Shelf Life of Food: An Overview

Catherine Cantley
University of Idaho (UI) Extension/
TechHelp Assistant Professor,
Department of Animal, Veterinary, and
Food Sciences

Janna Verburg-Hamlett
UI Extension/TechHelp Assistant
Professor and Food Safety Specialist,
Department of Animal, Veterinary, and
Food Sciences

Contents

University *of* **Idaho**
Extension

Introduction

A SIGNIFICANT CONTRIBUTOR to both fixed and variable costs in the food-manufacturing supply chain is shelf life. Shelf-life considerations directly impact

- the cost of primary packaging materials (as it relates to moisture, oxygen, and light barriers)
- the cost of production (production volumes are deliberately smaller with shorter shelf-life products)
- the cost of storage (including the size and utility requirements of warehouses)
- the markets that a company can supply (short shelf life often precludes companies from selling to remote and/or international locations or on e-commerce platforms)

Optimizing shelf life thus has significant value, but it can be time consuming and expensive.

Terminology

Shelf life is often indicated to the consumer (retail) or operator (food service) with a **Use By, Best Before, Sell By, Best By, or Expires On** date on the label (Figure 1). However, consumers don't always understand the labeling. In 2007, Kosa et al. found that only about 18% of consumers polled could correctly define "use by date," with nearly two-thirds relying on their senses

LABELING

Sell By/Before helps stores with inventory management, telling them how long to display the product for sale. It is not a safety date.

Best By/Before tells the consumer/operator how long a product will retain its best flavor or quality. It is not a purchase or safety date.

Use By/Before or Expires On is the last date recommended for the product to retain peak quality. It is not a safety date unless it's used on infant formula.

Freeze By/Before is a recommendation to the consumer to help maintain the product's optimal quality. It is not a safety date.

Figure 1. Labeling terminology commonly used by food manufacturers to communicate shelf-life expectations to the end user (Kosa et al. 2007; United States Department of Agriculture 2019).

to decide whether a food is safe to eat or not. Consumer uncertainty about the meaning of the dates likely contributes to about 20% of food waste in American homes (United States Food and Drug Administration [FDA] 2019b). Indeed, consumers discard food from fear, rather than proven scientific principles (Kosa et al. 2007; United States Food and Drug Administration 2019b; United States Department of Agriculture 2019). For this reason, the FDA supports the food industry's effort to standardize the use of the term **Best If Used By** to indicate acceptable sensory quality (not safety) (United States Food and Drug Administration 2019b).

However, manufacturers use many labels or strategies to indicate a product's shelf life when supplying a food service.

- For a shelf-stable, ready-to-eat (RTE) product that requires no further preparation, it may be sufficient to define that finished product's shelf life only in their specifications. Other products might require more specificity, like detailed storage conditions, preparation instructions, and holding conditions, in order to maintain appropriate food safety and optimal food quality.

- If a manufacturer performs bulk preparation of ingredients or a staged production, the product may require temporary packing in large bins or totes. In this case, it is necessary to define an **ingredient shelf life**—how long a tote may be staged under particular conditions (temperature, humidity, packaging format, etc.) before food safety or quality diminishes. Once all of the ingredients are processed and packaged in a consumer-ready form, the product has a **finished product shelf life** or simply **shelf life** (Figure 2).

- A manufacturer selling internationally or to certain distributors may find it prudent to define a **ship-by date,** because many customers or importers require that a particular percentage of the shelf life remain upon receipt of the goods, which they will factor in with the distribution time.

- If the product in question is frozen, a manufacturer may specify a **thawed shelf life** (the length of time a product will remain safe and of high quality) while "slacked out" (allowing a food's temperature to gradually increase under controlled conditions) and stored in refrigeration.

- Finally, **hold time** is a very important consideration in food service (analogous to **sell-by/before** in retail display cases). This refers to the amount of time a hot product can be held hot (typically under heat lamps, in steam tables, or in holding cabinets) or a cold product can be displayed on a chilled prep rail before its quality deteriorates or becomes unsafe.

SPECIFICATIONS

Ingredient Shelf Life is the length of time that an ingredient or bulk tote remains acceptable in ideal (or realistic) storage conditions.

Finished Product Shelf Life is the length of time that a finished product remains acceptable under specified, ideal storage conditions.

Ship-By Date corresponds to the customer-specific percentage of the remaining shelf-life requirement.

Thawed Shelf Life is the length of time a frozen, finished product remains acceptable after its removal from frozen storage conditions to refrigerated storage.

Hold Time is the length of time a product may be held after being prepared in its final format and readied for service.

Figure 2. Specification terminology commonly used by food manufacturers to communicate shelf-life expectations to distributors and food service.

Regulations

Except for infant formula, federal regulations do not require product dating. Indeed, they prohibit manufacturers from presenting misleading information on packaging, but shelf-life dates are not mandatory (United States Food and Drug Administration 2019b; USDA 2019). Regarding infant formula, regulation code 21 CFR PART 107 (finalized June 9, 2014) stipulates that manufacturers test these products for nutrient content in the final product

stage before they enter the market and when they exit (at shelf-life expiration) (United States Food and Drug Administration 2014; FDA 2019a). Although the United States Department of Agriculture (USDA) does not require quality or food-safety data labels for products under its jurisdiction, it mandates the display of a pack date for poultry products and thermally processed, commercially sterile products for traceability in the event of an outbreak of foodborne illness (USDA 2019).

The Approach

The American Society of Testing and Materials (ASTM International 2019) defines the goal of shelf-life determination as estimating the time at "which a consumer product is no longer usable, unfit for consumption, or no longer has the intended sensory characteristics." Shelf-life testing (Figure 3) often includes microbiological stability testing to ensure food safety, sensory and consumer testing to ensure consumer acceptability, and analytical testing to

ensure standard of identity (e.g., infant-formula nutrition standards). A **Best-If-Used-By** date will never be longer than the proven food-safety limits. However, a labeled shelf life may be significantly shorter than food-safety limits, if quality deteriorates faster than safety.

Many shelf-life dates are set with poor scientific rigor or poorly defined standards. This is because many manufacturers either don't know how to design a shelf-life program or lack the budget to outsource it. Well-designed and -executed scientific studies more reliably lead to a thorough analysis of failure mode. Depending on the product and process, this may be a single mode of failure or it may involve several. After determining specifically how a product changes with time, product-development or quality specialists will initiate a root-cause analysis and finally begin manipulating the product formulation, processing, or packaging to extend the shelf life to a desired time frame.
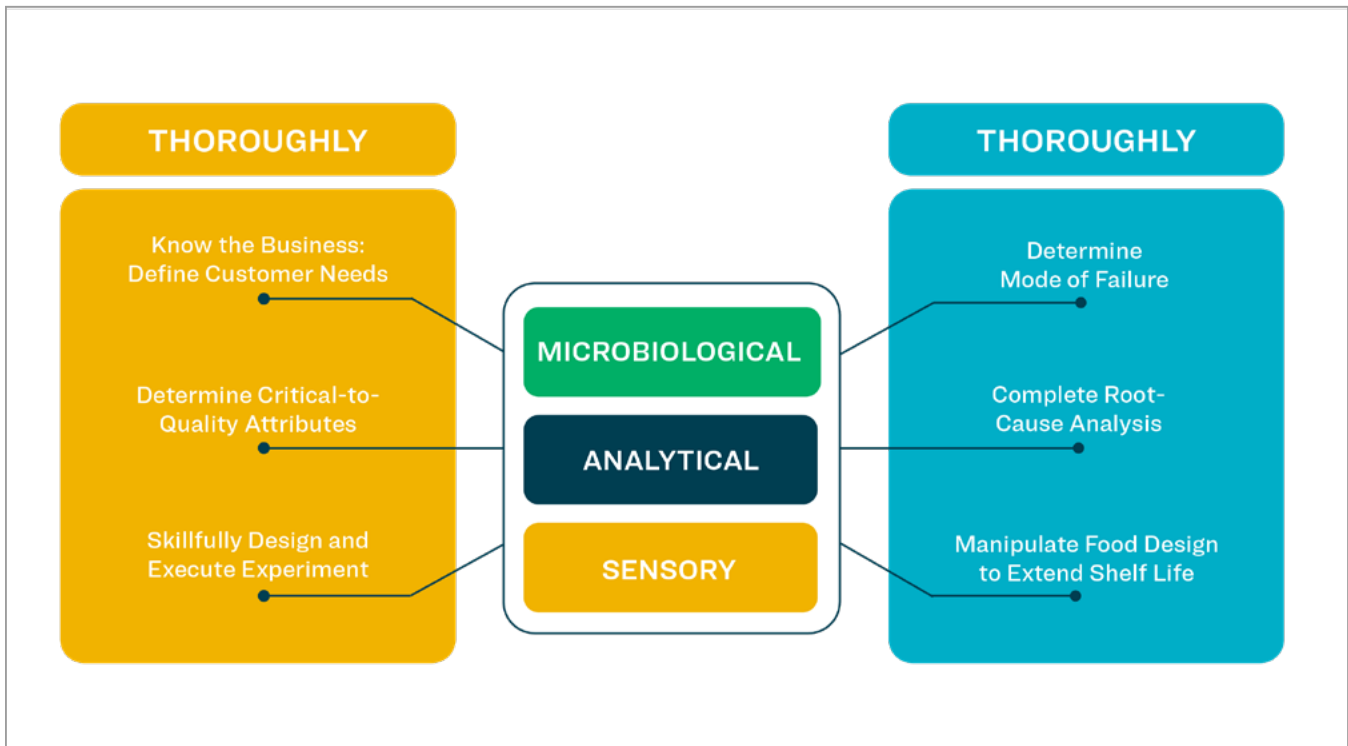

Figure 3. Approach to conducting shelf-life studies using microbiological, analytical, and sensory sciences. Thoroughly prepare by investigating customer needs, determining critical-to-quality attributes, and designing experiments that incorporate the appropriate level of variation. After experimentation, thoroughly investigate the cause of expiration by conducting root-cause analysis. Manipulate food design and packaging to extend shelf life as needed.

Experimental Design

Defining shelf-life success criteria is critical to do up front, because it informs the experiment's design. Common success criteria will be discussed in subsequent sections, but when the product eventually fails to meet success criteria, the *true shelf life* is the last sample time in which results were favorable. For example, if an analyst receives favorable results at 1, 6, and 7 weeks and unfavorable results at 8 weeks, then the true shelf life of the product is 7 weeks.

When planning for how long to conduct a shelf-life experiment, the analyst needs to start with business expectation. This may come from customers, consumers, similar products on the market, distribution needs, risk aversion, and/or business need. Many companies decrease the *labeled shelf life* to 20%–30% less than the *true shelf life*. They take this approach to account for production and distribution variability that may impact quality or safety. Thus, the experimental testing time may be 20%–30% longer than the shelf life indicated on the product package.

Depending on how well the company already understands the mode of failure, the number of sampling points in between time-zero and end-of-life will vary. To save cost, it is often not prudent to sample heavily at the beginning of a product's life. Rather, it is more pragmatic to back-load the sampling times. An example sampling schedule is shown below:

%	0	10	20	30	40	50	60	70	80	90	100	110	120	130
	X				X		X	X	X	X				X

Figure 4. An example of sampling times for shelf-life experiments, according to the percentage of predicted shelf life.

A good experimental design aims to minimize and/or account for variability. When conducting a shelf-life study, a number of factors are accounted for:

- Ingredient Variability
  - » Seasonality of ingredients (including changes to nutritional characteristics such as fats, sugars, and solids)
  - » Variety (if using raw agricultural products)
  - » Supplier variability

- Processing Variability
  - » Microbial load
  - » pH, acid content
  - » Degree of sanitation
  - » Temperatures
  - » Dwell times
  - » Pressures
  - » Piece-to-piece variability resulting from batch processing or side-to-side or top-to-bottom differences in bedded product

- Storage Conditions
  - » Temperatures
  - » Humidity
  - » Air flow
  - » Packaging
  - » Light exposure

- When testing color, texture, and flavor, it is often important to control:
  - » Serving temperature
  - » Testing area and lighting
  - » Competing odors
  - » Time of day
  - » Serve and presentation order (being aware of halo effects and positional bias)
  - » Palate cleansers

Sample size at each time point depends on the confidence level desired (90%–95% is common in the food industry), the amount of variability present in the product (standard deviation), and the size of difference a scientist is attempting to identify (Figure 5). After determining product variability and defining the desired detectable attribute-intensity difference, an experimenter calculates the number of samples needed to identify statistically relevant shifts in product quality. In this case, an increase in statistical confidence from 90% to 95% requires seven more samples for each time point in the study. Thus, a researcher will want to carefully consider the trade-offs in resources and accuracy.

CONFIDENCE LEVEL	90%	95%
DIFFERENCE	1.00	
STANDARD DEVIATION	2.40	
Alpha divided by 2	0.05	0.025
	1.645	1.960
SAMPLE SIZE REQUIRED	16	23

$\leftarrow n = \left(\dfrac{Z\sigma}{E}\right)^2$

Figure 5. Example of sample-size calculations at 90%- and 95%-confidence intervals.

Microbiological Testing

Products in which shelf life is determined microbiologically include many dairy products (pasteurized milk, cheese, dairy-based desserts), meat products (fresh meat, cooked meats such as sausage), fresh egg products (eggs, pasta) (Hough 2010), and refrigerated ready-to-eat items such as beverages, dressings, chilled deli salads, dips, and spreads. It is also crucial in canned meats, fruits, and vegetables, where anaerobic pathogens may grow and/or produce spores.

Products are analyzed for applicable quality indicators. This may include aerobic plate count, gram-negative bacteria, lactic acid bacteria, coliforms, pathogenic bacteria, and yeast and mold counts. The tests done are highly dependent on the specific product and the organisms common to it. For example, the dominating bacteria of minimally processed vegetables are *Pseudomonadaceae* spp., *Enterobacteriaceae* spp., environmental pathogens, and some species of lactic acid bacteria (Ragaert et al. 2007). The microorganisms frequently implicated in cheese expiration are *Clostridium* spp., *Staphylococcus aureus*, *Listeria monocytogenes*, and yeast and mold (Jalilzadeh et al. 2015). The microbiological testing time varies between two and five days. Companies that operate a lean operation might find this microbiological analysis procedure frustrating and costly.

One of the major factors that indicates how long a product will last is its initial microbial load (at the beginning of a product's shelf life). Emerging technology points to a number of promising tests that *more rapidly* quantify these loads, including adenosine triphosphate bioluminescence, deoxyribonucleic acid marker testing, polymerase chain reaction techniques, immunoassays (such as enzyme-linked immunosorbent assay) and separation techniques, electrochemical methods, and spectroscopic techniques, including UV-vis NIR spectroscopy and FTIR spectroscopy (Ziyaina et al. 2020).

In addition to passive shelf-life studies, scientists may be interested in launching more active approaches, including challenge studies or growth studies. These studies consist of (1) inoculating a food product with pertinent pathogens or indicator organisms, (2) storing it under controlled conditions (that may intensify extrinsic factors), and (3) subsequent modeling of growth. They can be used to test different processing or storage conditions.

Analytical Testing

There are two areas in which analytical testing may be the key limiting factor in a product's shelf life: nutritional testing and identifying chemical reactions and by-products.

In some foods, nutritional testing is critical, as it may be inextricably linked to its standard of identity. One example of this, as required by the FDA, is human milk replacement formulas (infant formulas) (Hough 2010). This testing may include degradation of vitamins, nutraceutical active ingredients, or other unstable compounds.

Food-safety experts understand the reaction kinetics for microorganisms and their relationship to water activity, pH, and sodium chloride. Many scientists study sorption isotherms (the relationship between water activity and moisture content) to understand the "sweet spot" for lipid oxidation, nonenzymatic browning, mold and yeast growth, and bacterial growth.

The food industry focuses on developing analytical methods because they are much less expensive and time-consuming than microbiological or sensory tests. Reaction by-products such as peroxides in lipolysis can be very helpful to understand rancidity. If the product is a fresh agricultural product, shelf life may be correlated to moisture and weight loss, which can be measured analytically.

Some correlations are being developed in which complex reactions or microbial spoilage are quantified using benchtop analytical measurements. For example, color has been successfully correlated with spoilage organisms over shelf life in milk products (Ziyaina et al. 2019), optical measurements of visual defects have been correlated with enzymatic browning in minimally processed vegetables (Ragaert et al. 2007), off-aromas have been detected using high-performance liquid chromatography in cheese (Marilley and Casey 2004), and electronic noses are making news in meat-quality assessment (Wojnowski et al. 2017). In the last ten years any number of these technologies have been used across most food categories. Where available and well developed, industry has adopted these methods. This is especially true for colorimeters and spectroscopy. However, in many cases the analytical:sensory correlations are not strong or simply do not exist. Especially in the area of off-odors and flavors, the chemical compounds, often in the parts-per-trillion level, are easier for a human to detect and articulate than an instrument.

Sensory Testing

The number of products in which shelf life is dependent on sensory properties far exceeds those set by microbiological or nutritional/analytical properties (Hough 2010). Sensory is "a scientific discipline used to evoke, measure, analyze, and interpret reactions to those characteristics of foods and materials as they are perceived by the senses of sight, smell, touch, taste, and hearing" (Sidel and Stone 1993). At about eighty years old, sensory is a fairly young science and not as ingrained into manufacturing as quality standards such as grade A. As a result, many researchers are not adept at conducting this type of research. It is common to see somewhat arbitrary scales used to measure sensory quality standards. Using vague terminology such as **limit of marketability**, **acceptable**, and/or **inedible** will not aid companies in making definitive business decisions. Ambiguous standards lead to ambiguous solutions.

There are two main tools in a sensory scientist's toolbox: descriptive tests and affective tests. Descriptive tests are conducted with a highly trained panel. They are used to precisely quantify the intensity of appearance, aroma, flavor, texture, and aftertaste attributes over time. The type of data that results from a well-trained panel is objective, precise, and unbiased. As a result, sample sizes for this type of measurement can be small (n = 10, reps = 2–3). This type of data is used in the aforementioned correlations with instrumental measures (e.g., rancid aroma/flavor intensity correlates with peroxide values). Affective tests, however, are conducted with consumer panelists (or operators if selling to food service). They are typically much larger in scale (n = 50–200, reps = 1–3, depending on type of product and type of test). Consumer testing indicates whether or not the average consumer/operator detects differences and, if so, whether they respond favorably or negatively to those changes.

Sensory shelf-life program designs vary greatly, correlating with company culture, including degree of risk aversion and the degree of their reliance on data to make decisions. If the company is very conservative with shelf life, they may require product at end-of-life (finished product, hold time, or thawed shelf life) to be indistinguishable from a fresh control. Discrimination tests are typically employed here. If the company wants to understand how a product changes over time, it will often use a trained descriptive panel (either internal employees or an external research firm) to quantify those attribute intensities. If a company is very interested in how their customers or consumers perceive those changes and how it changes their purchase behavior, it will likely use consumer testing (either internally or in a central location test). Resources such as number of internal employees, internal testing facilities, and/or budgetary constraints will greatly influence the testing conducted.

In many experimental designs, a control is selected to validate the data set. This takes some level of expertise. A sensory technologist evaluates the trade-offs of sample size requirements versus risk: Choose a new control at each sample time? What are the confounding factors? Collect enough product for the entire study at once? What are the facility's storage limitations? Store a control sample under different conditions? What factors in the data might that confound? As in any good study, it is best to keep as many factors constant as possible.

Table 1. Sensory and consumer testing matrix, factoring in risk aversion and resources. CLT = central location test.

Craves Data	Resources	Sensory and Consumer Tests
Low	Low	Difference from Control
Low	High	Tetrad, Duo-Trio, Forced Choice, Triangle, Affective/ Liking (CLT)
Low	Low	Affective/Liking (Internal Taste Test)
Medium	Medium	Trained Panel, Affective/Liking (Internal Taste Test)
High	High	Trained Panel Affective/Liking (CLT)

Accelerated Shelf-Life Studies

Accelerated shelf-life studies are appealing, due to current pressures to decrease new product development times and to address challenges associated with testing special-purpose foods, such as military combat rations, which may exceed ten years. Yet these studies are notoriously complicated and faulty. This is because the reaction kinetics for complex rheological and textural properties are not easily modeled, other biological changes happen during food storage, and many of the reactions interact with each other (Hough 2010). To be sure, the statistical conundrum is not for the faint of heart. In general, the simpler the formulation, the more likely an accelerated shelf-life study's validity.

If a company is bold enough to attempt one, its leadership must first understand the mode(s) of failure. This is critical because an accelerated shelf-life study involves speeding up the reaction kinetics for those particular modes of failure. Second, it must make a number of assumptions, including a big one—that a single acceleration factor affects all mode-of-failure-reactions identically (Hough 2010). In addition, the acceleration factor may actually cause changes to other ingredients that are different from normal storage conditions. In fact, attributes intended to be held constant during the experiment may change before the attribute(s) the scientist desires change.

Conclusion

Food safety is paramount, but quality-based shelf-life testing is not required by regulatory bodies (except for infant formula). While not mandated, this testing is critical to ensure that companies maintain high brand loyalty and customer satisfaction, while maximizing their margins. A good shelf-life study accounts for a variety of factors, including product, process, storage, and testing conditions. The type of study done is often determined by what the mode of failure is, which is limited by microbial growth, standard of identity, or sensory quality.

References

ASTM International. 2019. *ASTM E2454-19a, Standard Guide for Sensory Evaluation Methods to Determine the Sensory Shelf Life of Consumer Products*. ASTM International: West Conshoshoken, PA. doi: 10.1520/ E2454-19A.

Hough, G. 2010. *Sensory Shelf Life Estimation of Food Products*. CRC: Boca Raton, FL.

Jalilzadeh, A., Y. Tunçtürk, and J. Hesari. 2015. "Extension Shelf Life of Cheese: A Review." *International Journal of Dairy Science* 10(2): 44–60. doi: 10.3923/ijds.2015.44.60.

Kosa, K.M., S.C. Cates, S. Karns, S.L. Godwin, and D. Chambers. 2007. "Consumer Knowledge and Use of Open Dates: Results of a Web-Based Survey." *Journal of Food Protection* 70(5): 1213–19. doi: 10.4315/0362-028X-70.5.1213.

Marilley, L., and M.G. Casey. 2004. "Flavours of Cheese Products: Metabolic Pathways, Analytical Tools and Identification of Producing Strains." *International Journal of Food Microbiology* 90(2): 139–59. doi: 10.1016/S0168-1605(03)00304-0.

Ragaert, P., F. Devlieghere, and J. Debevere. 2007. "Role of Microbiological and Physiological Spoilage Mechanisms during Storage of Minimally Processed Vegetables." *Postharvest Biology and Technology* 44(3): 185–94. doi: 10.1016/j.postharvbio.2007.01.001.

Sidel, J.L., and H. Stone. 1993. "The Role of Sensory Evaluation in the Food Industry." *Food Quality and Preference* 4(1–2): 65–73. doi: 10.1016/0950-3293(93)90314-V.

United States Department of Agriculture. 2019. *Food Safety and Inspection Service*. "Food Product Dating." https://www.fsis.usda.gov/food-safety/safe-food-handling-and-preparation/food-safety-basics/food-product-dating. Accessed 28 April 2020.

United States Food and Drug Administration. 2014. "Infant Formula: Safety Do's and Dont's." https://www.fda.gov/consumers/consumer-updates/infant-formula-safety-dos-and-donts. Accessed 28 April 2020.

United States Food and Drug Administration. 2019a. *CFR - Code of Federal Regulations Title 21*. "Part 107. Subpart B. Section 107.20(c)." https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=107. Accessed 28 April 2020.

United States Food and Drug Administration. 2019b. "Confused by Date Labels on Packaged Foods?"

Wojnowski, W., T. Majchrzak, T. Dymerski, J. Gębicki, and J. Namieśnik. 2017. "Electronic Noses: Powerful Tools in Meat Quality Assessment." *Meat Science* 131(September): 119–31. doi: 10.1016/j.meatsci.2017.04.240.

Ziyaina, M., B. Rasco, T. Co, G. Ünlü, and S.S. Sablani. 2019. "Colormetric Detection of Volatile Organic Compounds for Shelf-Life Monitoring of Milk." *Food Control* 100(June): 220–26. doi: 10.1016/j.foodcont.2019.01.018.

Ziyaina, M., B. Rasco, and S.S. Sablani. 2020. "Rapid Methods of Microbial Detection in Dairy Products." *Food Control* 110: 107008. doi: 10.1016/j.foodcont.2019.107008

UI PEER REVIEWED