The Cold-Chain Evidence Gap

When a ready-to-eat food passes its factory tests but fails the shelf-life burden of proof: Listeria, distribution temperatures and the 2026 EU inspection threshold

José M. López, BSc, MD, PhD, LLD (C)

Food Safety, Industrial Microbiology & Regulatory Affairs

Abstract

Since 1 July 2026, amended European microbiological criteria have changed the legal consequences of an incomplete Listeria monocytogenes shelf-life demonstration for ready-to-eat foods capable of supporting its growth. The industrial difficulty is not simply selecting between absence in 25 g and 100 colony-forming units per gram. It is proving that a particular formulation, process, packaging system and realistically foreseeable distribution profile will remain compliant throughout its declared life. This article examines the often-neglected interface between challenge testing, cold-chain variability, environmental reservoirs, sampling uncertainty, supplier changes and the evidentiary discretion of competent authorities. It separates binding law from technical guidance, compares the EU and Spain with Great Britain, Northern Ireland and the United States, and proposes an auditable approach to managing product families and reformulation. A research agenda explores whether traceable temperature histories and probabilistic growth models can improve decisions without disguising uncertainty. The commercial conclusion is that shelf life is not merely a date printed on a pack; it is a defensible, version-controlled safety argument.

Keywords: *Listeria monocytogenes; ready-to-eat foods; shelf-life validation; cold chain; challenge tests; official controls; microbiological criteria; regulatory evidence*.

1. The regulatory event that turns a laboratory result into a business decision

Commission Regulation (EU) 2024/2895, adopted on 20 November 2024 and applicable from 1 July 2026, amended the Listeria monocytogenes food-safety criterion for category 1.2 ready-to-eat (RTE) foods able to support growth, excluding foods intended for infants and special medical purposes. The amendment addresses a gap in the former regime: where a producer could not demonstrate that L. monocytogenes would remain at or below 100 CFU/g throughout shelf life, the absence-in-25-g criterion had applied before the food left the immediate control of the producing operator, but the regulation did not express the same criterion for such foods after that point. The amendment extends the absence criterion to products placed on the market during shelf life where the required demonstration has not been made. [1]

This is not a universal zero-tolerance rule for every food, nor an automatic prohibition of all detectable Listeria in every RTE product. The applicable criterion depends on the legally defined food category and whether the producer has made a satisfactory demonstration. Nor does a laboratory result below 100 CFU/g on the day of manufacture itself establish compliance at the end of shelf life. The amended text expressly makes the demonstration a condition of the 100-CFU/g pathway. [1,2]

The practical tension lies in the words “to the satisfaction of the competent authority.” They require a technically credible evidence package but do not establish a single experimental recipe that mechanically guarantees acceptance. Producers therefore face two connected tasks: validate microbial behaviour and document why their validation represents the products and temperatures encountered in commerce.

1.1 A narrow amendment with broad consequences

The text changes a microbiological criterion, but its effects propagate through product development, customer specifications, laboratory contracts, cold-chain logistics, shelf-life declarations and the handling of positive environmental findings. A cheese slicing business, chilled prepared-meal manufacturer, fish processor or ready-to-eat protein producer may be exposed differently, depending on post-lethality handling, formulation and consumer instructions. A food intended to be cooked before consumption must not automatically be treated as RTE; however, commercial descriptions such as “heat and eat” do not settle the classification without examining reasonably foreseeable use and the applicable definition.

The producer must distinguish legal compliance, internal release criteria and customer purchasing specifications. A retailer may contractually demand non-detection in a specified sample even where the legal framework permits another route. Such a contractual limit can be commercially decisive but should not be described as the statutory criterion.

2. The industrial fault line: nominal storage temperature versus foreseeable exposure

A shelf-life experiment conducted at a convenient constant refrigerator temperature can be internally reproducible and externally misleading. Food moves through production buffers, loading docks, transport vehicles, distribution centres, retail display units and domestic refrigerators. Each segment contributes a time–temperature history, with potential excursions and uneven temperature distributions. The relevant exposure is the product temperature, not simply the vehicle’s air-temperature set point. The Commission’s updated guidance directs operators to consider reasonably foreseeable conditions of distribution, storage and use; the EURL technical guidance provides a framework for challenge testing and durability studies. [3,4]

This distinction has economic consequences. If a product is validated only at a benign temperature but sold through channels where higher temperatures are reasonably foreseeable, a claimed shelf life may lack an adequate evidentiary basis. Conversely, an unrealistically severe temperature profile can generate a technically conservative but commercially unworkable shelf life. The correct response is to justify a defensible distribution scenario using relevant evidence, not to select whichever temperature produces the desired commercial answer.

2.1 A worked example: why a short excursion matters

Consider an illustrative product with a starting concentration of 1 CFU/g and an assumed net growth rate of 0.20 log10 CFU/g per day under one storage regime. Over ten days, the simplified model predicts a 2-log increase, or approximately 100 CFU/g. If the effective rate were instead 0.30 log10 per day, the same initial concentration would reach approximately 1,000 CFU/g. These numbers are deliberately hypothetical; they are not measured rates for a named food, nor do they include lag phases, changing temperatures, uncertainty in initial contamination or strain heterogeneity. Their purpose is to show that apparently modest differences in growth assumptions can alter the regulatory conclusion by an order of magnitude.

The calculation must not be used as a substitute for a validated model. A realistic non-isothermal prediction requires temperature-dependent growth parameters, food-specific factors, potential lag adaptation and uncertainty bounds. In practice, the initial contamination is rarely known precisely; analytical non-detection

cannot be treated as zero. The prudent interpretation is a distribution of plausible outcomes, not a single deterministic line.

2.2 The missing information is often operational rather than microbiological

Manufacturers may possess extensive laboratory data but limited visibility of downstream temperatures. The retail contract may specify 0–4 °C while telemetry records only the distribution-centre environment; domestic storage may be inferred from assumptions rather than observed data. A sophisticated challenge test cannot correct an unrepresentative temperature profile. This is an evidence-governance problem involving commercial partners, not merely a laboratory problem.

3. Product classification is the first critical decision

Regulation (EC) No 2073/2005 distinguishes RTE foods that support growth from those that do not, and specifies separate categories for foods intended for infants and special medical purposes. Product characteristics such as pH, water activity, shelf life and demonstrated growth potential can affect classification. The producer should establish the classification for the actual finished product, taking account of variability, not simply assign a category based on a marketing label or an ingredient's reputation. [2,3]

A fermented or acidified food may exhibit strong inhibition in one formulation but weaker inhibition following reduced salt, changed acidulant, altered buffering capacity or a new supplier. A low average pH is not necessarily representative of heterogeneous inclusions or local microenvironments. Similarly, a recipe that appears hostile to Listeria may become more permissive after packaging or processing changes. The relevant evidence must track the material product, not a static brand name.

3.1 The danger of treating a product family as one formulation

Product-family approaches can reduce testing costs when technically justified, but equivalence should be demonstrated across worst-case members. Differences in pH, water activity, preservative concentration, fat distribution, particle size, packaging atmosphere, post-process contamination and storage temperature can all influence growth. Grouping products only because they share a commercial category is a weak basis for extrapolation.

Change	Possible microbiological effect	Evidence to revisit
Salt reduction	Higher water activity or altered stress response	Measured aw; growth potential
Protein reformulation	Changed buffering and nutrient availability	Matrix-specific challenge design
New packaging film	Different oxygen or gas transmission	Headspace and storage verification
Supplier change	Different background microbiota or impurities	Incoming specification and comparability
Longer shelf life	More time for growth	Updated end-of-life demonstration

4. What a challenge test establishes today

A challenge test deliberately inoculates a representative food with the target organism under controlled conditions to estimate growth potential or other relevant parameters. It is a research and validation instrument, not a routine finished-product release test. The European Union Reference Laboratory's February 2026 amended technical guidance should be read alongside EN ISO 20976-1 and the Commission's December 2025 guidance. These documents distinguish study objectives and address technical features such as strain selection, inoculation, sampling, storage profiles and interpretation. [3,4]

The critical issue is fitness for purpose. A study designed to estimate growth potential is not automatically interchangeable with one designed to evaluate maximum growth rate. A durability study following naturally occurring contamination provides different evidence from deliberate inoculation. Predictive microbiology is valuable where the model is appropriate to the matrix and range of conditions, but a numerical output cannot substitute for material characterisation or an appropriate validation strategy.

4.1 Where apparently professional studies fail

Common failure modes include testing only the most inhibitory batch; using an inoculation or handling procedure that changes the food; selecting temperatures lower than those reasonably foreseeable; failing to account for formulation variability; extrapolating beyond the validated shelf life; and interpreting a small number of non-detects as proof that growth is impossible. Laboratory accreditation is important but is not itself proof that the chosen protocol answers the manufacturer's legal question.

A manufacturer should specify the decision sought in the laboratory brief: category classification, growth-potential estimate, end-of-life criterion, or evaluation of a particular reformulation. The brief should identify target markets and expected storage conditions. The laboratory should explain which conclusions the resulting study can and cannot support. When the operator and laboratory disagree about an assumption, that disagreement belongs in the record rather than being silently resolved in favour of the preferred shelf-life claim.

4.2 The difference between verification and validation

Validation asks whether the designed control system can achieve the intended safety outcome under specified conditions. Verification checks whether the system is being operated as intended and remains effective. Routine environmental swabs, batch testing, trend analysis and review of deviations contribute to verification; they do not replace a scientifically appropriate shelf-life validation. A factory can have months of satisfactory environmental results while lacking evidence that a low-level contamination event would remain below the applicable limit through the distribution chain.

5. Environmental reservoirs: the factory can be clean on paper and contaminated in practice

Listeria is particularly challenging in chilled RTE operations because environmental persistence and recontamination after a lethality step can dominate risk. Drains, inaccessible equipment interfaces, conveyor undersides, hollow structures, condensation routes, damaged seals and maintenance interventions may connect wet reservoirs with food-contact zones. An environmental monitoring programme must therefore be designed around the plant's actual traffic, water movement and equipment geometry rather than a uniform number of swabs per room.

Negative results from a poorly targeted programme are weak reassurance. Sampling plans should identify hygienic zones, timing during production, post-cleaning checks where appropriate, high-risk niches, and escalation rules. A positive finding requires interpretation in light of organism identity, location, product exposure and recurrence. Repeated positives from related sites may justify deeper investigation, including strain-level characterisation where available; an isolated result does not by itself prove a persistent clone.

5.1 The inspection paradox of better monitoring

A factory that samples aggressively may find more environmental positives than one that rarely looks. The raw count of positives is therefore a poor comparative indicator unless sampling intensity, site selection and analytical methods are considered. Regulators and customers should examine whether findings trigger effective root-cause analysis, sanitation redesign, verification and risk-based product disposition. A low positivity rate obtained by avoiding difficult locations is not evidence of superior control.

Conversely, monitoring cannot become an excuse for uncontrolled production. A positive food-contact-surface result during production may require immediate product risk assessment, intensified sampling, corrective action and, depending on the facts, containment or withdrawal. The response must be proportionate to the hazard and grounded in traceable evidence.

6. Sampling uncertainty and the legal meaning of non-detection

A microbiological result is conditional on sample mass, sampling design, analytical method, organism distribution and detection capability. “Not detected in 25 g” means not detected in the tested analytical unit under the specified method; it does not prove absolute absence from an entire production lot. Likewise, a quantified concentration in a sample cannot be treated as a perfect representation of a heterogeneous batch.

The amended category 1.2 criterion specifies sampling plan n = 5, c = 0 and reference methods EN/ISO 11290-1 for detection and EN/ISO 11290-2 for enumeration, according to the applicable pathway. The sampling plan and interpretation rules must be read from the regulation itself. A manufacturer’s private monitoring plan may be more extensive, but it should not be confused with the statutory criterion. [1,2]

When an official finding conflicts with private testing, the scientifically useful questions include whether the same lot and material were sampled, how samples were taken and transported, which method was used, whether quantification or detection was relevant, and whether the results represent heterogeneous contamination. Legal arguments about procedure should not obscure the need to control a credible hazard; equally, the existence of a hazard does not eliminate procedural safeguards or the need for a reasoned administrative decision.

7. The administration - industry tension: evidence to whose satisfaction?

The competent authority must assess whether the producer’s demonstration is sufficient for the applicable legal criterion. That judgment can involve scientific discretion, especially where the evidence combines challenge tests, predictive models, product-family extrapolation and temperature assumptions. Discretion is not arbitrariness. An adverse finding should identify the relevant legal obligation, the factual deficiency, the material uncertainty and the reason the evidence does not support the operator’s claimed pathway.

The operator, for its part, should present a coherent dossier rather than a collection of certificates. It should connect product identity, critical parameters, process controls, study selection, worst-case conditions, monitoring and change control. If a competent authority rejects a study, the practical priority is to understand the specific technical reason and whether additional data or a revised risk assessment can resolve it. Formal review and appeal routes depend on national procedural law and the nature of the measure.

7.1 Spain: EU rules, decentralised enforcement

Regulation (EC) No 2073/2005 and its amendments are directly applicable in Spain. The European Commission establishes the relevant EU framework; EFSA provides scientific risk assessments within its remit; and competent national and autonomous-community authorities undertake official controls according to their functions. A laboratory report, supplier statement or private certification cannot expand the legal scope of a microbiological criterion. Nor should an inspector’s preference for a particular commercial laboratory be confused with a universally prescribed method unless a legal or technically justified requirement supports it.

7.2 Great Britain, Northern Ireland and the United States

The EU amendment cannot simply be assumed to apply in Great Britain. Operators must check the currently applicable domestic microbiological-criteria framework and relevant Food Standards Agency or Food Standards

Scotland guidance for the actual product and date. Northern Ireland requires separate legal analysis under its applicable arrangements; an EU–UK comparison that treats the United Kingdom as one regulatory territory is unreliable.

In the United States, FDA's preventive-controls framework, current good manufacturing practice requirements and guidance on control of L. monocytogenes in RTE foods must be analysed independently. US enforcement and adulteration concepts are not reducible to the EU's 100-CFU/g demonstration pathway. Exporters should maintain market-specific legal assessments and customer requirements rather than transferring a European compliance conclusion across borders. [5,6]

8. The hidden economics of a weak shelf-life dossier

Shelf-life validation is often treated as a compliance expense, but its true economic value lies in avoiding false confidence and unnecessary restrictions. A weak dossier can trigger shortened declared life, loss of retailer listings, increased testing, blocked stock, costly rework, claims disputes or recalls. An overly conservative dossier can also waste food and reduce the geographical reach of distribution. Neither cost justifies lowering the evidentiary standard; both justify investing in better data.

8.1 A transparent illustrative calculation

Assume a chilled product has annual sales of 2,000,000 units, an average contribution margin of €0.35 per unit, and a 3% distribution write-off rate associated with insufficient remaining shelf life. That implies 60,000 units written off and €21,000 in foregone contribution, excluding manufacturing cost, disposal and logistical effects. If scientifically supported process and logistics changes reduce write-offs to 1.5%, the illustrative contribution improvement is €10,500 per year.

The more important calculation is expected risk-adjusted cost: study expense plus process changes, expected spoilage, expected stock loss and the probability-weighted cost of an adverse safety event. The latter is difficult to estimate and should not be reduced to a deceptively precise number. The model is useful for comparing investment decisions, not monetising an acceptable level of consumer harm.

9. An auditable decision pathway for industrial operators

The first task is a legal and technical inventory of RTE products, classified by growth potential, consumer use, market and shelf-life claim. The second is to identify products for which the 100-CFU/g pathway is claimed and determine whether evidence supports that conclusion under foreseeable distribution conditions. The third is to prioritise the weakest product families: those with changed recipes, heterogeneous matrices, limited temperature data, frequent environmental findings or historical shelf-life extensions.

For each family, establish a version-controlled evidence record linking product specifications, manufacturing sites, process parameters, laboratory studies, model versions, distribution assumptions and approval decisions. Introduce change-control triggers for supplier substitution, salt or acid reduction, new packaging, new thermal process, different equipment or new destination market. The change-control question is not whether the product still tastes the same; it is whether the previous safety argument remains valid.

Decision gate	Minimum question	Documented output
Legal classification	Which RTE category and criterion apply?	Signed classification rationale
Matrix identity	Does the study represent current manufacture?	Batch and formulation comparison
Temperature	Are profiles foreseeable and justified?	Data sources and assumptions

Growth evidence	Does the evidence support end-of-life compliance?	Validated study and limitations
Environment	Are recontamination routes controlled?	Zoned monitoring and corrective-action trends
Change control	What invalidates the prior conclusion?	Revalidation trigger register
Market access	Do destination rules or customer terms differ?	Territory-specific compliance review

10. Research & Innovation Opportunity

10.1 The gap is not a lack of predictive microbiology

Challenge testing, predictive microbiology, temperature logging, environmental mapping and probabilistic risk assessment are established disciplines. The proposed gap is narrower: many industrial evidence packages do not integrate product-specific growth data with the distribution temperatures actually observed and the manufacturing changes that affect the relevance of prior studies. This is a hypothesis about evidence integration and decision quality, not a claim that the underlying technologies are new. Before any proprietary claim, a structured review of published research, standards, patents and commercial systems would be necessary.

10.2 Hypothesis

Unvalidated hypothesis: a version-controlled shelf-life evidence system combining laboratory growth-potential studies, product-level time–temperature distributions, environmental-risk indicators and explicit uncertainty bounds will reduce both false reassurance and unnecessary shelf-life restriction compared with static single-temperature qualification alone. Its plausibility derives from known temperature dependence of microbial growth and established risk-modelling methods. Its incremental predictive performance remains to be demonstrated.

10.3 Proposed study design

A first phase would recruit multiple RTE product families with different pH, water activity, packaging and processing characteristics, including examples with known formulation changes. A second phase would obtain appropriately anonymised temperature records across manufacturing, transport, retail and consumer-relevant segments, with careful separation of air and product temperature. A third phase would conduct challenge tests and, where justified, durability studies following applicable guidance and EN ISO 20976-1. A fourth phase would compare blinded expert decisions based on conventional dossiers against decisions based on integrated evidence, using independent external validation.

Prespecified failure criteria should include poor out-of-sample calibration, missing excursions, excessive model sensitivity to uncertain initial contamination, inability to distinguish product-family members, and evidence packages that fail regulatory review despite improved predictive metrics. A useful result would be a transparent decision protocol and data specification, not an algorithm that purports to authorise products. Any research involving pathogenic challenge testing requires appropriately equipped, competent laboratories and biosafety procedures.

10.4 Industrial relevance and limitations

The potential industrial value lies in more targeted validation expenditure, better logistics negotiations, earlier identification of invalidated studies and defensible shelf-life decisions. It is especially relevant to chilled protein-rich foods, sliced cheeses, fish products and composite RTE meals. It would not replace sanitation, environmental monitoring, legal classification or official controls. Any INTABIOTECH application should be

based on separately validated product-specific evidence, not an implied safety claim for a commercial ingredient.

11. Conclusions

The 2026 amendment does not make food safety reducible to one microbiological number. It exposes a deeper industrial obligation: the producer must be able to explain why the finished food, as actually manufactured and foreseeably distributed, will meet the relevant criterion throughout its declared shelf life. A technically correct test at an irrelevant temperature, or a valid study on an obsolete formulation, may fail that obligation.

The strongest operators will treat shelf-life evidence as a living control system linking formulation, factory hygiene, logistics and legal accountability. Competent authorities should demand credible evidence and explain deficiencies precisely; industry should not confuse favourable routine results with proof of end-of-life compliance. The frontier is not another generic laboratory test. It is the disciplined integration of microbiological evidence with the real industrial conditions under which food is sold.

References and primary sources

[1] Commission Regulation (EU) 2024/2895 of 20 November 2024, amending Regulation (EC) No 2073/2005; applicable 1 July 2026. https://eur-lex.europa.eu/eli/reg/2024/2895/oj

[2] Commission Regulation (EC) No 2073/2005 on microbiological criteria for foodstuffs (as amended). https://eur-lex.europa.eu/eli/reg/2005/2073/oj

[3] European Commission, DG SANTE (December 2025). Guidance document on Listeria monocytogenes monitoring and shelf-life studies for ready-to-eat foods. https://food.ec.europa.eu/document/download/44257174-bf8c-4214-a60d-6790a7ca4109_en

[4] European Union Reference Laboratory for Listeria monocytogenes (26 February 2026). Technical Guidance Document on challenge tests and durability studies for assessing shelf-life of ready-to-eat foods, version 4 amendment 1. DOI: 10.5281/zenodo.18783767. https://zenodo.org/records/18783767

[5] Food Safety Authority of Ireland. Listeria monocytogenes factsheet and 2026 regulatory change. https://www.fsai.ie/business-advice/running-a-food-business/food-safety-and-hygiene/microbiological-hazards/common-food-poisoning-bacteria/listeria-monocytogenes-factsheet

[6] US FDA. Control of Listeria monocytogenes in Ready-To-Eat Foods: Guidance for Industry. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/control-listeria-monocytogenes-ready-eat-foods-guidance-industry

[7] Regulation (EC) No 852/2004 on the hygiene of foodstuffs. https://eur-lex.europa.eu/eli/reg/2004/852/oj

[8] Regulation (EC) No 178/2002 laying down general principles and requirements of food law. https://eur-lex.europa.eu/eli/reg/2002/178/oj

[9] Regulation (EU) 2017/625 on official controls. https://eur-lex.europa.eu/eli/reg/2017/625/oj

[10] ISO 20976-1:2019. Microbiology of the food chain — Requirements and guidelines for conducting challenge tests of food and feed products — Part 1. https://www.iso.org/standard/69673.html

[11] European Commission. Official controls and enforcement. https://food.ec.europa.eu/horizontal-topics/official-controls-and-enforcement_en

[12] Safefood (25 May 2026). Listeria Q&A with FSAI, FSA and industry specialists, discussing shelf-life validation, environmental monitoring and implementation. https://www.youtube.com/watch?v=P6wTD3I_ez8

Evidence classification: [1], [2], [7]–[9] are binding legal instruments as applicable; [3]–[5], [10]–[12] provide guidance, standards and institutional interpretation; industrial failure analyses and decision recommendations are expert synthesis.

© 2026 José M. López. All rights reserved.