Accepted Manuscript

Number of consumers and days of display necessary for the assessment of meat colour acceptability

Rodrigo Augusto Cortêz Passetti, Virginia Celia Resconi, Cihan Çakmakçı, María del Mar Campo, Jackeline Karsten Kirinus, Ludmila Couto Gomes Passetti, Ana Guerrero, Ivanor Nunes do Prado, Carlos Sañudo

PII: S0963-9969(19)30187-5
DOI: https://doi.org/10.1016/j.foodres.2019.03.036
Reference: FRIN 8345

To appear in: *Food Research International*

Received date: 12 October 2018
Revised date: 13 March 2019
Accepted date: 15 March 2019

Please cite this article as: R.A.C. Passetti, V.C. Resconi, C. Çakmakçı, et al., Number of consumers and days of display necessary for the assessment of meat colour acceptability, Food Research International, https://doi.org/10.1016/j.foodres.2019.03.036

Number of consumers and days of display necessary for the assessment of meat colour acceptability

Rodrigo Augusto Cortêz Passetti[a,b,*], Virginia Celia Resconi[a], Cihan Çakmakçı[a,c], María del Mar Campo[a], Jackeline Karsten Kirinus[a,d], Ludmila Couto Gomes Passetti[e], Ana Guerrero[a], Ivanor Nunes do Prado[b], Carlos Sañudo[a]

[a]*Departamento de Producción Animal y Ciencia de los Alimentos, Instituto Agroalimentario de Aragón (IA2), Universidad de Zaragoza-CITA, Miguel Servet 177, 50013 Zaragoza, Spain.*

[b]*Department of Animal Science, Universidade Estadual de Maringá, Av. Colombo, 5790, Jd. Universitário, Maringá, 87020-900, Paraná, Brazil.*

[c]*Department of Animal Science, Faculty of Agriculture, Van Yuzuncu Yil University, Zeve-Kampus, 65080, Van, Turkey.*

[d]*Universidade Estadual de Santa Catarina, Santa Catarina, Brazil, Rua Beloni Trombeta Zanin, Santo Antônio, Chapecó, 89815-630, Santa Catarina, Brazil.*

[e]*Department of Animal Science, Universidade Federal de Sergipe, Avenida Marechal Rondon, s/n Jardim Rosa Elze, São Cristóvão,49100-000 Sergipe, Brazil.*

**Corresponding author. E-mail address: racpassetti@gmail.com (R. A. C. Passetti). 3011 8931; fax: +55 44 3011 8977.*

Conflict of interest: none

*Corresponding author. E-mail address: racpassetti@gmail.com (R. A. C. Passetti). 3011 8931; fax: +55 44 3011 8977.

ABSTRACT

Visual assessment is regarded as the gold standard to evaluate meat colour shelf-life, but it is costly and time consuming. To address this issue, this paper aims to evaluate the number of consumers and days of display that are necessaries in order to assess the colour shelf-life of meat, presented with different methods, all using images. Photographs of thirty-six lamb steaks were taken just after cutting (day 0) and on each of the following days until the 14th day of display under standardized conditions. Images were presented in three different manners: 1) with days of display and animals in random order (Random); 2) days of display in sequential and animals in random order (Sequential); and, 3) days of display and animals in sequential order (Animal); they were presented to 211 consumers who evaluated visual acceptability on a 9-point scale. At day zero, visual acceptability scores were the highest in Animal, followed by Sequential, and then by the Random (P < 0.05) method. Scores decreased over time for all methods tested (P < 0.05). The Random method presented the highest standard deviation; however, an increase in standard deviation among consumers along days of display was observed for all methods tested (*P* < 0.05). Shelf-life determined by regression varied according to the method of presentation (7.83, 7.00 and 7.54 days for Random, Sequential and Animal, respectively). A minimum number of 4 day points before and 4 day points after neutral scores had been reached (scores = 5.0) were necessary in order to obtain a robust model. The minimum number of required consumers (α = 0.05; d = 0.1 and β = 0.2 or 0.1) varied according to methodology: it was 81 to 109 consumers for Random, 69 to 92 for Sequential, and 55 to 74 for Animal. Our study indicates that an optimal number of days and evaluators can be calculated depending on the manner of sample presentation. These findings should be taken into account in further studies that aim to balance data reliability with the cost involved in meat colour analyses.

Keywords: digital images; freshness; lambs; redness; sensory; shelf-life

1. Introduction

More than any other sensorial attribute, the colour of meat bears a decisive influence on consumers' willingness to purchase it (Mancini & Hunt, 2005). During display or storage, the red-cherry colour on the steak surface changes to brown, which consumers regard as less acceptable. Therefore, colour is strongly related to commercial shelf-life (Passetti et al., 2017; Prado et al., 2015; Warren et al., 2008). Although colour can be objectively determined using simple and non-destructive methods, such as the physical CIE L*a*b* space method, studies with consumers are still necessary (Holman, Mao, Coombs, van de Ven, & Hopkins, 2016; Yang et al., 2016). Visual assessments are costly and time-consuming, but they are still regarded as the gold standard in order to estimate consumer perception and to help define the shelf-life and acceptability of a specific product within a given market and time period (AMSA, 2012; Mancini & Hunt, 2005). Such consumer studies are nevertheless complex, expensive and time consuming.

Shelf-life assessments can be designed with basic or reversed methodologies, among other approaches (Passetti et al., 2017). Basic (or sequential) storage design is the simplest and most common one used to perform sensory shelf-life experiments in foods: consists of evaluations of a single batch of samples at various storage times (Hough, 2010). The drawback, however, can be that evaluators become aware of the aim of the experiment and start expecting that samples will become more deteriorated with the passage of time: this, in turn, can lead to biased results (Østli, Esaiassen, Garitta, Nøstvold, & Hough, 2013; Passetti et al., 2017). Other disadvantages of this methodology include: availability on the part of evaluators for attending sessions on several fixed days; potential modification of criteria when evaluating the samples over time, plus the additional cost involved if the evaluators are remunerated (Giménez, Ares, & Ares, 2012). Reverse design, on the other hand, consists in evaluating a set of samples randomly distributed according to storage times. They can all be assessed together in a single evaluation session, thereby minimizing the effort and resources required for carrying out the experiment (Hough, 2010). However, reverse design requires homogeneous samples for every storage time, and meat is highly variable (even within a muscle from a single animal). In the particular case of visual assessment, this problem could be addressed by asking the subjects to evaluate photographs instead of the direct visualization of the samples (Passetti et al., 2017). This allows for reverse design using the same steaks stored at different times, which represents an additional advantage in the case of a limited and expensive product such as meat. Furthermore, by using photographs, measurement conditions can be more optimally standardized, and the images can be saved to reproduce the study in different locations or for other purposes.

Another methodological lies in the attempt to estimate the optimum number of consumers and days of display necessaries to carry out these studies, while at the same time limiting the cost and time factors without compromising robustness of results. Recent studies of hedonic visual shelf-life evaluation in ruminant meat (Table 1) have resorted to trained, semi-trained or untrained evaluators ranging from 6 to 56 people. Some studies analyze the samples daily, every two, three and four days, and the period of study can range from 8 to 20 days of display (Eiras et al., 2017; Guerra-Rivas et al., 2016; Langroodi et al., 2018; Passetti et al., 2017; Possamai et al., 2018; Prado et al., 2015; Ripoll, Alcalde, Arguello, Cordoba, & Panea, 2018). The numbers of evaluators in these tests is lower than the figures recommended by textbooks for hedonic tests using consumers (AMSA, 2012; Giménez et al., 2012). The difficulty in increasing the number of evaluators could be associated with the relatively long interval of time until discolouration sets in, thereby implying a low availability of consumers when basic designs are used. To address this problem, evaluations using digital images could represent a more practical alternative that could help increase the number of persons, since the images can be viewed at any time, and within a shorter time period (Holman et al., 2016).

In acceptability methodology studies, a way of calculating the optimal number of consumers based on standard error data has been proposed (Hough et al., 2006). However, the standard error may differ according to product type, attribute measured (appearance *versus* eating attributes such as flavour), and design approach. Regarding the points chosen for evaluation in shelf-life studies, a total of seven times equally distanced from one another has been suggested, including the initial point regarded as the fresh sample and with shelf-life estimations in the middle of the total time period evaluated (Hough, Calle, Serrat, & Curia, 2007). However, until now, a comparison of results obtained by modifying the number of time points had not been carried out.

This study presented herein was performed to calculate the optimum number of consumers and days of display necessary to evaluate lamb meat shelf-life using three different manners of digital image presentation.

2. Material and methods

2.1. Local conditions and ethical compliance

Lambs were raised at the Animal Experimentation Service of the University of Zaragoza (SEA), Spain (latitude 41º41′N) during autumn of 2015. The area is located in the Ebro Valley, which possess a dry Mediterranean climate with an average annual temperature of 15 ºC and an average annual rainfall of 317 mm. This study was performed in accordance with the recommendations for the care and use of experimental animals laid out by the University of Zaragoza (R.D. 53/2013).

2.2. Animals, slaughter, and meat sampling

A total of 36 male and female lambs born from medium wool local breeds and their crosses, at the SEA facilities of the University of Zaragoza were used for this study. After weaning at 38.4 ± 4.3 days of age, animals with average initial body weight of 14.3 ± 3.1 kg were distributed into 16 slatted floor pens. All animals had *ad libitum* access to water, commercial concentrate (with 17.6% of crude protein and 3.8% of ether extract) and cereal straw. At 86 days of age, lambs were slaughtered with an average weight of 22.8 ± 4.2 kg at a licensed commercial slaughterhouse (MercaZaragoza). The day before slaughter, all lambs were transported together for less than 5 km. After slaughter, carcasses remained under cooling for 24 hours. The left rack was removed from T6 to T13 vertebrae and transferred to the Meat Laboratory at the Faculty of Veterinary Medicine of the University of Zaragoza. *Longissimus thoracic* steaks (2.5 mm, 8^{th} - 10^{th} vertebra) were cut and placed individually on plastic trays (12.5 x 10 x 3.6 cm), overwrapped with a retractile oxygen-permeable plastic film (water vapour transmission rate of 5–10 g/m^2.day at 38°C and 90% RH and an O_2 transmission rate of 650–750 cm^3/m^2.day at 25° C and 0% RH; TavFood packing S.L., Spain) and stored in a refrigerator at 4 ºC in the dark for fifteen days.

2.3 Display and photography

Standardized conditions for photography's were prepared (Chan, Moss, Farmer, Gordon, & Cuskelly, 2013; Passetti et al., 2017). Photos were taken at day 0 (immediately after cutting the steak), the following day (day 1), and subsequently on a daily basis until the 14^{th} day. A CANON Power Shot G10 digital camera mounted on a photographic structure (at a height of 45 cm over the samples) under the lighting from two fluorescent tubes was used. Following preliminary tests to guarantee that meat samples appeared in the photos entirety, the camera was set on standard mode without flash with 1/200 shutter speed, F/5, ISO 400 aperture size and 18 mm focal distance. For each sample, one of three photos shots attempts was selected. Images were stored and transferred to computer as JPEG file and recoded to be presented in different orders (Passetti et al., 2017).

2.4. Visual assessment

A total of 540 photographs (16 animals x 15 time points) were presented in three different orders: Random, in which the days of display and animals were presented in randomized fashion; Sequential, in which the days of display were presented in sequential order and the animals were presented in random order; and Animal, in which both the days of display and animals were presented in sequential order. The images were projected in rooms with 20-30 people who attended each day of evaluation. The 540 images were initially evaluated one day using one of the methods and, the following sessions/days (with one week interval among them), using the other methods. The order of presentation of the three methods was balanced. Consumers were instructed to evaluate the acceptability of colour on a hedonic scale from 1 ("I dislike it extremely") to 9 ("I like it extremely"), ignoring other visual aspects such as size, marbling, etc. Each image was presented every 5 s. After 45 evaluations, consumers rested for 1 min. Sixty Spanish, 77 Brazilian and 74 Turkish consumers participated at each country's university facilities.

2.5. Statistics analyses

The one-way ANOVA of consumer-assigned average scores and its respective standard deviation were carried out using the IBM Statistical Package for the Social Science (SPSS version 22). Lamb was considered as a block effect, and days of display and methodology were considered as fixed factors, and interactions were included in the model. Interactions were observed between days of display and methodology; those effects were therefore evaluated by a GLM analysis for each day, and differences among means were assessed by using the Tukey Test ($P < 0.05$). To calculate the numbers of consumers required for future studies, was applied the methodology proposed by Hough et al. (2006): calculate the root mean square error from the analysis of variance divided by the scale length (RMSL) and considering an α level or type I error of 0.05, a β level or type II error of 0.1 and 0.2 and a d level of 0.1 or 0.2, which implies a difference of 10% or 20% that is sought in the experiment, using the GRANMO software. The error mean square was obtained from the global model (considering all methods) and from further ANOVAs conducted within each method of image presentation, considering days as a fixed factor.

To analyse the number of days, regressions were performed between days of display and acceptability scores, evaluating the entire data and reducing days, and subsequently evaluating the implications of those reductions by the R^2 values and the shelf-life (number of days when scores reached 5.0) obtained with the corresponding equations.

3. Results

3.1 Visual acceptability scores and standard deviation

Visual acceptability varied according to the chosen methodology, and decreased along days of display (Table 2). On day zero of display, photos presented in Random method were less accepted than in the Sequential and the Animal methods ($P < 0.05$). Mean values started below 7.0 (I like it moderately) in Random and Sequential on day zero of display, while means close to 8.0 (I like it very much) were observed in the Animal method. Samples were lightly appreciated (scores ≥ 6.0) until Day 5 (Random and Animal) and Day 4 (Sequential) of display, while evaluators neither like or dislike (scores ≥ 5.0) samples until Day 7 in all the methodologies. Mean scores ≥ 4.0 (I dislike it lightly) were observed until Day 10 in Random and until Day 9 in the Sequential and Animal methodologies. Mean values ≥ 3.0 (I dislike it moderately) were observed until Day 13 in the Random, Day 12 in the Sequential and Day 11 in the Animal methodology. Mean scores > 2.0 (I dislike it very much) were observed on the Day 14 in all methodologies.

Standard deviation among consumers are presented in Table 3. The Random method presented the highest standard deviation, followed by the Sequential, then the Animal methodology ($P < 0.05$). Standard deviation increased until Day 7 for Random and Sequential methods (1.16 to 1.58 and 0.68 to 1.08) and until Day 6 for Sequential method (0.65 to 1.09). For all methods standard deviation were similar until Day 13, but then decreased to 1.32, 0.98 and 0.99 for the Random, Sequential and Animal methods on Day 14.

3.2 Effect of methodology on shelf-life determination

The Random and Sequential methods presented an intermediary correlation ($R^2 = 0.576$ and 0.568 respectively) with days of display (Table 4). The Animal method presented the highest correlation ($R^2 = 0.712$) with days of display. Shelf-life was determined by the number of days which scores were = 5.0 (neither like nor dislike). The Random method presented the longest shelf-life, 7.83 and 7.63 days when considering days 0 or 1 as the first day of display period, respectively. The Animal method presented shelf-life of 7.54 or 7.41 days, and the Sequential method shelf-life of 7.00 or 6.87 days when considering days 0 or 1 as the first day of display period, respectively.

3.3 Number of required days

Regression analyses between acceptability scores and days of display were performed, and reductions of the number of days were assessed. The exclusion of Day Zero of evaluation resulted in a reduction of shelf-life determination by 0.13 (Sequential and Animal) and 0.20 (Random) days (Table 4). Taking all days as a reference point, evaluations carried out every two days decreased the shelf-life of the samples by 0.02, 0.20 and 0.13 days for the Random, Sequential and Animal methodologies, respectively. Evaluations every three days reduced shelf-life by 0.28, 0.13 and 0.15 days for the Random, Sequential and Animal methodologies. Evaluations carried out every four days decreased shelf-life by 0.43 days for the Random method, and increased shelf-life by 0.07 days for the Sequential and Animal methodologies. Evaluations every five days increased shelf life by 0.36, 0.14 and 0.27 days for Random, Sequential and Animal methods. Evaluations every six days, increased shelf-life by 0.09 days for the Random methodology and reduced shelf-life by 0.23 and 0.10 for the Sequential and Animal methods. Evaluations every 7 days increased shelf-life by 0.41, 0.17 and 0.16 for the Random, Sequential and Animal method, respectively.

Another way of establishing the optimal number of days of analysis for future studies was performed by regression analysis with display days finishing on the 8th, 9th, 10th, 11th, 12th, 13th and 14th day evaluated (Table 5). Regressions performed until the day after mean scores were lower than 5.0 resulted in low R^2 and shorter shelf-life. An increment of the number of evaluation days gradually increased R^2 values and stabilized after the 11th day.

3.4 Number of required consumers

Table 6 shows that the root mean square error divided by scale length (on a hedonic scale 1-9) was 0.226, 0.208, 0.186 for the Random, Sequential and Animal methods and amounted to 0.207 considering all methodologies together (Global). In our study, differences in means that were sought in the experiment (d) were 10% and 20%, and Type I of error was 0.05, whereas two scenarios for Type II of error (β) of 0.2 and 0.1 were considered. Considering d of 20%, the minimum number of consumers necessary were 21, 18, 14 and 17 for $\beta = 0.2$, and 28, 23, 19 and 23 for $\beta = 0.1$ (for Random, Sequential, Animal and Global, respectively). However, reducing the differences in means expected in the experiment to 10% drastically increased the numbers of consumer's necessaries to 81, 69, 55 and 68 for $\beta = 0.2$, and 109, 92, 74 and 91 for $\beta = 0.1$ (for Random, Sequential, Animal and Global, respectively).

4. Discussion

4.1Visual acceptability scores and standard deviation

Myoglobin, the protein responsible for meat colour, occurs in three forms: deoxymyoglobin, with a purplish-red colour, associated with vacuum-packaged meat or with meat immediately after cutting; oxymyoglobin, presenting a cherry-red colour associated with fresh meat, and metmyoglobin, a brown colour which is associated with spoiled or non-fresh meat (Mancini & Hunt, 2005). According to Brugiapaglia and Destefanis (2009) consumers are able to distinguish meat colour prior to (purplish-red) and after blooming (cherry-red).We hypothesized that acceptability scores would be lower on day zero than on Day 1 of display, because a cherry-red meat colour is considered most desirable by consumers (Carpenter, Cornforth, & Whittier, 2001).

However, this result was only observed when photos were presented in random order, which was also the order in which acceptability was the lowest. Passetti et al. (2017) observed differences between the methods of presentation (Sequential *vs* Random) specially in the first days of evaluation. When consumers know the day of display, they feel more confident in assigning higher scores: they associate the first day of evaluation with the greatest degree of freshness, and this, in turn, biases their assessment of the colour itself. This bias increases when consumers assign even higher scores on day zero when they know they will be evaluating the same meat through all days of display (Animal).

Thus, depending on the method of presentation, first-day scores can be higher or lower. As a consequence, the number of days in which meats receives a score higher than 5.0 points can differ, thereby directly impacting shelf-life determination.. This make the first day of evaluation a critical point for studies using basic storage design, in the case of those that use hedonic scales in which shelf-life is defined with scores equal to 5.0 points, since lower values mean dislikes (Eiras et al., 2017; Passetti et al., 2017; Prado et al., 2015). Merely excluding day zero in basic storage design methodologies would not solve this problem, since, in several countries, meat that has not yet bloomed is still purchased in butchers' shops just after cutting.

The association between colour and freshness varies from consumer to consumer: this can be observed in the standard deviation among the consumers scores (Passetti et al., 2017). In our study, the Sequential and Animal methodologies presented a lower standard deviation than Random presentation:, thus, results are more homogeneous when the additional information concerning the degree of freshness is provided. In contrast to our results, Passetti et al. (2017) did not observe a difference in standard deviation among methodologies for the majority of days evaluated: this could be due to their use of semi-trained evaluators, whereas in the present study we resorted to untrained consumers. Semi-trained consumers might display more uniform criteria of acceptability, thereby leading to lower dispersion amongst themselves (Hough et al., 2006).

Standard deviation increased over time and then reduced in the last days of evaluation: this was similar to the findings observed by other authors (Arnold, Scheller, Arp, Williams, & Schaefer, 1992; Passetti et al., 2017). The decrease in the standard deviation after a certain day of display reflects the agreement among evaluators that meat is no longer acceptable, thus after that point there is no benefit in prolonging the evaluation.

4.2 Method effect on shelf-life determination

The Random method resulted in the most extensive shelf-life, which agrees well with the findings of Passetti et al. (2017). When days were presented in random order, acceptability was higher because consumers associated meat freshness with colour alone, since no additional information was provided. When consumers were provided with information regarding days of display, shelf-life was lower. However, the Animal method resulted in a more extended shelf-life than Sequential, and also presented the highest R^2 value among the three methodologies applied. This could be explained by the higher scores at day zero. The additional information regarding the animal could lead to more confidence in assigning higher value scores. Both the Sequential and Animal methods are possibly biased due to the knowledge provided regarding the number of days of storage. Consumers are conditioned during evaluation to score less than the day before, but higher than the following day. The higher R^2 value observed in our study and by Passetti et al. (2017) for these methods reinforces the hypothesis that consumers evaluation increasingly become less dependent of the colour appreciation in itself. On the other hand, in the Random method, the greater standard deviation observed among consumers could be explained by the greater difficulty they experience in making decisions. The constant changes among samples, and differences such as size, intramuscular fat, etc., might interfere in the evaluation process. The standard deviation is lower in the Animal method, because the same sample is shown throughout display, thus making it easier for consumers to focus solely on the aspect of discoloration.

4.3 Numbers of required days

Evaluation with photos overcomes the difficulties of basic storage design which requires that evaluators come every day to the laboratory(Passetti et al., 2017). In our experience, these methods nevertheless still present some inconveniences in view of the high number of photos to be evaluated, which make it unattractive for consumers to participate. In order to be more practical, the number of photos should be reduced by excluding certain evaluation days, yet without compromising shelf-life results.

Excluding the day zero of evaluation decreased R^2 values (0.023 and 0.018 points) for Sequential and Animal, and increased R^2 value by 0.004 points in Random, while shelf-life determination was reduced by 0.13 (Sequential and Animal) and 0.20 (Random) days. Thus, removing the day zero of displays had 7 times more impact on shelf-life determination in the Random method than in Sequential and Animal methods: a change of 0.001 points in R^2 value reflected a change of 0.050 days of shelf-life in the Random method, whereas a change of 0.001 points in R^2 reflected a change of 0.006 and 0.007 days of shelf-life in the Sequential and Animal methodologies, respectively. In the Sequential and Animal methods, the first day of evaluation was day zero of display (meat just after cutting), while in the Random method it was not. As previously discussed the onset of evaluation is a critical point on basic storage design. Our results demonstrate that day zero is more significant in Random designs. Thus, if studies want to include day zero, the Random order of presentation is the most suitable method in view of consumer purchase habits.

Photos presented every two days increased the R^2 values of 0.018, 0.025 and 0.019 points in the Random, Sequential and Animal methods. Shelf-life determination was only minimally affected in the Random method (0.02 days), whereas it decreased by 0.20 days in the Sequential method and increased by 0.13 days in the Animal method. Our results demonstrate that, in the reverse design model (Random), the number of presented photos could be reduced by the half (8 day points) without compromising the robustness of the regression model. On the other hand, decreasing the number of day points to six or less (photos presented every 3, 4, 5 and 6 days) resulted in significant R^2 values decreases of the regression model, and notable changes in shelf-life determination.

In basic storage design, samples are daily/regularly evaluated, and the inflection point (neutral point) can be estimated based on the results obtained. Therefore, based on the evolution of the scores, researchers can decide on the duration of storage time. Certain previous studies used hedonic scales to evaluate meat displayed for 10 days, and the inflection point (scores < 5.0) occurred after 5-6 days of display (Eiras et al., 2017; Passetti et al., 2017; Prado et al., 2015). Possamai et al. (2018) displayed goat meat for 15 days, and used a purchase intention of 50% to reflect consumer rejection, which occurred after 7 days of display, whereas Ripoll et al. (2018) displayed suckling kid meat for 8 days, and their shelf-life calculation lay between 6 and 8 days. However, deciding upon the storage time for reverse storage designs, for which the inflection point is unknown, can be challenging.

So far no study has provided methodologic recommendations to help decide until when samples must be evaluated: thus, in the case of reverse storage design, researchers must resort to past studies or instrumental measurements as a reference to estimate the duration of storage period. Hough et al. (2007), proposed a total number of 7 points equally distributed along the inflection points, independently of the food product. In agreement with that recommendation, our results showed that a minimum of 8 points are necessary to calculate the shelf-life of meat based on regression analysis (Table 4), including day points with scores lower than 5.0. When decreasing the number of days evaluated up to the 14th day (a total of fifteen day points) to the 8th day (a total of nine day points), a significant reduction of R^2 values occurred (Table 5). This demonstrates that the scores below 5.0, which set in after Day 7 of display (Table 2) play an essential role in providing the regression analysis with robustness, because 5.0 is the inflection point and lower scoresreflect consumer aversion to the product. Thus, we could assume that a minimal number of 4 points below and 4 points above the inflection point should be used to design the visual assessment of meat shelf-life.

4.4 Number of required consumers

The present study was conducted using several samples and resorting to a high number of consumers (211), more than those usually reported in visual acceptability studies of ruminant meat (Table 1). Therefore, we expected to be able to provide a good measure of consumer variability. Our results of the suggested number of consumers suggested (Table 6), especially in the Random method, were in agreement with the findings regarding other types of products and hedonic attributes, including eating quality (Hough et al., 2006), probably because those authors used the Random method in their acceptability studies. In shelf-life studies, however (Table 1), where sequential method or basic design is the most commonly used, as pointed out in the review by Giménez et al. (2012). Since variability is reduced in the Sequential and Animal methods, fewer consumers are necessary for testing. For example, for an $\alpha = 0.05$, $\beta = 0.1$ and $d = 0.1$, in the Random method 109 consumers are necessary and 92 and 74 are needed in the Sequential and Animal method, respectively. As explained in Hough et al. (2006) this number represents the consumers that will test all the samples, as a unique block. If we need to add further factors, such as different nationalities, age, etc., 109 consumers should be used in the different groups to be compared. The problem of limited number of evaluators available at one's disposal at a certain time could be solved by using digital images, since photos can be stored and analyses can be repeated whenever convenient.

5. Conclusions

Photographs presented in random order provide a model design that more realistically reflects consumer acceptance of meat color. This also helps avoid bias introduced by additional information regarding days of display, which plays an especially important role during the first days of evaluation. In order to obtain a robust regression model for the shelf-life determination of meat, it is recommended to have a minimum number of 8 points distributed before and after the inflection point (scores = 5.0). Our findings revealed that the number of evaluators used in previous published studies was lower than the recommendable number of evaluators for a consumer study. The use of digital images offers a practical alternative for future research studies of shelf-life assessment performed by consumers.

Acknowledgements

R.A.C. Passetti and J.K. Kirinus were supportedby a scholarship from CNPq (Conselho Nacional de desenvolvimento científico e tecnológico), Brazil. V.C. Resconi was supported by a contract from the Juan de la Cierva-Incorporación Program (Ministerio de Economía y Competitividad), Spain; and L.C.G. Passetti by a Post-Doc a CAPES Fellowship (Brazil). Authors thanks toJ. Pomed, M. Barahona and J. L. Olleta for their technical assistance.

References

AMSA. (2012). *AMSA Meat Color Measurement Guidelines*. 201 West Springfield Avenue, Suite 1202 Champaign, Illinois USA 61820.

Arnold, R. N., Scheller, K. K., Arp, S. C., Williams, S. N., & Schaefer, D. M. (1992). Visual and Spectrophotometric Evaluations of Beef Color Stability. *Journal of food science, 57* (2), 518-520. http://dx.doi.org/10.1111/j.1365-2621.1992.tb05530.x

Brugiapaglia, A., & Destefanis, G. (2009). Sensory evaluation of meat.t. colour. using photographs. *Italian Journal of Animal Science, 8* (2), 480-482. https://doi.org/10.4081/ijas.2009.s2.480

Carpenter, C. E., Cornforth, D. P., & Whittier, D. (2001). Consumer preferences for beef color and packaging did not affect eating satisfaction. *Meat Science, 57* (4), 359-363. http://dx.doi.org/10.1016/S0309-1740(00)00111-X

Chan, S.-H., Moss, B. W., Farmer, L. J., Gordon, A., & Cuskelly, G. J. (2013). Comparison of consumer perception and acceptability for steaks cooked to different endpoints: Validation of photographic approach. *Food Chemistry, 136* (3-4), 1597-1602. https://doi.org/10.1016/j.foodchem.2012.04.069

Eiras, C. E., Guerrero, A., Valero, M. V., Pardo, J. A., Ornaghi, M. G., Rivaroli, D. C., Sanudo, C., & Prado, I. N. (2017). Effects of cottonseed hull levels in the diet and ageing time on visual and sensory meat acceptability from young bulls finished in feedlot. *Animal, 11* (3), 529-537. http://dx.doi.org/10.1017/S1751731116001749

Giménez, A., Ares, F., & Ares, G. (2012). Sensory shelf-life estimation: A review of current methodological approaches. *Food Research International, 49* (1), 311-325. http://dx.doi.org/10.1016/j.foodres.2012.07.008

Guerra-Rivas, C., Vieira, C., Rubio, B., Martinez, B., Gallardo, B., Mantecon, A. R., Lavin, P., & Manso, T. (2016). Effects of grape pomace in growing lamb diets compared with vitamin E and grape seed extract on meat shelf life. *Meat Science, 116*, 221-229. http://dx.doi.org/10.1016/j.meatsci.2016.02.022

Holman, B. W., Mao, Y., Coombs, C. E., van de Ven, R. J., & Hopkins, D. L. (2016). Relationship between colorimetric (instrumental) evaluation and consumer-defined beef colour acceptability. *Meat Science, 121*, 104-106. https://doi.org/10.1016/j.meatsci.2016.05.002

Hough, G. (2010). *Sensory shelf life estimation of food products*: Crc Press.

Hough, G., Calle, M. L., Serrat, C., & Curia, A. (2007). Number of consumers necessary for shelf life estimations based on survival analysis statistics. *Food Quality and Preference, 18* (5), 771-775. https://doi.org/10.1016/j.foodqual.2007.01.003

Hough, G., Wakeling, I., Mucci, A., Chambers IV, E., Gallardo, I. M., & Alves, L. R. (2006). Number of consumers necessary for sensory acceptability tests. *Food Quality and Preference, 17* (6), 522-526. http://dx.doi.org/10.1016/j.foodqual.2005.07.002

Langroodi, A. M., Tajik, H., Mehdizadeh, T., Moradi, M., Kia, E. M., & Mahmoudian, A. (2018). Effects of sumac extract dipping and chitosan coating enriched with Zataria multiflora Boiss oil on the shelf-life of meat in modified atmosphere packaging. *LWT - Food Science and Technology, 98*, 372-380. https://doi.org/10.1016/j.lwt.2018.08.063

Mancini, R. A., & Hunt, M. C. (2005). Current research in meat color. *Meat Science, 71* (1), 100-121. https://doi.org/10.1016/j.meatsci.2005.03.003

Østli, J., Esaiassen, M., Garitta, L., Nøstvold, B., & Hough, G. (2013). How fresh is fresh? Perceptions and experience when buying and consuming fresh cod fillets. *Food Quality and Preference, 27* (1), 26-34. https://doi.org/10.1016/j.foodqual.2012.05.008

Passetti, R. A. C., Torrecilhas, J. A., Ornaghi, M. G., Mottin, C., Oliveira, C. A. L. d., Guerrero, A., del Mar Campo, M., Sañudo, C., & Prado, I. N. (2017). Validation of photographs usage to evaluate meat visual acceptability of young bulls finished in feedlot fed with or without essential oils. *Meat Science, 123*, 105-111. https://doi.org/10.1016/j.meatsci.2016.09.009

Possamai, A. P. S., Alcalde, C. R., Feihrmann, A. C., Possamai, A. C. S., Rossi, R. M., Lala, B., Claudino-Silva, S. C., & Macedo, F. A. F. (2018). Shelf life of meat from Boer-Saanen goats fed diets supplemented with vitamin E. *Meat Science, 139*, 107-112. http://dx.doi.org/10.1016/j.meatsci.2018.01.011

Prado, I. N., Campo, M. M., Muela, E., Valero, M. V., Catalan, O., Olleta, J. L., & Sanudo, C. (2015). Effects of castration age, protein level and lysine/methionine ratio in the diet on colour, lipid oxidation and meat acceptability of intensively reared Friesian steers. *Animal, 9* (8), 1423-1430. http://dx.doi.org/10.1017/S1751731115000580

Ripoll, G., Alcalde, M. J., Arguello, A., Cordoba, M. G., & Panea, B. (2018). Consumer visual appraisal and shelf life of leg chops from suckling kids raised with natural milk or milk replacer. *Journal of the Science of Food and Agriculture, 98* (7), 2651-2657. http://dx.doi.org/10.1002/jsfa.8758

Warren, H. E., Scollan, N. D., Nute, G. R., Hughes, S. I., Wood, J. D., & Richardson, R. I. (2008). Effects of breed and a concentrate or grass silage diet on beef quality in cattle of 3 ages. II: Meat stability and flavour. *Meat Science, 78* (3), 270-278. https://doi.org/10.1016/j.meatsci.2007.06.007

Yang, X., Woerner, D. R., Hasty, J. D., McCullough, K. R., Geornaras, I., Sofos, J. N., & Belk, K. E. (2016). An evaluation of the effectiveness of FreshCase technology to extend the storage life of whole muscle beef and ground beef. *Journal of Animal Science, 94* (11), 4911-4920. http://dx.doi.org/10.2527/jas.2016-0508

Table 1
Recent studies with visual hedonic evaluation of ruminant meat during display

Samples/animal (n)	Display period analysed	Shelf-life results	Hedonic scale	Shelf-life estimation	Evaluators	Design	Reference
Longissimus steaks/ steers (n=8)	Daily (0 to 9 days)	5 to 6 days	Visual acceptability (9 points)	Values below 5	17 consumers	Basic design	(Prado et al., 2015)
Longissimus steaks/young bulls (n=10)	Daily (1 to 10 days)	4 to 5 days	Colour acceptability (9 points) Willingness to buy (yes/no)	Values below 5 Survival model (50% rejected samples)	37 consumers	Basic design	(Eiras et al., 2017)
Longissimus chops/lamb (5 different chops per day from 12 animals/treatment)	0, 4, 7, 11 and 14 days	11 to >14 days	General appearance (5 points)	Values above 3	6 trained panellists	Basic design	(Guerra-Rivas et al., 2016)
Quadriceps femoris steaks/beef (not mentioned)	0, 4, 8, 12, 16 and 20 days	No determined	Colour acceptability (9 points)	No determined	10 trained panellists	Basic design	(Langroodi et al., 2018)
Longissimus steaks (real samples and photos)/ young bull (n=8 to 16)	Daily (1 to 11 days)	6.3 to 8.1 days	Colour acceptability (9 points)	Regression analysis (when reach 5)	17 semi-trained	Basic and reversed design	(Passetti et al., 2017)
Longissimus steaks Goat (n=5 to 10)	Daily (0 to 14 days)	7.1 to 10.5 days	Purchase intention (yes/no)	Survival model (50% rejected samples)	18 consumers	Basic design	(Possamai et al., 2018)
Leg chops/sucking kids (n=4 selected samples)	1, 3, 6 and 8 days	6 to 8 days	Purchase intention (yes/no)	Survival model (50% rejected samples)	56 consumers	Basic design	(Ripoll et al., 2018)

Table 2
Scores of visual acceptability[1] following the method[2] of image presentation of lamb steaks through display

Days	Random	Sequential	Animal	P-value	SEM
0	6.47Cab	6.97Ba	7.80Aa	0.001	0.049
1	6.81Ba	6.90Ba	7.66Aab	0.001	0.047
2	6.63Bab	6.57Bab	7.37Ab	0.001	0.047
3	6.39Bbc	6.30Bbc	6.94Ac	0.001	0.046
4	6.40Abc	6.051Bcd	6.59Acd	0.001	0.045
5	6.07Ac	5.74Bd	6.24Ad	0.001	0.045
6	5.48Ad	5.22Be	5.57Ae	0.003	0.045
7	5.31d	5.08e	5.26e	0.068	0.045
8	4.73e	4.55f	4.68f	0.214	0.046
9	4.26f	4.19fg	4.20g	0.738	0.046
10	4.11Afg	3.81Bgh	3.78Bh	0.003	0.047
11	3.86Agh	3.59Bhi	3.38Bi	0.001	0.047
12	3.59Ahi	3.23Bij	2.96Aj	0.001	0.047
13	3.41Ai	2.90Bjk	2.60Cj	0.001	0.047
14	2.92Aj	2.55Bkl	2.22Ck	0.001	0.045
P-value	0.001	0.001	0.001		
SEM	0.029	0.033	0.037		

[1] Scale from 1: "I dislike it extremely", to 9: "I like it extremely"
[2] Random: days of display and animals in random order (reversed design); Sequential: days of display in sequential and animals in random order (basic design); and Animal: days of display and animal in sequential order
[a-l] Different letters in the same column mean significant differences (Tukey 0.05).
[A-C] Different letters in the same line mean significant differences (Tukey 0.05)..

Table 3
Standard deviation of visual acceptability[1] following the method[2] of image presentation of lamb steaks through display

days	Random	Sequential	Animal	*P* Value	SEM
0	1.16Afg	0.95Be	0.68Cg	0.001	0.019
1	1.12Afg	0.95Be	0.74Cg	0.001	0.019
2	1.15Afg	0.95Be	0.78Cfg	0.001	0.017
3	1.25Adef	1.01Bbcde	0.87Cef	0.001	0.017
4	1.27Ade	1.00Bcde	0.90Cef	0.001	0.016
5	1.37Acd	1.00Bcde	0.96Cde	0.001	0.018
6	1.45Abc	1.09Babcd	1.02Bbcde	0.001	0.018
7	1.58Aab	1.14Babc	1.08Babcd	0.001	0.019
8	1.58Aab	1.14Babc	1.13Bab	0.001	0.019
9	1.58Aab	1.18Ba	1.16Ba	0.001	0.018
10	1.62Aa	1.18Ba	1.19Ba	0.001	0.018
11	1.59Aab	1.15Bab	1.18Ba	0.001	0.018
12	1.61Aa	1.11Babcd	1.15Ba	0.001	0.019
13	1.59Aab	1.10Babcd	1.11Babc	0.001	0.019
14	1.32Acd	0.98Bde	0.99Bcde	0.001	0.018
P-value	0.001	0.001	0.001		
SEM	0.008	0.008	0.008		

[1] Scale from 1:" I dislike it extremely", to 9: "I like it extremely"
[2] Random: days of display and animals in random order (reversed design); Sequential: days of display in sequential and animals in random order (basic design); and Animal: days of display and animal in sequential order
[a-g] Different letters in the same column mean significant differences (Tukey 0.05).
[A-C] Different letters in the same line mean significant differences (Tukey 0.05).

Table 4:

Shelf-life calculated by regression (y = number of days and x = score 5.0) of visual acceptability[1] scores following the method[2] of image presentation of lamb steaks, depending on the day points evaluated

Days evaluated	Methodologies	Equation	R^2	p-value	Shelf-life
All days 0; 1 ;2; 3; 4; 5; 6; 7; 8; 9;10; 11; 12; 13 and 14 (*N*=15)	Random	$y = -0.008x^2 -0.180x + 6.900$	0.576	0.001	7.83 days
	Sequential	$y = -0.003x^2 -0.284x +7.134$	0.568	0.001	7.00 days
	Animal	$y = -0.003x^2 -0.379x +8.030$	0.712	0.001	7.54 days
Excluding day 0 1 ;2; 3; 4; 5; 6; 7; 8; 9;10; 11; 12; 13 and 14 (*N*=14)	Random	$y = -0.002x^2 -0.282x + 7.269$	0.580	0.001	7.63 days
	Sequential	$y = -0.001x^2 -0.325x +7.280$	0.545	0.001	6.87 days
	Animal	$y = -0.436x +8.235$	0.694	0.001	7.41 days
Every 2 days 0; 2; 4; 6; 8; 10; 12 and 14 (*N*=8)	Random	$y = -0.010x^2 -0.146x + 6.752$	0.594	0.001	7.81 days
	Sequential	$y = -0.003x^2 -0.285x +7.077$	0.593	0.001	6.80 days
	Animal	$y = -0.003x^2 -0.379x +7.976$	0.731	0.001	7.41 days
Every 3 days 0; 3; 6; 9 and 12 (*N*=5)	Random	$y = -0.012x^2 -0.122x + 6.607$	0.521	0.001	7.55 days
	Sequential	$y = -0.004x^2 -0.267x +7.023$	0.541	0.001	6.87 days
	Animal	$y = -0.006x^2 -0.341x +7.870$	0.695	0.001	7.39 days
Every 4 days 0; 4; 8 and 12 (*N*=4)	Random	$y = -0.017x^2 -0.058x + 6.580$	0.543	0.001	7.40 days
	Sequential	$y = -0.006x^2 -0.241x +7.004$	0.570	0.001	7.07 days
	Animal	$y = -0.008x^2 -0.313x +7.845$	0.714	0.001	7.61 days
Every 5 days 0; 5 and 10 (*N*=3)	Random	$y = -0.031x^2 +0.074x + 6.474$	0.476	0.001	8.19 days
	Sequential	$y = -0.014x^2 -0.175x +6.966$	0.524	0.001	7.14 days
	Animal	$y = -0.018x^2 -0.218x +7.800$	0.670	0.001	7.81 days
Every 6 days 0; 6 and 12 (*N*=3)	Random	$y = -0.012x^2 -0.091x + 6.474$	0.558	0.001	7.92 days
	Sequential	$y = -0.003x^2 -0.270x +6.966$	0.606	0.001	6.77 days
	Animal	$y = -0.005x^2 -0.339x +7.800$	0.742	0.001	7.44 days

[1] Scale from 1: "I dislike it extremely", to 9: "I like it extremely"
[2] Random: days of display and animals in random order (reversed design); Sequential: days of display in sequential and animals in random order (basic design); and Animal: days of display and animal in sequential order

Table 5
Shelf-life calculated by regression (y = number of days and x = score 5.0) of visual acceptability[1] scores following the method[2] of image presentation of lamb steaks, evaluated daily until Day 8 to Day 14 of display

Days evaluated	Methodologies	equation	R^2	p-value	Shelf-life
Until day 8 (*N*=9)	Random	$y = -0.041x^2 +0.095x +6.589$	0.271	0.001	7.49 days
	Sequential	$y = -0.015x^2 -0.187x +7.022$	0.288	0.001	6.94 days
	Animal	$y = -0.021x^2 -0.228x +7.855$	0.454	0.001	7.43 days
Until day 9 (*N*=10)	Random	$y = -0.038x^2 +0.070x +6.609$	0.367	0.001	7.49 days
	Sequential	$y = -0.013x^2 -0.197x +7.031$	0.349	0.001	7.03 days
	Animal	$y = -0.018x^2 -0.247x +7.871$	0.523	0.001	7.34 days
Until day 10 (*N*=11)	Random	$y = -0.028x^2 -0.002x +6.680$	0.420	0.001	7.71 days
	Sequential	$y = -0.012x^2 -0.210x +7.043$	0.406	0.001	6.96 days
	Animal	$y = -0.015x^2 -0.276x +7.899$	0.575	0.001	7.47 days
Until day 11 (*N*=12)	Random	$y = -0.020x^2 -0.066x +6.750$	0.464	0.001	7.84 days
	Sequential	$y = -0.008x^2 -0.242x +7.077$	0.448	0.001	6.98 days
	Animal	$y = -0.011x^2 -0.308x +7.934$	0.618	0.001	7.51 days
Until day 12 (*N*=13)	Random	$y = -0.014x^2 -0.116x +6.812$	0.504	0.001	7.78 days
	Sequential	$y = -0.006x^2 -0.260x +7.100$	0.490	0.001	6.96 days
	Animal	$y = -0.008x^2 -0.333x +7.965$	0.654	0.001	7.54 days
Until day 13 (*N*=14)	Random	$y = -0.009x^2 -0.165x +6.877$	0.534	0.001	7.94 days
	Sequential	$y = -0.004x^2 -0.275x +7.120$	0.530	0.001	6.99 days
	Animal	$y = -0.005x^2 -0.358x +8.000$	0.685	0.001	7.58 days
All days (*N*=15)	Random	$y = -0.008x^2 -0.180x + 6.900$	0.576	0.001	7.83 days
	Sequential	$y = -0.003x^2 -0.284x +7.134$	0.568	0.001	7.00 days
	Animal	$y = -0.003x^2 -0.379x +8.030$	0.712	0.001	7.54 days

[1] Scale from 1: “I dislike it extremely”, to 9: “I like it extremely”

[2] Random: days of display and animals in random order (reversed design); Sequential: days of display in sequential and animals in random order (basic design); and Animal: days of display and animal in sequential order

Table 6
Number of consumers needed for visual acceptability tests of lamb steaks according the method of image presentation, considering an α of 0.05

Method	RMSL	d	β	
			0.2	0.1
Random	0.226	0.2	21	28
		0.1	81	109
Sequential	0.208	0.2	18	23
		0.1	69	92
Animal	0.186	0.2	14	19
		0.1	55	74
Global	0.207	0.2	17	23
		0.1	68	91

α: probability of Type I error.
RMSL: root mean square error divided by scale length.
d: difference in means that is sought in the experiment (scale 0–1).
β: probability of Type II error.

Highlights

- Method of presentation and display day information affected acceptability scores.
- Not bloomed initial point (day zero) affected shelf-life calculation by regression.
- It was required 4 points after and before rejection day for shelf-life determination.
- Display shelf life of meat studies used fewer consumers than suggested.
- Digital images simplify the usage of reversed designs and more consumers.