.2 for RM1 and 6.6 for RM2), thus demonstrating a satisfactory performance ($|z| \leq 2$ for 10 out of 12 panels).

These two RMs could be used as reference samples for the winey-vinegary and rancid defects and also for different quality grade (virgin, lampante) VOOs, if provided at different intensities. The availability of RMs properly selected and tested could also encourage the cooperation between tasting panels (ring test intra/inter member state) in harmonization activities and facilitate the performance of panels, representing useful tools for panel quality control.

In the future, in view of a possible commercialization of these two olfactory reference materials, it would be useful to test different dilutions and study their stability over time to provide precise indications about their shelf-life and the best conditions of storage and use.

It is hoped that this work will pave the way for research that could lead to the commercialization of sensory reference materials also related to other fundamental defects such as, for example, musty and fusty-muddy, both artificially produced and for olfactory use only, like those validated in this work, and “natural” (i.e generated by biotechnology in virgin olive oils). Again, the main challenge will be the reproducibility, the possibility of standardization and their shelf life.

FUNDING SOURCES

This work was supported by the Horizon 2020 European Research project OLEUM “Advanced solutions for assuring the authenticity and quality of olive oil at a global scale”, which received funding from the European Commission within the Horizon 2020 Framework Programme (2014-2020), grant agreement No. 635690. We are grateful to all panel members who performed sensory analysis of virgin olive oils from each institution involved.

DECLARATION OF COMPETING INTEREST

The authors of this article declare that they have no financial, professional or personal conflicts of interest that could have inappropriately influenced this work.

AUTHORSHIP CONTRIBUTION STATEMENT

A. Bendini: Conceptualization, Formal analysis, Investigation, Methodology, Writing—review & editing. S. Barbieri: Conceptualization, Formal analysis, Investigation, Methodology, Writing—original draft. R. Aparicio-Ruiz: Writing - review and editing. A. Koidis: Formal analysis, Writing—review & editing. K. Brkić Bubola: Conceptualization, Formal analysis, Investigation, Methodology, Writing—original draft. M. Bučar-Miklavčič: Conceptualization, Formal analysis, Investigation, Methodology, Writing—original draft. F. Lacoste: Conceptualization, Formal analysis, Investigation, Methodology. U. Tibet: Conceptualization, Formal analysis, Investigation, Methodology. O. Winkelmann: Conceptualization, Formal analysis, Investigation, Methodology. D. L. Garcia-Gonzalez: Conceptualization, Formal analysis, Investigation, Methodology, Writing—review & editing. T. Gallina Toschi: Writing—review & editing, Funding acquisition, Project administration.

RESEARCH DATA POLICY

DATA AVAILABILITY

Bendini, Alessandra; Barbieri, Sara; Aparicio-Ruiz, Ramon; Koidis, Anastasios; Brkić Bubola, Karolina; Bučar-Miklavčič, Milena; Lacoste, Florence; Tibet, Ümmühan; Winkelmann, Ole; García-González, Diego Luis; Gallina Toschi, Tullia (2024) OLEUM project. Data about inter-laboratory validation of newly formulated reference materials for olive oil sensory analysis. University of Bologna. DOI https://doi.org/10.6092/unibo/amsacta/7668. [Dataset].

REFERENCES

Aparicio-Ruiz R, Barbieri S, Gallina Toschi T, García-González DL. 2020. *Foods* **9**, 1870. https://doi.org/10.3390/foods9121870

Barbieri S, Aparicio-Ruiz R, Brkic Bubola K, Bucar-Miklavcic M, Lacoste F, Tibet U, Winkelmann O, Bendini A, Garcia-Gonzalez DL, Gallina Toschi T. 2021. *Int. J. Gastron. Food Sci*. **25**, 100402. https://doi.org/10.1016/j.ijgfs.2021.100402

Codex. 1981. Codex stan 33. Adopted in 1981. Revised in 1989. 2015, 2017. Amended in 2009, 2013.

EA. 2003. European Cooperation for Accreditation. EA-4/14 INF:2003. The selection and use of reference materials, 2003.

EC. 2020. European Commission. Study on the implementation of conformity checks in the olive oil sector throughout the European Union. Publication office of European Union.

EEC. 1991. Official Journal of the European Communities. European Community, Commission Regulation 2568/91 on the Characteristics of Olive Oil and Olive Residue Oil and on the Relevant Methods of Analysis; Official Journal of the European Communities: Brussels, Belgium, 1991; Vol. L248, 1–83.

EU. 2022a. Official Journal of the European Union, Commission Delegated Regulation (EU) 2104/2022 of 29 July 2022 supplementing Regulation (EU) No 1308/2013 of the European Parliament and of the Council as regards marketing standards for olive oil, and repealing Commission Regulation (EEC) No 2568/91 and Commission Implementing Regulation (EU) No 29/2012; Official Journal of the European Union: Brussels, Belgium, 2022; Vol. L284, 1–22.

EU. 2022b. Official Journal of the European Union, Commission Implementing Regulation (EU) 2105/2022 of 29 July 2022 laying down rules on conformity checks of marketing standards for olive oil and methods of analysis of the characteristics of olive oil; Official Journal of the European Union: Brussels, Belgium, 2022; Vol. L284, 23–48.

IOC. 2024. International Olive Council. Sensory analysis of olive oil - Method for the organoleptic assessment of virgin olive oil, T.20/Doc. No 15/Rev. 11, 1-19.

IOC. 2023a. International Olive Council. Guide for the selection, training and quality control of virgin olive oil tasters-qualifications of tasters, panel leaders and trainers, T.20/Doc. no.14/Rev.8, 1-24.

IOC. 2023b. International Olive Council. Guidelines for the Accomplishment of the Requirements of the Norm ISO 17025 by of Sensory Testing Laboratories with Particular Reference to Virgin Olive Oil, T.28/Doc. No 1/Rev. 8, 1-8.

ISO. 2015a. ISO Guide 30. Reference Materials - Selected Terms and Definitions. Geneva: International Organization for Standardization.

ISO. 2015b. ISO Guide 31. Reference materials - Contents of certificates, labels and accompanying documentation. Geneva: International Organization for Standardization.

ISO. 2017. ISO Guide 35. Reference Materials - Guidance for characterisation and assessment of homogeneity and stability. Geneva: International Organization for Standardization.

ISO/IEC 17025. 2017. General requirements for the competence of testing and calibration laboratories. Geneva: International Organization for Standardization.

Karambelkar N. 2018. *Certified Reference Materials a few guidelines*. Cutting Edge, 28-33.

Krasner SW. 1995. *Wat. Sci. Tech*. **31**, 265-272. https://doi.org/10.1016/0273-1223(95)00486-7

Ortega-Gavilán F, García-Mesa JA, Marzal-Fernández JC, Moreno-Ballesteros FJ, Rodríguez-García FP, González-Casado A, Cuadros-Rodríguez L. 2022. *Food Chem.* **380**, 132195. https://doi.org/10.1016/j.foodchem.2022.132195

Ortega-Gavilán F, Valverde-Soma L, Rodríguez-García FP, Cuadros-Rodríguez L, Gracia Bagur-González M. 2020. *Food Chem.* **322**, 126743. https://doi.org/10.1016/j.foodchem.2020.126743