

FDA's Response to External Peer Review - Data Review¹ on
FDA's "Draft Report for Peer Review: Risk-Ranking Model for Product Tracing as Required by
Section 204 of FSMA (September 2015)"

August 2020

¹ "Summary Report: External Peer Review of FDA's Draft Risk-Ranking Model for Product Tracing as Required by Section 204 of FSMA – Supporting Data Review." Versar, Inc., March 18, 2016.

I. INTRODUCTION

The FDA Food Safety Modernization Act (FSMA) section 204 (21 U.S. Code § 2223) requires the Agency to designate high-risk foods for which additional recordkeeping requirements are necessary to protect public health. The Agency developed a Risk-Ranking Model for Product Tracing (RRM-PT)² as a data-driven decision support tool to assist the Agency in the process of designating a Food Traceability List as required by FSMA Section 204. The RRM-PT considered the characteristics of both foods and hazards, i.e., food-hazard pairs. The model scores food-hazard pairs according to data and seven criteria: (1) Frequency of Outbreaks and Occurrence of Illnesses, (2) Severity of Illness, (3) Likelihood of Contamination, (4) Growth Potential, with Consideration of Shelf Life, (5) Manufacturing Process Contamination and Industry-wide Intervention, (6) Consumption, and (7) Cost of Illness. These criteria are consistent with the requirements in section 204(d)(2)(A) (21 U.S. Code § 2223(d)(2)(A)).

Two panels of independent external experts provided peer reviews of the draft model and the underpinning data used to generate risk scores with the model, respectively. This report describes the comments from the data review. In the data review, four experts were selected to answer 12 charge questions and to evaluate and provide written comments on a draft report and risk scores, as well as on the underpinning data used to generate risk scores. This peer review focused on the data used to populate the model, including data from the literature and expert elicitation, and the scientific rationale and justification for the scoring of food-hazard pairs in the model.

² Following this peer review, FDA changed the name of the model in 2019 to Risk-Ranking Model for Food Tracing.

Peer reviewers and affiliations:

Jean-Louis Cordier, Ph.D.

CH-1806 Saint-Légier-La Chiésaz
Switzerland

Thomas Montville, Ph.D.

Environmental Laboratory
School of Environmental and Biological Sciences
Rutgers, the State University of New Jersey
New Brunswick, New Jersey 08901

Sara Monteiro Pires, Ph.D.

Research Group for Risk-Benefit
National Food Institute, Technical University of Denmark
1601 Copenhagen V, Denmark

Thomas Ross, Ph.D.

School of Land and Food – Tasmanian Institute of Agriculture
University of Tasmania
University of Tasmania, Australia

II. CHARGE QUESTIONS

FDA has developed a draft risk ranking model to inform the designation of high-risk foods for which additional recordkeeping requirements are appropriate and necessary to assist product tracing, as required under section 204(d)(2) of the FDA Food Safety Modernization Act (FSMA). The draft model was developed through an iterative process that involved, among other things, using the FSMA statutory factors to define criteria and scoring functions of the criteria, and collecting data relevant to the scoring criteria for food-hazard pairs to identify those foods which should be designated as high-risk for future consideration in policy decision.

The focus of this review is on the underlying data used in the model, within the context of the overall risk ranking approach, criteria and results. Note: a separate panel is reviewing the model.

Charge Questions:

1. In order to apply the FSMA factors it is necessary to first take into account both the characteristics of foods and known or reasonably foreseeable hazards, i.e., food-hazard pairs. The food categorization scheme involves a list of 1,286 food-hazard pairs (candidates) and 335 foods (secondary commodities) linked to approximately 50 food categories (primary commodities).
 - a. Is the food classification scheme appropriate and adequate to identify a comprehensive list of foods representative of FDA-regulated products?
 - b. Is the granularity of the food classification appropriate and supportable by available data (particularly the secondary commodities)?
 - c. Are the method and data used to identify the food-hazard pairs adequate?
 - d. If not, please describe another practical alternative to this food classification scheme that might be considered. Please describe what changes, if any, might be considered and provide examples of additional food-hazard pairs (if any) that might be considered in the future.
2. Please evaluate the scoring definitions for the seven criteria:
 - a. Are the definitions appropriately defined for microbial hazard, chemical hazard, and undeclared allergens? If not, please describe changes that might be considered and why.
 - b. For Criterion 3 (C3), the same definition is used for ready-to-eat (RTE) and not-ready-to-eat (NRTE) foods. Should a different criterion weight to C3 be applied to the food risk score for an RTE vs NRTE food-hazard pair? If so, please specify the weighing scheme that might be considered (e.g., C3 weight of 15 for RTE food vs. weight of 10 for NRTE food).

3. The draft model integrates a substantial amount of public health and commodity-specific data about food-hazard pairs, as well as information regarding manufacturing and processing of different foods.
 - a. Are any of the data sources used not appropriate for any of the seven criteria? If so, please explain which data source(s) should not be used and why.
 - b. Are there data sources not yet used but that should be considered? If so, please provide specific examples of data source(s) for each criterion, and explain why the additional data sources might enhance the criteria scoring.
4. Data weighting is used for Criteria 1 and 3, where more weight is given to more recent data and studies more relevant to the U. S. food supply. Are there any of the weighting factors not appropriate? If so, please explain what other weighting factors should be considered.
5. Consumption data from the NHANES What We Eat in America database were used and, for certain commodities (e.g., ice) where no data or limited data were available, expert opinion was used in scoring Criterion 6. Is the timeframe used (2011-2012 or 2009-2010 survey cycle) appropriate? Is it appropriate to use the highest consumption rate from day 1 or day 2 in the scoring? If not, please explain what changes should be considered and what other data sources for consumptions might be considered to improve consumption estimates.
6. Please select primary commodities and hazards within your expertise and review the scoring of a few food-hazard pairs in the criteria files. Are any of the references used not appropriate? Are there any underlying data not appropriately used? Is any score assigned to a criterion not appropriate? If so, please provide suggestions on how the process of data collection and documentation might be refined.
7. The draft model relies on expert elicitation to fill data gaps, in particular for the scoring of Criterion 5 and for the scoring of criteria for chemical hazards and undeclared allergens. Is the expert elicitation process (which involves external panels and FDA subject matter experts) adequate to address data gaps? Is the scoring logic used to combine data and expert opinions and the order of preference in section 4 (e.g., Figure 4.3) appropriate? If not, please explain what changes might be considered and why.
8. Susceptible populations are considered as part of the definitions in Criterion 2. Are there other ways to account for susceptible populations more specifically in the model? For example, would it be appropriate to scale down the % consumption (Criterion 6) for undeclared food allergens-pairs to reflect the proportion of allergic consumers? Please explain.

9. Given the underlying data supporting the scoring, what are the considerations to take into account when identifying high risk vs. not high risk food-hazard pairs or foods?
10. How often should the model be updated, considering the data sources and data currently available and data that might become available in the future?
11. Is the draft report clear in its description of the risk ranking approach, criteria, data and results, and model limitations? If not, please identify aspects that are unclear or could be more transparent.
12. Do you have any additional comments? Please share them in your review.

III. SUMMARY OF PEER REVIEWERS COMMENTS AND FDA RESPONSE

GENERAL IMPRESSIONS

Comments.

Three reviewers expressed support for the model, data, and general approach, stating “This report describes an impressive effort to rank foodborne diseases and foods on the basis of an extensive list of criteria.”, “The FDA has done an admirable job in building a science-based model for risk ranking and in developing a report that makes it accessible to a reader with no previous experience in risk ranking or modeling.”, and “The aim of the document to determine the risk of food-hazard pairs is clear and as such it is a useful and valuable approach which could benefit public health authorities as well as other stake-holders of the food chain.”, respectively. A fourth reviewer focused comments on approaches to the task of risk ranking, suggesting “it would be a valuable experiment to remove the epidemiological data from the scoring, and use it instead to assess the reliability of the relative risk estimates of the risk assessment model based on Criteria 2 – 7 only, for example.” Reviewers also made some suggestions and additional comments/concerns discussed in more detail in their comments to the specific charge questions.

FDA Response:

We appreciate the comments and suggestions provided by the reviewers. We conducted a sensitivity analysis varying criteria weights which provides insights into the impact of each criterion on the results. The RRM-PT by statutory mandate must consider all six factors required in FSMA 204 (21 U.S. Code § 2223), i.e., all seven criteria in the RRM-PT are necessary for RRM-PT model. Responses to the additional comments are discussed below under our response to each of the charge questions.

QUESTION 1.

Comments 1a.

The reviewers commented that the choice of the categorization scheme is justified, and it appears that all FDA-regulated food products can be placed in this classification scheme. The list of foods seems appropriate having been based on commodity groups that fall within FDA jurisdiction. At the level of primary commodities, it is helpful to categorize broad food categories allowing to fine tune at a more granular level. The described rationale for use of RFR-adapted primary categories is reasonable, and the definition of secondary commodities under these primary ones facilitates the use of data and information at different levels. The list is representative of (and more detailed than) the FDA’s RFR commodities definitions. One reviewer raised concerns about the specificity of multi-ingredient foods, that no information is provided about the original contaminated ingredient in multi-ingredient foods, and that not all

potential food-hazard combinations have been included. Another reviewer suggested we include additional foods not regulated by FDA.

FDA response:

We thank the reviewers for their comments on the appropriateness of the food classification scheme. Since the review, we have updated the terminology such that Commodity Category is now used instead of “Primary Commodity,” and “Commodity” is now used instead of “Secondary Commodity.” We have further updated the list of commodities to define them at similar levels of granularity to the extent possible, while taking into consideration the characteristics of the food, the hazard, and the manufacturing process.

Multi-ingredient foods are included in the model where data and knowledge (from outbreaks, recalls, surveillance studies, and experts) suggest that they should be. Contamination of multi-ingredient foods can arise from inclusion of contaminated ingredients but also can arise from environmental contamination during preparation of the multi-ingredient food. In scoring Criterion 1, we attribute an outbreak to a food-hazard pair according to information on the contaminated ingredient, which is consistent with the general premise of the reviewers’ suggestion. Where such attribution is not the case, the outbreak is attributed to the multi-ingredient food in the model.

Regarding hazards associated with multi-ingredient foods in the model, we have included food-hazard pairs based on data and information from outbreak investigations, product sampling, the scientific literature, external expert panels, FDA subject matter experts, and external review panels. For example, we have included *L. monocytogenes* as a potential hazard for prepared foods such as RTE deli salads and sandwiches (arising from environmental contamination). Although we did not attempt to identify every possible food-hazard pair, the pairs that have been identified for these commodities are based available data and expert knowledge. The scores for these multi-ingredient food-hazard pairs in turn provide a solid representation for these commodities in the risk ranking, because the risk score for a commodity is driven by the higher-ranked food-hazard pairs associated with the commodity.

Comments 1b.

The peer reviewers made a variety of comments and suggestions regarding granularity of the food classification and availability of data for scoring. One reviewer was concerned about the availability of data at the secondary commodity level, particularly for multi-ingredient foods and suggested focusing on ingredients in these foods. Another reviewer was concerned about grouping foods together that had dissimilar risks, in particular all finfish in the food-hazard pair associated with methyl-mercury. A third reviewer disagreed with some of the classifications and felt that the distinction between secondary and tertiary levels was not clear. A fourth reviewer noted differences across criteria in the numbers and types of references providing data.

FDA response:

We thank the reviewers for the specific suggestions on how to improve the granularity of the food classification and the definitions of commodities.

We agree with the reviewer that ingredients in a multi-component food should be considered separately in the model and we have done so. For example, ingredients in sandwiches such as leafy greens, tomatoes, cheese, egg, and seafood are considered separately under their respective commodity categories. We also considered multi-ingredient foods, such as in RTE deli salads, because of data suggesting the importance of potential hazards such as those introduced from the food preparation/manufacturing environment, e.g., *L. monocytogenes*, as described above. In terms of contamination data for multi-ingredient foods, we used data collected for the multi-ingredient food (e.g., Luchansky et al. 2017 survey of *L. monocytogenes* in RTE foods such as deli salads and sandwiches), where available, to determine the prevalence or contamination rate for scoring Criterion 3; if such data were not available, we considered the potential for ingredient(s) to be contaminated as appropriate, such as in the expert opinion used to score Criterion 3 for the pair “RTE deli salads - *Cyclospora cayetanensis*”.

We agree with the reviewer that food classification for food-hazard pairs should be appropriate for the hazard and we agree with the reviewer that only large finfish (not all finfish) should be paired with methyl mercury. When this hazard is considered for risk ranking, we will revise the secondary commodities associated with finfish accordingly.

We agree with the reviewer that some of the commodities can be better defined and we have since revised some commodity descriptions. We agree that “what is valid for a particular food is valid for others with a very similar ecology and manufacturing conditions, although not involved so far in outbreaks” and that is why one of the first steps we took was to develop a comprehensive list of food-hazard pairs based on not only outbreaks but also detection of the hazards in foods and expert knowledge. For the foods mentioned by the reviewers as problematic, we have revised the relevant commodity definitions considering the specific concerns expressed.

We agree with the reviewer that details of packaging and storage conditions (as well as shelf life) may influence growth potential and this information is part of the overall information we considered to evaluate pathogen growth. We acknowledge that quantitative data on outbreaks and prevalence are not available for a large proportion of food-hazard pairs. We relied on expert opinions to fill data gaps, such as for scoring Criterion 1 for chemical hazards associated with chronic health effects. In addition to sampling data, we used other sources of information, such as Reportable Food Registry (RFR) reports, recalls, detection reported in the literature, detection association with outbreak investigations as well as expert knowledge for scoring Criterion 3

(likelihood of contamination). In the revised model, for food-hazard pairs involving microbial hazards and marine and plant biotoxins, if there is no reported outbreak for a pair, a score of 0 is assigned to Criterion 1.

For the comments on Criterion 4, we acknowledge the scoring in the draft model relied heavily on a few references and on expert opinions. Subsequent to the peer review, we carried out additional literature searches and through a contract conducted an expert consultation specific to Criteria 3 and 4 to fill data gaps. As recommended by the reviewer, we have reviewed and utilized online predictive microbiology databases and tools, including ComBase, which includes dozens of predictive growth models and tens of thousands of observations on microbial growth rates in foods (ComBase Consortium 2019), and the USDA Pathogen Modeling Program (PMP) (USDA ARS 2019). We have used relevant data and models from these predictive modeling databases, as well as growth studies published in the literature, to score Criterion 4 for relevant food-hazard pairs where appropriate.

Comments 1c.

Reviewers commented on the method and data used to identify food-hazard pairs indicating the list of pairs was built from the ground up, starting with the primary data from the literature or other sources and that this provides a sound scientific basis for all of the food-hazard pairs; the list includes all hazards that historically have been associated with each food; and the methodology is reasonable and potentially allows for the inclusion of most combinations of foods-hazards, Reviewers also shared some concerns and suggestions including a) concern that the model relies primarily on outbreak data for identifying food-hazard pairs; b) concern the method does not allow for the flexibility of allowing considering new, unseen events; c) concern about consistency in identifying food-hazard pairs; and d) concern that inclusion of food-hazard pairs for which evidence of a known or reasonably foreseeable risk is not strong, would significantly influence the risk score for that food.

FDA response:

We thank the reviewers for the positive comments on the appropriateness of the method and data and for the specific suggestions on how the list of the food-hazard pairs could be refined. We have taken these suggestions into consideration in revising our methodology and the list of food-hazard pairs and we have revised the Methods document accordingly.

We agree with the reviewers that it's important to consider data and information beyond outbreaks and recalls. Indeed, we have identified food-hazard pairs based on not only outbreaks and recalls but also sampling data, reported detection of the hazard in the food in the literature, association from risk factors identified in case-control study of sporadic illness in the U.S., food-hazard combinations subjected to risk assessments worldwide, knowledge of subject matter

experts, etc. (see more details in Table C-1 “Considerations in identifying a new food-hazard pair” in the Methods document.

We acknowledge that unseen hazards (i.e., unseen in any lines of evidence described in Table C-1 in the revised Methods document) have not been considered in the model. However, our method does provide substantial flexibility for considering new food-hazard pairs, including the consideration of the potential contamination of foods and ingredients in primary production, processing, and distribution (e.g., environmental contamination). Potential for failure in time-temperature control and potential inadequacy in the control and prevention of contamination of primary ingredients are considered as part of Criterion 5 on the state of industry-wise interventions. Potential for failure in time-temperature control, such as suboptimal time-temperature control during distribution reported in the literature (Pouillot *et al.* 2010), is also considered in Criterion 4 on assessing growth potential of a bacterial pathogen in food.

We have reviewed our methodology, the list of food-hazard pairs, and the concerns identified with specific food-hazard pairs by reviewers and revised each, as appropriate, to provide better consistency. For example, we have revised food-hazard pairs involving *Cronobacter* spp. to include only foods consumed or possibly consumed by infants. We have combined or created new food commodities, as needed, for example, combining tofu and tofu products as a single commodity paired with *Clostridium botulinum* and the other hazards and for seafood, creating two commodities, “Finfish, reduced oxygen-packaged” and “Finfish (dried or salted),” that are paired with *C. botulinum* (U. S. FDA 2019), and “Open ocean finfish (histamine-producing species)” that is paired with Scombroid toxin (Histamine). We have added food-hazard pairs, such as “RTE deli salads - *Cryptosporidium parvum* or other spp.”, which together with the pair involving *Cyclospora cayentanensis* provides a good representation of the available evidence for potential association of food-borne protozoan parasites with this commodity. We also added a pair “RTE deli salads – STEC O157”. We have further considered microbial ecology, manufacturing processes, behavior of microorganisms in food-hazard pairs identification and scoring of criteria 4 and 5. We clarified the revised process by which food-hazard pairs are identified in the Methods document, which includes an appendix to provide a comprehensive list of considerations and information sources for identifying new food-hazard pairs .

Inclusion of a food-hazard pair does not necessarily mean that evidence supports a relatively high score for that pair; pair scores depend on the evidence available across all seven criteria. Further, food-hazard pairs are scored independently rather than relatively. This means that inclusion of one pair in the model does not impact the risk score of any other pair. Through a parallel peer review for the model, we revised the aggregation methodology used to score commodities from the food-hazard pairs associated with each. This new methodology is insensitive to the number of food-hazard pairs associated with a food; the risk score for the commodity is driven by the highest-scored food-hazard pair(s) involved. This means that

inclusion of “trivial” food-hazard pairs (those with very weak evidence or very low food-hazard scores) will not generally influence the commodity risk score. This also means that it is not necessary to include the full range of combinations for all commodities and hazards (estimated at ~24,500 potential food-hazard pairs by the reviewer) in the model as long as the most significant ones, in terms of evidence of known or foreseeable risk, are included.

We agree with the reviewers that outbreaks from other nations can provide useful information and such information (e.g., in technical reports from the European Food Safety Authority, Food Standards Australia New Zealand, Health Canada, etc.) was used in our efforts to identify potential food-hazard pairs. The scoring of the food-hazard pairs, however, is based on data relevant to the U.S. population from consumption of foods available to consumers in the U.S. We have added a pair “Fruits (frozen) - Hepatitis A virus”, and updated the strawberry pair to “Berries (fresh) - Hepatitis A virus”.

Comment 1d.

Reviewers provided some alternative approaches to identifying relevant food-hazard pairs for the model.

FDA response:

We thank the reviewers for the additional comments on practical alternatives to the draft food classification scheme that might be considered. We considered the “Failure Mode Effect Analysis” approach (e.g., Kottapalli *et al.* 2017), and decided not to use it because it does not allow us to readily use available data such as outbreak and contamination data to identify food-hazard pairs. Furthermore, some aspects of what could go wrong and how it could occur are reflected in the scoring definition of Criterion 5 that assesses the current state of industry-wide preventive controls and interventions. As described above, we do not consider meat and poultry products (except game meats) in the model because they are not under FDA jurisdiction. We have implemented the suggestions on cross-check between categories (e.g., see responses above related to tofu and tofu products, fresh berries and frozen fruits). We agree with the reviewers that additive scoring schemes even based on exponential scales have limitations. We choose this scoring scheme that allows us to account for the statutory factors and operationalize them based on data and information relevant to seven criteria in the model. While the scoring scheme does not explicitly enable a “reset to zero” for biocidal processes, this aspect is taken into account in scoring Criterion 5, where the scoring definition involves whether control measures are available and adequate, whether there is evidence for consistent implementation in the industry. We agree with the reviewers on the importance of engaging stakeholders and we have done so through issuing a Federal Register notice (79 FR 6596 (Feb. 4, 2014)) to seek stakeholders’ comments on a draft multi-criteria decision analysis approach, as well as on a variety of issues such as food classification, scoring definitions and other aspects. We have addressed comments from stakeholders in the model development and refinement process.

QUESTION 2

Comments 2a.

The reviewers in general commented that the scoring definitions are well defined for the seven criteria, and that the definitions are appropriately defined for microbial hazard, chemical hazard, and undeclared allergens. The reviewers also provided specific questions/comments and suggestions for further refinement and clarification of the definitions.

FDA response:

We thank the reviewers for the comments about the appropriateness of the scoring definitions for the seven criteria. We revised the description for the Criterion 1 scoring definition to clarify that the number of reported outbreaks and the number of outbreak-related cases are used in the scoring. We agree with the reviewers on the complicated nature of estimating disease incidence for chemical hazards and undeclared allergens, and that assessment of available relevant epidemiological and animal studies is crucial for scoring Criterion 1 for the estimated occurrence of foodborne disease due to exposure to these hazards.

We agree with the reviewers that foodborne illness sequelae (sometimes severe) should be considered. Indeed, by using the cost per case for the hazard from Minor *et al.* 2015 (with 2018 update by FDA for cost adjustment to reflect inflation), the impacts from various sequelae associated with various microbial pathogens are considered in the scoring of Criterion 7. Minor *et al.* 2015 generated estimates of cost of foodborne illness and lost quality-adjusted life years (QALYs) and lost quality-adjusted life days (QALDs), which take into consideration severity as well as sequelae of foodborne illnesses and involves using the average quality weight of a foodborne illness, measured as the index of EuroQol 5 Dimensions (EQ-5D) with preference weights adapted for the U.S. population, analogous to the disability weights approach used in the Global Burden of Disease. In the model, Criteria 2 and 7 reflect different aspects of the public health impact of hazards in foods; together they provide a more complete characterization. As an example, hospitalization rates for *Salmonella* spp. (nontyphoidal) infection and *Campylobacter* infection captured in Criterion 2 are estimated as 27.2% and 17.1%, respectively, while QALD loss including the losses arising from sequelae such as Irritable Bowel Disease and Guillian-Barré Syndrome, captured in Criterion 7, are 1.62 and 2.09, respectively. Similarly, differences in the health impacts of *Salmonella* spp. (nontyphoidal) and *L. monocytogenes* infections are captured in the model; in this case differentiation is provided by Criterion 7. Cost per case is vastly different, \$6,563 and \$1,797,753 (2018 estimate) for *Salmonella* spp. (nontyphoidal) and *L. monocytogenes* infection, respectively. (Hospitalization and mortality rates for infection with these pathogens, while different, are both sufficiently large that food-hazard pairs involving them receive the same categorical score of 9 for Criterion 2).

We have clarified in the text of the Methods document that, where available, hospitalization and mortality rates are used in scoring Criterion 2; when these data are not available, scoring is based on the qualitative definition and expert opinions with consideration of relevant studies. For Criterion 3, the scoring definition refers to % of samples and when sampling data are available, they are used preferentially in the scoring. Regarding the definition for Criterion 4, we have clarified in the Method document that knowledge on the customary shelf life of the product and how the product is stored includes considerations of the potential for growth under moderate abuse conditions.

Regarding the comment on consumption, the revised Criterion 6 takes into account two risk-influencing aspects of consumption: amount consumed per serving, which influences the risk per serving and the percentage of the U.S. population that consumes the food, which influences the per annum risk. We considered a lower weight for Criterion 6, suggested by one reviewer, in the sensitivity analysis provided that explores non-equal weighting options for the seven criteria - see details in Methods document.

We have identified staphylococcal enterotoxins as the hazard for food-hazard pairs involving *S. aureus*, and placed mycotoxins (a toxin of microbial origin) under the microbial hazard category to allow consideration of the impact of industry-wide interventions on the formation of mycotoxins, including control measures for storage conditions and shelf-life. Regarding the additional undeclared allergen hazards (“sub-types” referred to by the reviewer) defined, e.g., undeclared allergens (other than nuts), these are used to pair with a food that itself is an allergen (e.g., mixed nuts) where the potential hazard of undeclared allergens are those other than nuts in the pair “Mixed Nuts – Undeclared Allergens (other than nuts).” We have added a footnote to Appendix B in the Methods document to clarify.

Comments 2b.

The peer reviewers disagreed as to whether to apply differing weights to Criterion 3 for ready-to-eat (RTE) vs. not-ready-to-eat (NRTE) food-hazard pairs.

FDA response:

Taking into consideration all the comments from the reviewers, we decided not to apply a different weight to Criterion 3 for a ready-to-eat (RTE) vs. a not-ready-to-eat (NRTE) food-hazard pair. We agree that because of the diverse nature of the different hazard types (e.g., microbial pathogens, heat-stable toxins produced in the food from growth of certain bacterial pathogens, chemical hazards not sensitive to heat inactivation), adjustment to the scoring scheme required to accommodate this issue is more complex than a simple weighting for RTE vs. NRTE could accurately reflect. We also agree that the likelihood of contamination in the finished product is the outcome of food supply chain and the control measures applied during manufacturing up to consumption, and data collected for Criterion 3 reflect the percent

contamination rate in the finished products consumers purchase and should already reflect these differences. Further, scoring of Criterion 5 includes considerations related to whether the food is RTE and NRTE foods. In Criterion 5, for example, industry-wide interventions are indeed considered for the products in the form that is available to the consumers. Consumer cooking is not considered in the model.

QUESTION 3

Comments.

The reviewers commented that in general the data sources are comprehensive and appropriate to support criteria scoring according to criteria definitions. One reviewer noted that the references cited in the draft report did not clearly communicate the wide range of credible sources that could be used by FDA to score the criteria. Several reviewers suggested specific additional references for FDA to consider. Two reviewers re-iterated comments made on criteria definitions (Question 2); one asked that more information be provided in the Methods document as to the reasoning used to develop the scoring definitions.

FDA response:

We thank the reviewers for the comments about the appropriateness of the data sources used. We appreciate the specific suggestions made to improve the ways some of the data sources are used and suggestions on additional data sources and references to consider (which we have done). We have also developed a separate (comprehensive) reference list that captures all data sources evaluated and used in scoring the model.

We agree with the reviewers that outbreaks generally represent a fraction of total cases, and this proportion varies and is particularly small for some pathogens, such as the examples indicated in the comments. We agree that estimating the “true” incidence of a wide variety of foodborne diseases associated with consumption of specific foods would require data on both outbreaks and sporadic cases. However, currently, data are not available for sporadic cases and this is a limitation. We have reviewed various attribution references (e.g., Batz *et al.* 2012, Pires *et al.* 2009, Pires 2013, Scallan *et al.* 2015). The Scallan *et al.* 2015 study estimated the burden of foodborne illness of 31 pathogens in the U.S. using disability-adjusted life year (DALY) as a metric. However, because the study reported annual DALYs per pathogen from all food sources, it is not feasible to determine DALYs or estimated outbreak and sporadic cases for specific food-hazard pairs. We also choose cost per case reported by Minor *et al.* 2015 (with adjustment for 2018 costs) as a data source for scoring Criterion 7, and use the QALD values reported (Minor *et al.* 2015) as the basis to determine severity, in a sensitivity analysis on alternative scoring for Criterion 2. As described above, the QALD value reflects average quality weight of a foodborne illness (measured as the EQ-5D index with preference weights adapted for the U.S. population).

Using QALD values (and thus taking into account quality weights) serves a purpose similar to using disability weights. Furthermore, Minor *et al.* 2015 reported QALDs and cost of illness values for the 31 pathogens reported by Scallan *et al.* 2015, as well as for an additional pathogen, two marine biotoxins, and foodborne allergic reaction. Thus, we decided to use data reported by Minor *et al.* 2015 (with adjustment for 2018 costs). Regarding the suggestion for using underreporting and underdiagnosis multipliers for Criterion 1 as well as Criterion 7, we believe that in a multicriteria decision analysis methodology, it's more robust to keep Criteria 1 and 7 mutually independent as much as possible, which includes minimizing overlapping in data. Therefore, we decided to use the multipliers for scoring only Criterion 7, not for both criteria 2 and 7.

We agree with the reviewers that the task of scoring severity and cost of illness is particularly difficult for chemical hazards because, as the reviewers indicated, some are (considered potentially) carcinogenic, and the probability of a given individual person developing disease is determined by many other factors. We plan to further evaluate the estimates from the global burden of foodborne diseases (FERG, http://www.who.int/foodsafety/areas_work/foodborne-diseases/ferg/en/) in particular the burden of some chemical hazards (at the regional level) in updating the scoring of food-hazard pairs involving chemical hazards associated with adverse chronic health effects, if these hazards are considered in a finalized rule.

Regarding the comments on Criterion 2, we agree with the reviewers that Scallan *et al.* 2011 did not specifically consider sequelae and the reported hospitalization and mortality rates were based on laboratory confirmed cases. We considered the aspects related to sequelae and QALY loss (and QALD loss) in scoring Criterion 7 by using the cost per case values from Minor *et al.* 2015. The sequelae for the pathogens indicated by the reviewers (e.g., *Salmonella* infections can lead to reactive arthritis and other diseases, *Campylobacter* to these and to Guillian-Barré Syndrome, STEC O157 to end stage renal disease including hemolytic uremic syndrome) are considered by Minor *et al.* in developing QALYs and cost per case values. The approach we take helps ensure that overall a complete picture of severity is considered in the model while keeping criteria 2 and 7 independent and as mutually exclusive as feasible. As described above, instead of using disability weight, we use QALD values as a basis for scoring severity and conducted sensitivity analysis for an alternative scenario for scoring Criterion 2. Regarding the comments on Criterion 7, we understand the reviewers' concern that, because underdiagnosis/underreporting multipliers reported in Scallan *et al.* 2011 were derived to correct all illnesses (outbreak and sporadic cases included) and that outbreak-related cases are more likely to be reported, applying the multipliers to outbreak-related cases may not be appropriate for all pathogens. This is a limitation due to the lack of data on sporadic cases. As the reviewers pointed out, there are also limitations in general for using outbreak cases for attribution and this may be improved in the future when more data on sporadic cases become available.

Regarding underreporting factor for Norovirus, we agree with the reviewers' reasoning and that based on the 5,461,731 foodborne cases estimated in Scallan *et al.* 2011, and 2009-2010 shows 69,145 Norovirus outbreak-related reported cases (Hall *et al.* 2013), an estimated underreporting factor of 79.0 might be derived. Since Scallan *et al.* (2011) did not report underreporting and underdiagnosis multipliers for Norovirus, we will consider the reviewers' suggestion in the future when revised multipliers are reported in a peer-reviewed paper. We thank the reviewers for the comment regarding the use of the Norovirus cost per case for seafood parasites. This was based on suggestions from external subject matter experts. We will re-evaluate symptoms and duration of disease for seafood parasites and, if appropriate, revise the value for seafood parasites in the future. Regarding the comment on the scientific basis for the experts' input, we relied on a structured process for expert elicitations, where, through contracts, each of the expert panels consisted of subject matter experts with a range of relevant expertise and the experts provided input through a structured process (see more details in the response to Question 7 comments below).

As the reviewers pointed out, the list of food-hazard pairs were identified using outbreak data, as well as other data sources such as detection of the hazard in food reported in the literature, recalls, FDA sampling data, published risk assessments, and subject matter expert knowledge. In some cases, pairs may have been identified because of screening or survey conducted by an organization. In identifying food-hazard pairs based on known or foreseeable association of the food and the hazard, we did not pre-determine whether or not the pair would score high or low in the model (the scoring would be determined in a subsequent step by using data and information available for the seven criteria). As described in the Methods document, the process used to identify a comprehensive list of food-hazard pairs without requiring an *a priori* assumption about the score the food-hazard pair might receive is a strength of the methodology.

We thank the reviewers for the additional suggestions on data sources, which we have considered in updating the model. For example, we have revised the scoring definition for Criterion 4 to include a quantitative measure of growth (e.g., likely growth means growth ≥ 3 log₁₀CFU based on published study or predictive microbiology models). The growth determination uses available data on growth based on published data on observed growth in a food-hazard pair and, where appropriate, by expected growth estimated using predictive microbiology database and modeling tools such as ComBase (ComBase Consortium 2019) and the USDA PMP (USDA ARS 2019). We use knowledge on the customary shelf life of the product and how the product is stored (including consideration of potential time temperature abuse during retail and home storage). We have searched the large curated datasets deposited in ComBase, reviewed growth models for various food and pathogen combinations in ComBase and the USDA PMP program, and use the data and model predictions as appropriate for relevant food-hazard pairs, according to matching food-hazard pairs, conditions and characteristics of food (e.g., pH and *a_w*). We thank the reviewers for the suggestion on using supermarket sales as a

potential data source for consumption evaluation. Instead, where data are not available in the National Health and Nutrition Examination Survey (NHANES) What We Eat in America (WWEIA) database, we rely on consumption subject matter experts and other surveys reported in the literature.

QUESTION 4

Comments.

Two reviewers indicated the weighting factors were appropriate. One reviewer agreed with the approach but thought that different factors should be derived for each hazard. One reviewer wasn't sure weighting factors should be applied.

FDA response:

We adopted a data weighting scheme used in the 2003 FDA/FSIS *L. monocytogenes* risk assessment (US DHHS FDA/USDA FSIS 2003) to address relevance of data, where more recent outbreaks and sampling data from more recent surveys have a greater impact on the criterion score. Sampling data are also weighted by sample size and geographic region. We agree with the reviewers that industry/regulators have taken measures to reduce the incidence of a specific foodborne illness by new methods and interventions over the past twenty years. The “date” weighting of outbreaks and sampling data acknowledges such changes may have occurred by placing greater emphasis (weight) on more recent data which are more reflective of the current status. We also agree with the reviewers that control measures once validated, are an indication of adequate control measures available. This is considered in scoring Criterion 5, where both (a) adequate control measures available and (b) evidence for consistent implementation in industry are considered including a long history of implementing adequate control measures.

We recognize that improvements in detection sensitivity or detection limits for chemical hazards can lead to apparent increases in prevalence for chemical hazards if prevalence is defined simply as “detected”. However, in this model “prevalence” is defined as presence at a level above an action level or allowable level. We recognize that the issue remains for some chemical hazards and undeclared allergens where an action level or allowable level has not been established. We agree with the reviewer that PCR-based pathogen detection technologies may not differentiate viable from non-viable cells. We address this issue by evaluating whether cultural confirmation was done as part of the sampling and testing, and determining the number of positive samples based on culture-confirmed positives where data are available (where a study reported only positive samples from PCR-based test, this is documented). We agree that changes in controls, regulations, implementation, or hazard detection across twenty years and their impact on the frequency of outbreaks, numbers of illnesses per outbreak, and prevalence of hazards in foods, likely vary across food-hazard pairs. However, we do not agree that this means it would be better

to derive hazard-specific weights for the outbreak and prevalence data for this risk ranking model. Instead, the current strategy emphasizes more recent data (or discounts older data) to the same extent (same weights) for all hazards thereby giving more recent data the same importance in the scoring across all food-hazard pairs.

QUESTION 5

Comments.

The reviewers commented that the time frame seems appropriate. Several reviewers commented that because the purpose is to estimate the proportion for the population that is exposed (i.e., consumes each food), they had no concerns with using the rate from the day with highest consumption and indicated that the use of highest consumption is appropriate since it will produce the most conservative model. One reviewer commented that given the size of the samples, the average seems a more reliable and logical way of synthesizing the available data to make the most representative assessment and suggested the justification/explanation of the use of the maximum (or sum?) and more evaluation of the effects of that decision are required for transparency.

FDA response:

We thank the reviewers for confirming the appropriateness of the time frame for consumption data from the NHANES WWEIA database (2009-2010 and 2011-2012 data for the 2015 draft report for peer review). According to the same principle, in the updated model (2019 Methods report) we now use NHANES consumption data from the latest three cycles (2015-2016, 2013-2014, and 2011-2012) to estimate consumption.

Regarding the comments on use of the sum of consumption rates (vs. mean or maximum consumption rate), we choose the sum based on considerations of the NHANES data and how the data are used to calculate consumption rate for a commodity in the model. A commodity in the model may include several different food items (e.g., the commodity “Leafy greens” includes lettuce, spinach, etc.). The consumption analysis involves first calculating a consumption rate for each of food items that belong to the commodity. Consumption rate for the commodity is calculated by summing the consumption for each of the associated food items (i.e., NHANES food codes) that are mapped to the commodity. For example, consider three food items (A, B, and C) that are mapped to Commodity Z, and the consumption rates for these food items are 17.6%, 2.57% and 0.865%, respectively. It is most appropriate to use the sum of consumption rates (21.0%) for scoring the consumption of Commodity Z (rather than the mean or maximum values) because it represents most accurately the consumers who consumed any one of the three foods in Commodity Z. Using the mean or the maximum rate would underestimate the consumption rate for the commodity.

Based on other comments from the reviewers (see below), we include amount consumed (in addition to consumption rate) in the scoring definition for Criterion 6.

QUESTION 6

Comments.

The reviewers provided specific comments on the data and references used for scoring a number of food-hazard pairs, after their in-depth evaluation of risk scores and the appropriateness of cited references and data used for a subset of food-hazard pairs in selected primary commodities and hazards within their areas of expertise.

FDA response:

We thank the reviewers for evaluating the scoring as well as the data and rationale for selected food-hazard pairs, and for providing comments on the appropriateness of the underpinning data used and how the process of data collection and documentation might be refined. We thank the reviewers on the positive feedback, such as on the food-hazard pairs involving *C. botulinum*. Based on the reviewers' suggestions, we have refined the data and scoring for food-hazard pairs and organized our responses in a few sets as follows.

Response 6-1.

We thank the reviewers for the in-depth review of the references cited in the scoring of the finfish – methyl mercury pair in the draft report. We have considered the references and rationale suggested by the reviewers on the food-hazard pair involving Methyl mercury, and plan to use the data to score the pair “Finfish (Large, species associated with highest mercury levels) - Methyl mercury”. Regarding the comments related to how the references are cited and availability of the references, we have systematically updated how the references are cited and now provide the full citation (including the title of the document) for each of references evaluated and used in the model.

We thank the reviews for providing a number of additional references to strengthen the data for scoring the selected food-hazard pairs. We have considered the references suggested by the reviewers and also included additional authoritative references in model revision. We have corrected errors in reference citations or usages. To address reference-related issues the reviewers indicated here and elsewhere, we have conducted additional literature search as well as conducting an expert consultation specific to criteria 3 and 4 to fill data gaps (Beuchat *et al.* 2016). For example, to support the scoring of Criterion 4 for the *L. monocytogenes* in cold smoked finfish pair, we now include several new studies and predictive models (*e.g.*, Dalgaard and Jørgensen 1998, IFT 2001, Jinneman *et al.* 2007, Pouillot *et al.* 2007, and Pouillot *et al.*

2009, Mejlholm *et al.* 2010). To support the scoring of the pair “Canned fruits and vegetables – *C. botulinum*”, we added two new references with relevant information of growth and inhibition of microorganisms in foods (Montville and Mathews 2013) and microbiology of canned foods including low acid canned foods (Gavin and Weddig 1995). We now use the correct cost per case of botulism (\$1,849,016 from Minor *et al.* 2015 with 2018 update). We have included ICMSF Book 5 (ICMSF 2005) and Book 8 (ICMSF, 2011) as references for information on microbial ecology, the role and impact of contamination along the food chain, the impact of sampling and testing, verification and validation of control measures, etc., and considered it as a reference in the scoring of Criteria 5. Regarding references listed in Appendix E of the draft report for peer review, we have conducted further QA/QC of data obtained from the references. In updating the model, through a contract with RTI International we have developed an SQL database (a relational database) to compile and connect data for the seven criteria and all the references.

Response 6-2.

With regard to *Cronobacter* spp., since this hazard poses risk primarily to infants who do not usually consume scones, we have removed the pair “Scones - *Cronobacter* spp.” and similarly the pair “Confections or coatings - *Cronobacter* spp.” from the model. Regarding the suggestions on the pairs for dried milk, we agree with the reviewers on the risk associated with *Salmonella* and keep the pair “Dried milk – *Salmonella* spp.” We also keep the pairs “Dried milk – *Cronobacter* spp.” and “Dry instant breakfast – *Cronobacter* spp.” because there might be a potential for infant to consume dried milk or dry instant breakfast in their diet. We did remove the pair “Confections or coatings - *Cronobacter* spp.” from the model because the chance of infant consuming confections is extremely remote. Regarding the scoring of Criterion 1 for the pairs “Infant formula – *Cronobacter* spp.” and “Infant formula – *Salmonella* spp.”, we agree with the reviewers that a number of outbreaks have occurred worldwide. Based on updated outbreak data in the U.S., the criterion score=1 is now assigned for both pairs in the model.

Response 6-3.

Regarding outbreak data associated with *Salmonella*, we have updated the data to now include outbreak data up to mid-2019. Accordingly, outbreak data included in the model now span 1999-2019 to include outbreak data in a 20-year period, where the most recent outbreaks are weighted most heavily. As part of updating the model and scoring, we have further refined how outbreaks are assigned to a food-hazard pair. Notably, an outbreak is assigned to a food-hazard pair first using contaminated ingredient information and, if contaminated ingredient was not identified from outbreak investigations, using information on the food vehicle implicated. For example, outbreaks that were assigned to the pairs “Fresh salsa – *Salmonella* spp.” and “Sandwiches – *Salmonella* spp.” in the draft report have been re-assigned according to information about the contaminated ingredient, where that information was available. For example, an outbreak where tomato was the contaminated ingredient in fresh salsa (implicated as a vehicle) is now assigned

to a pair involving tomato. The scoring of Criterion 1 for the four pairs have been updated accordingly; among the four *Salmonella* pairs, only two pairs (involving shell eggs and tomatoes) currently have score=9 for Criterion 1.

We acknowledge that pairs that fit into scoring 9 for Criterion 1 according to the scoring bins can differ in the number of outbreaks and cases, their impact as measured by Criterion 7 can be different depending on the hazard and the number of cases.

Regarding the comments on *Campylobacter*, we agree with the reviewers that it is a leading cause of foodborne illnesses in the U.S. We have evaluated approximately two dozen food-hazard pairs involving *Campylobacter* spp. and generated risk scores for these pairs based on the data and information available for the seven criteria.

Regarding the comment on products made with a wide variety of ingredients (such as fresh salsa and sandwiches), we acknowledge that including these products in the model introduces another layer of complexity. We consider multi-ingredient foods because of data suggesting the importance of potential hazards such as those introduced from the food preparation/manufacturing environment. An objective of the model development was to develop a comprehensive list of food-hazard pairs for each of the commodity categories of FDA-regulated human foods, to be used as candidates for scoring in the model. As described above, we have refined the scoring definitions to place greater emphasis on the ingredient than on the multi-ingredient food, e.g., for Criterion 1, contaminated ingredient identified in an outbreak is used first for outbreak assignment. As described in our response above, in some instances, there is potential for a hazard to be introduced during the preparation of multi-ingredient foods, such as environmental contamination of *L. monocytogenes* during manufacturing of RTE deli salads and sandwiches; these scenarios are considered in the model. We agree with the reviewers that for sandwiches, contamination data were very poor in the draft model for peer review. In the revised model, we now include more data from recent surveys and literature review, e.g., for the pair “Sandwiches – *L. monocytogenes*”, we now have data on approximately 2300 samples from a market basket survey in the U.S. (Luchansky *et al.* 2017).

We agree with the reviewers’ comments on arsenic in rice and that the link between exposure to arsenic and potential onset of cancer is one of the better established scientifically. Regarding the comments on the data used to support a score of 9 for Criterion 3, a reference is the FDA risk assessment on arsenic in rice and rice products (U.S. FDA 2016a). Considering that an action level in the U. S. has not been established, a score=9 is assigned for Criterion 3 for the pair “Rice – Arsenic (inorganic)” based on the data and expert judgment.

Response 6-4.

We agree with the peer reviewers on the potential risk of *Bacillus cereus* in pasteurized milk and have added this pair to the model. Furthermore, we have considered the references provided by the reviewers, as well as other references and relevant data and information to score seven criteria for the pair “Milk (fluid and white and Grade-A pasteurized) – *Bacillus cereus* and enterotoxin”. Regarding the comment on *L. monocytogenes* in cooked meats as a potential pair, this pair is not included; only FDA-regulated human foods are evaluated in this model. As well, we appreciate the references provided. We have considered the suggestions and references (e.g., Christiansson *et al.* 1989, Te Giffel *et al.* 1996, Zwietering *et al.* 1996, Notermans *et al.* 1997) in the revised version of the model and report.

Regarding comments on the pair for *L. monocytogenes* in soft-ripened cheese, we concur with the reviewers on the comments on the rationale behind the scoring of Criterion 1; in the model update, there are additional outbreaks in the U.S. since the time of peer review, and the new data are reflected in the updated scoring of this criterion. Regarding the comments on Criterion 2, we acknowledge that susceptible population are at a greater risk for listeriosis and that the differences in susceptibility may be up to 100 times or greater according to dose-response analysis (e.g., Pouillot *et al.* 2015). However, we believe applying the scoring definition for Criterion 2 with consideration of the entire U.S. population, a score of 9 is still appropriate for *L. monocytogenes*. For Criterion 2, the scoring definition involves using hospitalization and mortality rates defined as the % of laboratory-confirmed cases; thus, we believe it is appropriate to use the death rate in Table 3 in Scallan *et al.* 2011 without adjustment. For comments on Criterion 3, in the revised model, we have updated the references (Bernini *et al.* 2013, Gombas *et al.* 2003, Lambertz *et al.* 2012, Luchansky *et al.* 2017, Martinez-Rios and Dalgaard 2018, and Ryser 2007) from which we extract data to calculate contamination rate for scoring Criterion 3 for the pair “Cheese (made from pasteurized milk), soft or soft ripened or semi-soft – *L. monocytogenes*”. For comments on Criterion 4, in the revised model, we have updated the references that now include examples suggested by the reviewers (Mejlholm and Dalgaard 2009, Mejlholm *et al.* 2010) and other relevant studies and predictive models (e.g., U.S. FDA 2017, ConBase Consortium 2019, Schwartzman *et al.* 2011, Tiwari *et al.* 2014, and Wemmenhove *et al.* 2016). We have double checked the cost per case and now have the corrected value from Minor *et al.* 2016 (with 2018 cost adjustment to reflect inflation).

Regarding comments on the pair for STEC in bagged leafy green salads, we considered the reviewers’ comments and we believe that a score=9 for STEC O157 is appropriate to distinguish its severity from the severity of STEC non-O157 (score=3). Of note, we have refined the name of the pairs to “Leafy greens (fresh-cut) – STEC O157” and “Leafy greens (fresh-cut) – STEC non-O157”. Minor *et al.* 2015 showed differences in severity, i.e., QALD of 4.61 for STEC O157 and 1.59 for STEC non-O157, which parallel the differences reported by Scallan *et al.* 2011, i.e., hospitalization rate of 46.2% for STEC O15 and 12.8% for STEC non-O157. Based on the

reviewers' suggestion, we have conducted a sensitivity analysis with the QALD values from Minor *et al.* 2015 as an alternative scoring for Criterion 2. The alternative Criterion 2 scenario resulted in a shift to a lower risk score for some of food-hazard pairs, particularly food-undeclared allergen pairs and pairs involving *Brucella* spp., *Cryptosporidium parvum* and other spp., *Salmonella* spp., *Shigella* spp., and *Y. enterocolitica*, *V. parahaemolyticus* and *V. cholerae*. Regarding Criterion 4, to further support the scoring, we have also added references, including predictive models from ComBase, e.g., Posada-Izquierdo *et al.* 2013 and Posada-Izquierdo *et al.* 2014.

Regarding comments on the pair for Hepatitis A in frozen berry fruits, for Criterion 1, the scoring was indeed based on outbreak data in the U.S. We recognize, as the reviewers noted, that there have been many outbreaks Hepatitis A from frozen berry fruits in Australia and Europe. However, only US outbreak data are used in the scoring because the risk ranking is for the U.S. population. Nonetheless, additional outbreaks attributed to Hepatitis A in frozen fruits (including berry fruits) has occurred in the U.S. since the peer review, and we have used the latest available data to update Criterion score for the pair "Fruits (frozen) – Hepatitis A". For Criterion 2, we have considered the reviewers' comments and re-evaluated the severity score for Hepatitis A. There are pros and cons in determining severity scores based on hospitalization and mortality rates vs. QALD values. We have decided to use the former from Scallan *et al.* (2011) in scoring severity for Criterion 2 and use the latter (which is accounted for in the cost of illness estimates) in scoring cost of illness for Criterion 7. This approach allows us to consider both hospitalization and mortality rates and QALDs in different ways in the model, while keeping criteria 2 and 7 as mutually independent as possible. For Criterion 3, we agree with the reviewers that available evidence suggests relatively low frequency and level of contamination (Maunula *et al.* 2013). In the revised model, score=1 for Criterion 3 is obtained based on data reported in Maunula *et al.* 2013 as well as the detection of Hepatitis A in frozen fruits in FDA investigations of Hepatitis A outbreaks linked to frozen berries, e.g., an outbreak in 2016 linked to frozen strawberries (U.S. FDA 2016b). There is an ongoing sampling assignment (U.S. FDA 2020) for Hepatitis A and Norovirus in frozen berries (including strawberries, raspberries, and blackberries) and data from this surveillance sampling can be used in the future in the scoring for this pair. For Criterion 5, we agree that highly effective interventions against Hepatitis A on fresh fruits (prior to frozen) are not yet available. We have considered the reviewers' comments in updating the model. In expert elicitation and QA/QC after the peer review, overall Criterion 5 score remains 3 for this pair.

QUESTION 7

Comments 7-1.

Peer reviewers indicated that in general the expert elicitation process (which involves external panels and FDA subject matter experts) was adequate to address data gaps, assuming numbers of experts are sufficient, and their experience is appropriate. One reviewer stated “In this project, I feel they [expert elicitations] were the best approach to address data gaps, and the panels seemed to be varied in the sense that they included experts with experience in different fields within food safety.” Another said, “The range of experts used seems appropriate and sufficient to generate credible responses to the missing data particularly when combined with a public comment process to enable further input and opinion.” One reviewer was concerned that the numbers of reviewers were too small, at least in some areas. With regard to the scoring logic used to combine data and expert opinions and the order of preference for use of expert opinion as opposed to other types of data (typically only used when no other data was available), one reviewer said “The scoring method used to combine available data and expert input seems reasonable” and another said “The hierarchy is appropriate”; the other two reviewers did not comment on this aspect. Several reviewers wanted more clarity with regard to the experts and the process and provided some comments of ways strategies and processes to optimize the accuracy of the elicitation. One reviewer was concerned that data were evaluated for food-hazard pairs for which no outbreaks, recalls or contamination were evident.

FDA response:

In general, the expert scoring reflected the consensus opinion of multiple SMEs. FDA contracted with RTI International and the Institute of Food Technologists (IFT) to provide technical and logistical support to assist in the process of developing the risk-ranking model; three expert elicitations were conducted to fill data gaps. Members of the external expert panels (Table 7-1) were chosen to include a wide range of relevant expertise and consisted of 13 SMEs in the fields of food safety, food production and processing, microbiology, risk assessment, toxicology and chemistry, and allergens. As an example of the process followed, in the IFT/RTI elicitation process, SMEs reviewed background materials on the model and detailed information on the scoring process for each criterion, individually scored criteria for the identified data gaps, and participated in panel discussions (face-to-face meetings and conference calls) convened by RTI/IFT, which included a structured process to determine consensus scores. As an example of the information formation provided to experts or scoring, each was provided with an overview of the scoring definitions and examples, such as how to use the order of preference for types of data (e.g., when to use contamination rate based on sampling data or when to use RFR data to score Criterion 3).

Table 7-1. Composition of expert panels in the external expert elicitation process

Area of Expertise	Affiliation	Subject Matter Expert
Food Production & Processing, Food Microbiology	Flying Food: Large-scale catering company providing airlines retail partners with meals and snacks	Dr. Paul Hall ^{a, b}
Food Production & Processing, Food Microbiology	JC Rosen Resources: Consulting services for produce and food industries, expertise throughout supply chain	Dr. Joan Rosen ^{a, b}
Food Production & Processing, Food Microbiology	Chiquita: Producer and distributor of fresh produce	Dr. Courtney Parker ^{a, b}
Food Safety and Microbiology	University of Georgia: Academia	Dr. Larry Beuchat ^{a, b, c}
Food Safety and Chemistry, Allergens	Nestle: Multinational food and beverage company	Dr. Jupiter Yeung ^{a, b, c}
Food Safety and Risk Assessment	Rutgers University: Academia	Dr. Donald Schaffner ^{a, b, c}
Toxicology, Food Chemistry	Coughlin & Associates: Consulting services in food, nutrition, chemical toxicology, safety	Dr. James Coughlin ^{a, b}
Food Safety and Epidemiology	University of Minnesota: Academia	Dr. Craig Hedberg ^a
Food Safety, Food Production and Processing	University of Minnesota: Academia	Dr. Frank Busta ^a
Toxicology	Harvard University: Academia	Dr. A. Wallace Hayes ^b
Toxicology	University of Minnesota: Academia	Dr. Ken Wallace ^b
Food Safety and Allergies	University of Nebraska - Lincoln: Academia	Dr. Steve Taylor ^c
Food Safety and Risk Assessment	University of California - Davis: Academia	Dr. Carl Winters ^c

^a These SMEs were also involved in the first expert panel.

^b These SMEs were involved in the second expert panel.

^c These SMEs were involved in the third expert panel.

Using a similar structured process, FDA SMEs conducted QA/QC of Criterion 5 scores in 2015 (before the peer review), and again in 2019 (after the peer review) taking into consideration the

current state of industry-wide interventions in light of FSMA implementation and other regulations and guidance. As an example of this process, in 2019 Versar, Inc. conducted a QA/QC of the scoring of Criterion 5 for all food-hazard pairs involving microbial hazards and acute chemical hazards with FDA SMEs. The pairs were divided into five commodity groups: dairy, eggs, and meat; plants products and beverages; produce; seafood; and multi-commodity foods. In total, 13 FDA SMEs with expertise in the five commodity groups reviewed the food-hazard pairs. The group makeup was static during the structured QA/QC process. SMEs first individually determined for each food-hazard pair in the commodity group, whether a change in industry-wide practices or implementation had occurred since 2015/2017 (depending on the data it was last scored) or whether knowledge of industry-wide practices or implementation had significantly changed. Next, SMEs scored the pairs for which they had identified a change. Finally, contractor expert elicitation professionals facilitated consensus-building among the SMEs to arrive at a consensus score. Overall, the FDA SMEs recommended changes to the Criterion 5 score for approximately 8% of food-hazard pairs. Some changes to scores from 2015/2017 were anticipated considering implementation of FSMA and other regulations and guidance since the external expert elicitations (in 2015 or 2017). With regard to capturing rationale for certain decisions, we have relied on a structured expert elicitation process that involved SMEs with relevant expertise, providing necessary background materials, and consensus scoring after individual scoring, to capture the collective knowledge and expertise of the SMEs who participated.

We agree with the reviewers and have revised the method for scoring Criterion 1, such that a food-hazard pair involving a microbial hazard, toxin of microbial origin, or biotoxin for which no outbreaks and no occurrence of illnesses have been reported, would have a score of C1=0 assigned. However, if a food-hazard pair is identified based on evidence other than outbreaks (such as detection of a pathogen in a food), we believe it is valid to use the model to score C2 through C7, even though C1=0 would be assigned. This is a strength rather than a weakness of the model because it allows us to use risk factors beyond outbreak association. With regard to *Cronobacter* spp., since this hazard poses risk primarily to infants who do not usually consume scones, we have removed this pair and the pair “Confections or coatings - *Cronobacter* spp.” from the model.

QUESTION 8

Comments:

The reviewers did not recommend scaling down consumption to adjust for the fact that a restricted subpopulation only may be susceptible or may be more susceptible or have more severe health impacts than the general population. Reviewers had several suggestions on how best to capture these differences in susceptibility and severity within the population.

FDA response:

We thank the reviewers for the comments and specific suggestions. Based on the reviewers' comments, we have decided not to scale the consumption scoring specifically for any hazard (undeclared allergen, chemical hazard or microbial hazard) for susceptible populations.

We score the food-hazard pairs based on consideration of both the average risk to the entire population and the risk to "at risk" subpopulation as appropriate, such as risk from *Cronobacter* spp. (where only foods for infants are considered) and risk from *L. monocytogenes* (where susceptible populations are considered in the scoring definition for criteria 2 and 7). We acknowledge that for the scoring definition for Criterion 2, where quantitative data are used, mortality and hospitalization rates are based on laboratory-confirmed cases. As the reviewer pointed out, the total estimated number of cases per year is considered in the scoring definition for Criterion 7. We believe that this approach has two advantages: i) takes into account severity based on confirmed cases and total estimated cases in two different ways (two different criteria), and ii) minimizes "double-counting" of severity because Criterion 2 and Criterion 7 address separate aspects of severity. We thank the reviewers for providing the Goulet *et al.* reference and have included it for scoring severity for *L. monocytogenes* (FAO/WHO 2004, Goulet *et al.* 2012, Pouillot *et al.* 2015).

We believe it is not feasible to generate two different risk scores, one for the general population and one for the susceptible population, as one reviewer suggested, because data available for the criteria usually are not specific to the susceptible population, *e.g.*, cases associated with outbreaks are for the general population, contamination rate data from food survey is usually for products available to the general population, and consumption data from NHANES are usually for the general population. Adverse health effects for the susceptible population as well as the general population are considered in the scoring definitions for Criterion 2 (severity) and Criterion 7 (cost of illness). Thus, we generate one score for each food-hazard pair using the scoring definition in the methodology that have been refined taking into account the peer review comments.

Based on the reviewer's suggestion, we have examined an alternative scenario in which the scoring definitions for Criterion 2 (severity of illnesses) measured by the loss of QALD as

reported by Minor *et al.* (2015). Significantly, the Minor *et al.* 2015 evaluation of severity resulted in much lower ratings to food allergens than those derived by external expert elicitations. The alternative Criterion 2 scenario resulted in a shift to a lower risk score for some of food-hazard pairs, particularly food-undeclared allergen pairs and pairs involving *Brucella* spp., *Cryptosporidium parvum* and other spp., *Salmonella* spp., *Shigella* spp., and *Y. enterocolitica*, *V. parahaemolyticus* and *V. cholerae*.

We agree with the reviewer that *Cronobacter* spp. only causes severe disease in infants. Since it is not known to cause adverse health effects in the general population, we have removed three *Cronobacter*-food pairs that involved foods not typically consumed by infants (e.g., confections without chocolate, and scones). For the remaining pair, e.g., “*Cronobacter*-Infant formula”, we keep score=9 for Criterion 2 for *Cronobacter* spp. as suggested.

Regarding considering a threshold for the presence of undeclared allergens, the definition for Criterion 3 on likelihood of contamination indeed has a similar consideration for chemical hazards, where contamination of a chemical hazard above an action level or allowable level is used. If regulatory guidance on thresholds for food allergens is established in the U.S., we will use the information in evaluating data for the likelihood of contamination for scoring Criterion 3.

While the model does not differentiate among the types of undeclared allergens, relevant allergens for each food are implicitly considered in the evaluation of the criteria. For example, the likelihood of cross-contact for undeclared allergens during manufacturing is considered in expert elicitations for Criterion 5 scoring, as suggested by the reviewers, by including subject matter experts with good knowledge on manufacturing processes for different commodities. Similarly, data on food contamination with undeclared allergens are generally reflective of the most likely undeclared allergens in that food (e.g., RFR and recall data). In instances where expert elicitation was used to score Criterion 3, these experts considered the kinds of information provided by the reviewers, e.g., high likelihood to find a dairy ingredient in a culinary preparation such as prepared ready-to-eat deli salads, or nuts in confectionery such as candy bars with chocolate, and low likelihood of finding shellfish in milk, based on their experience. Regarding the specific food mentioned by the reviewers, we plan to revise the hazard from “undeclared allergens (other than milk)” to “undeclared allergens” for butter, cream, and cheese because “milk” is not in the name of these foods. We also plan to revise the hazard from “undeclared allergens (other than nuts)” to “undeclared allergens” for hummus (made from chickpeas) according to the reviewers’ comment. We do not have a pair for “Butter milk – Undeclared allergens (other than milk)” because undeclared allergens have been represented in other dairy products. Our approach is to include representative food-hazard pairs, without including every possible food-hazard pair (which would not be feasible), noting that the aggregated risk score for a commodity is driven by the higher-ranked food-hazard pairs associated with the commodity.

QUESTION 9

Comments.

The reviewers provided a number of comments and suggestions regarding considerations to take into account when identifying high risk vs. not high risk food-hazard pairs or foods including that the FSMA regulation identifies what should be considered, compare risk scores, consider the risk score distribution and percentage of foods designated, check that the designation should make sense from an evidence point of view, and check for consistency. One reviewer suggested considerations of a HACCP approach, less weight be given to short shelf-life foods, and costs to industry and the nation be included. One reviewer reiterated their concern that hazards in foods that are not reflected strongly in the outbreak record (e.g., new hazards or food-hazard combinations and some known hazards such as *Campylobacter*) may not be adequately considered.

FDA response:

We have taken into consideration the reviewers' comments and suggestions in the way we describe model outputs to inform policy decisions. The seven criteria in the model encompass the contributions from epidemiological evidence (e.g., outbreaks) as well as other risk factors (e.g., known or reasonably foreseeable contamination of the hazard in the food, the food can support the growth of pathogens). We also provide model results in a variety of ways: score for each of the seven criteria, risk score for each of the food-hazard pairs, aggregated risk score for each of the commodities, distributions of food-hazard pair risk scores and commodity aggregated risk scores, a ranked list of food-hazard pairs based on risk score, and a ranked list of commodities based on aggregated risk scores. The HACCP considerations suggested by one reviewer are captured by the approach we use in developing the risk ranking model, in which we first identify a comprehensive list of food-hazard pairs and then evaluate and score each pair using well-defined criteria in the model. Especially in Criterion 5, the scoring definition encompasses whether adequate control measures are available and how well the control measures are implemented in the industry. This definition captures the reviewer's suggestions on identifiable Critical Control Points and their reliability. We thank the reviewers for suggesting the pros and cons with regard to placing more weight or less weight on products with a short shelf life in light of the real time decision during an outbreaks or contamination events. In the model, and shelf life is considered in estimating growth potential and the scoring of Criterion 4. While the impact of some food hazards on public health is not reflected strongly in the outbreak record (Criterion 1), their impact is captured to a large extent in the cost of illness Criterion 7 and their potential impact is captured by Criteria 2-6, for the potential public health impact of all food-hazard pairs.

As noted by one reviewer and consistent with a risk analysis framework, the (risk management) process for this policy decision is functionally separate from the (risk assessment) model development, implementation, and risk ranking results. We appreciate the reviewers' elaborating on the nature of risk and that the risk-ranking model has represented various aspects of the "risk equation". We agree that risk communication to understand the expectations of stakeholders is important. We have taken into consideration comments by stakeholders on a draft approach (issued for comments in a Federal Register notice (79 FR 6596 (Feb. 4, 2014))), and designed the model flexibly so different options can be explored.

In the revised model, we have refined commodity (analogous to "secondary commodity") definitions to better ensure a similar level of granularity across all commodity categories, taking into consideration the characteristics of the food, the hazard and the manufacturing process. We have also re-examined the food-hazard pairs included in the model for consistency, and revised where appropriate, while at the same time, recognizing that it is not feasible to include all potential combination of food-hazard pairs. Our approach is to include representative food-hazard pairs, without including every possible food-hazard pair (which would not be feasible), noting that the aggregated risk score for a commodity is driven by the higher-ranked food-hazard pairs associated with the commodity.

We agree with the reviewers that economic costs of an outbreak are not limited to medical costs. Indeed, although non-public health economic impacts such as potential industry costs and loss of market costs are not included in Criterion 7 of the model, such costs may be considered elsewhere in separate regulatory impact analysis.

QUESTION 10

Comments.

The reviewers generally suggested the model data be updated every 3-5 years, noting that the timing should take into account new information that becomes available that might signal changes in criterion scores such as outbreaks, consumption patterns, introduction of new processes, products or suppliers.

FDA response:

We thank the reviewers for the comments and suggestions. In the revised report, we have incorporated the latest data available, including: for Criterion 1, outbreak data up to mid-2019 for microbial hazards and up to 2017 for marine and plant biotoxins; for Criterion 6, consumption data from NHANES up to 2016 (the latest data cycle available); for other criteria, up to 2019 as available for peer-reviewed literature, technical reports, and expert elicitations QA/QC. We will

also consider the reviewers' suggestions in future updates of the model, considering the data sources and data that might become available in the future.

QUESTION 11

Comments.

Two of the peer reviewers indicated that the report was clear in its description of the risk ranking approach, criteria, data and results, and model limitations and one referred to comments made previously (earlier questions) on places where the report could be more transparent, more fully document sources, and limitations of the model. One reviewer felt the report would benefit from streamlining and significant editing.

FDA response:

We thank the reviewers for their comments. We have revised the Methods document taking into account the reviewer comments. For example, for the comment about general approach on uncertainty, we have streamlined the description of how uncertainty and confidence are scored. Because the evaluation is on availability and quality of the data, we have decided to remove the term "uncertainty" and keep only the term "confidence" to avoid redundancy, because results for the latter correspond inversely to those for the former. Furthermore, we have refined the definitions for confidence evaluation. The evaluation of confidence scores is based on the overall weight of evidence of the available and quality of the data used for scoring the criteria and by extension the food-hazard pairs. In the revised Methods document, we keep the description on calculating confidence scores as a section in chapter 5, rather than its own chapter because the content is relative short.

QUESTION 12

Comments.

One reviewer provided additional comments.

FDA response:

We thank the reviewer for the additional comments. Criterion 1 scores for food-hazard pairs involving undeclared allergens were filled by an allergen expert on the SME panel. We plan to include more experts and seek additional data, if available, for these pairs in the future. With regard to the Criterion 1 score for food-hazard pairs involving microbial hazards and marine and plant biotoxins, the score is based on data (not expert opinions) and it has been updated to reflect the most recent outbreak data for all the food-hazard pairs. For the comment Regarding severity scoring in Criterion 2, we have considered this comment (see for example responses to comment 6-6 and Question 8), and for the reasons described above, we have decided to continue to use

hospitalization and mortality rates as the basis for scoring Criterion 2 in the baseline model, and ran a scenario analysis for scoring Criterion 2 using QALD values (Minor *et al.* 2015) as an alternative data source.

When scoring of Criterion 5, experts take into account both statutory factors iii) and iv), that is, based on their expert knowledge they considered the point in the manufacturing process of food where contamination is most likely to occur, and evaluate the likelihood of contamination and steps taken during the manufacturing process to reduce the possibility of contamination. For example, whether the point of contamination is most likely to occur before or after a kill step is a treated differently. In scoring Criterion 5, a product that has an adequate kill step with processing that significantly minimizes pathogens in a food (“Strong” step), but is subsequently exposed to a post-lethality processing environment with high contamination potential would score C5=3. In comparison, a packaged product receiving an adequate kill step, with the no potential for post-lethality exposure (and thus low contamination potential), would score C5=1. In this way, the sequence of events in the manufacturing environment are considered.

FSMA section 204(d)(2)(A) mandates that the designation of a Food Traceability List must consider specific statutory factors, one of which is “the likely or known severity, including health and economic impacts, of a foodborne illness attributed to a particular food.” This factor, specifically cost of illness, is represented in the model through Criterion 7. As described in our Question 9 response, non-public health economic impacts such as potential industry costs and loss of market costs are not included in Criterion 7 of the model.

We clarify that the cost-per-case estimates from Scharff 2012 were used in an earlier draft (prior to 2015), as described in Appendix C of Appendix K to the Sept 2015 draft report. After the Minor *et al.* 2015 paper was published, the estimates from Minor *et al.* 2015 were used. In the revised report, we keep the reference to Scharff 2012 as one of the studies we evaluated during model development. However, we now clearly indicate that cost-per-case estimates from Minor *et al.* 2015 (with 2018 updates for cost adjustment to reflect inflation) are used for scoring Criterion 7, and that non-public health economic impact (such as industry costs) are not considered in scoring Criterion 7.

SECTIONS III & IV. Specific Observations on Draft Report for Peer Review: *Risk Ranking Model for Product Tracing as Required by Section 204 of FSMA (RRM-PT Draft Report)* within the context of the supporting data.

Comments.

The reviewers provided specific comments and editorial comments on the draft report and appendices.

FDA response:

We have considered each comment in the revision of the Methods document and appendices. For example, in the revised Methods document, we have made further clarification.

- We have clarified that the quantitative scoring definition for Criterion 1 (Methods document, Figure 2-2) is applicable to microbial hazards and marine and plant toxins that cause acute effects; it does not apply to undeclared allergens. We have further clarified that for C1, “the occurrence of illnesses includes outbreak-associated cases **only**” and further clarified when score=0 is assigned. Furthermore, in the description of Criterion 7, we indicate that although in principle, the total number of cases used for C7 scoring includes both outbreak and sporadic cases, few data were available on sporadic cases. Thus, outbreak data were used to calculate the scaled number of cases per year for Criterion 7 scoring using underreporting and underdiagnosis factors.
- Undeclared allergens are chemicals, but they are placed under a different hazard type in our model (and in FDA regulations) because of their unique characteristics, e.g., they are not contaminant *per se*, but rather, a normal component of food that need to be declared, and measures to prevent cross-contact are needed to minimize risk to allergic consumers.
- In the revised Methods document, we have clarified that mycotoxins are of microbial origin. We further indicate that mycotoxins (e.g., aflatoxin M1, fumonisins, Ochratoxin A, and patulin) are included in the model under microbial hazards because microbial growth and control of microbial growth are relevant in addressing mycotoxins.
- We have revised the scoring definition for Criterion 3 according to the reviewers’ suggestions, e.g., scoring bins are now “Low (>0 to ≤0.1%)” and “High (>1%)” – see Table 2-3 in the revised Methods document. Qualitative qualifier is included with quantitative value in the same bin as a reference for expert elicitation score where data (from sampling, RFR and recalls) are not available. We have clarified that in scoring for Criterion 3, data according to an order of preference: prevalence data is used, if available, followed by the average number of RFR reports, the average number of recalls per year, expert elicitation, or lastly, eLEXNET data. Average number/year is correct because data from multiple years are summed and the number/year is calculated for RFR reports and recalls. If the weighted prevalence is 0, it means sampling data are not available. If no prevalence data exist or prevalence data suggests a value of “0”, the scoring moves to using RFR report next, because if RFR report indicates detection of the hazard in the food, then the true prevalence is not zero.
- This geographic weight is to enable the use of sampling and survey studies conducted outside of the U.S., as a proxy to contamination rate for products consumed in the U.S. This is also because a large proportion of the U.S. food supply is imported from around

the world. And this approach was used in a previous interagency risk assessment (US DHHS FDA/USDA FSIS 2003). We did not use outbreak data from outside of the U.S. because an outbreak either occurs or does not occur in the U.S. population and usually are not “imported”. Similar rationale for using only U.S. consumption data for C6.

- Taking into consideration the peer review comments, we have refined the scoring definition for Criterion 4. The name of the criterion is now “Growth potential, with consideration of shelf life” (one indicator) instead of “Growth potential/shelf life” (two indicators). We have revised the scoring definition for Criterion 4, which now includes amount of growth (\log_{10} increase) given customary shelf life, e.g., growth in the range of 1-3 \log_{10} CFU growth over the customary shelf life of the food is an example of Moderate growth potential. The revised definition allows us to appropriately apply data from growth studies and predictive models such as those in ComBase and the USDA PMP program. For food-hazard pairs in which the hazard does not multiply in food (e.g., chemical, allergen, virus and parasite) or the food does not support pathogen growth (e.g., frozen food), C4=0 regardless of shelf life.

REFERENCES

- Batz, M. B., Hoffmann, S., & Morris Jr., J. G. (2012). Ranking the disease burden of 14 pathogens in food sources in the United States using attribution data from outbreak investigations and expert elicitation. *Journal of Food Protection*, 75(7), 1278-1291. doi:10.4315/0362-028X.JFP-11-418
- ComBase Consortium. (2019). ComBase and predictive models. Retrieved from <https://www.combase.cc/index.php/en/>. Accessed: December 10, 2019.
- Christiansson, A., Naidu, A. S., Nilsson, I., Wadström, T., & Pettersson, H. (1989). Toxin production by *Bacillus cereus* dairy isolates in milk at low temperatures. *Applied and Environmental Microbiology*, 55(10), 2595-2600.
- Bernini, V., Bottari, B., Dalzini, E., Sgarbi, E., Lazzi, C., Neviani, E., & Gatti, M. (2013). The presence, genetic diversity and behaviour of *Listeria monocytogenes* in blue-veined cheese rinds during the shelf life. *Food Control*, 34(2), 323-330.
- Beuchat, L. R., Schaffner, D. W., & Versar. (2016). Expert consultation on risk ranking model for product tracing data review and response. Report to FDA, November 22, 2016.
- Dalgaard, P., & Jørgensen, L.V. (1998). Predicted and observed growth of *Listeria monocytogenes* in seafood challenge tests and in naturally contaminated cold-smoked salmon. *International Journal of Food Microbiology*, 40(1-2), 105-115.
- Food and Agriculture Organization of the United Nations and World Health Organization (FAO/WHO). (2004). Risk assessment of *Listeria monocytogenes* in ready to eat foods - Technical report. Microbiological Risk Assessment Series 5. FAO/WHO, Rome. Retrieved from <http://www.fao.org/docrep/010/y5394e/y5394e00.htm> Accessed: August 4, 2020.
- Gavin, A., & Weddig, L. M. (1995). Microbiology of thermally processed foods. In *Canned Foods: Principles of Thermal process Control, Acidification and Container Closure Evaluation* (6th ed., pp. 7-17). Washington, DC: National Food Processors Association.
- Gombas, D. E., Chen, Y., Clavero, R. S., & Scott, V. N. (2003). Survey of *Listeria monocytogenes* in ready-to-eat foods. *Journal of Food Protection*, 66(4), 559-569.
- Goulet, V., Hebert, M., Hedberg, C., Laurent, E., Vaillant, V., De Valk, H., & Desenclos, J.-C. (2012). Incidence of listeriosis and related mortality among groups at risk of acquiring listeriosis. *Clinical Infectious Diseases*, 54(5), 652-660.
- Institute of Food Technologists (IFT). (2001). Chapter II. Potential hazards in cold-smoked fish: *Listeria monocytogenes*. Processing Parameters Needed to Control Pathogens in Cold Smoked Fish - A Report of the Institute of Food Technologists for the Food and Drug Administration. Retrieved from <https://www.fda.gov/files/food/published/Processing-Parameters-Needed-to-Control-Pathogens-in-Cold-Smoked-Fish.pdf>. Accessed: August 4, 2020.

- International Commission on Microbiological Specifications for Foods (ICMSF). (2005). *Microorganisms in Foods 6 Microbial Ecology of Food Commodities*, 2nd ed. Kluwer Academic/Plenum Publishers, New York.
- International Commission on Microbiological Specifications for Foods (ICMSF). (2011). *Microorganisms in Foods 8 Use of Data for Assessing Process Control and Product Acceptance*. Springer, New York.
- Jinneman, K. C., Wekell, M. M., & Eklund, M. W. (2007). Incidence and behavior of *Listeria monocytogenes* in fish and seafood. In: Ryser, E.T., Marth, E.H. (eds.), *Listeria, Listeriosis, and Food Safety*, 3rd ed. CRC Press, Boca Raton.
- Kottapalli, B., Butler, S., Peers, M., & Holzhueter, D. (2017). Application of principles of Failure Modes, Effects, and Criticality Analysis to fluid milk food safety plan. Poster P.111, Society for Risk Analysis (SRA) Annual Meeting, December 10-14, 2017. Arlington, Virginia.
- Lambertz, S. T., Nilsson, C., Bradenmark, A., Sylven, S., Johansson, A., Jansson, L. M., & Lindblad, M. (2012). Prevalence and level of *Listeria monocytogenes* in ready-to-eat foods in Sweden 2010. *International Journal of Food Microbiology*, 160(1), 24-31. doi:10.1016/j.ijfoodmicro.2012.09.010
- Luchansky, J. B., Chen, Y., Porto-Fett, A. C., Pouillot, R., Shoyer, B. A., Johnson-DeRycke, R., Eblen, D. R., Hoelzer, K., Shaw Jr, W. K., Van Doren, J. M., Catlin, M., Lee, J., Tikekar, R., Gallagher, D., Lindsay, J. A., Lm MBS Team, & Dennis, S.(2017). Survey for *Listeria monocytogenes* in and on ready-to-eat foods from retail establishments in the United States (2010 through 2013): Assessing potential changes of pathogen prevalence and levels in a decade. *Journal of Food Protection*, 80:903-921.
- Martinez-Rios, V., & Dalgaard, P. (2018). Prevalence of *Listeria monocytogenes* in European cheeses: A systematic review and meta-analysis. *Food Control*, 84, 205-214.
- Maunula, L., Kaupke, A., Vasickova, P., Soederberg, K., Kozyra, I., Lazic, S., Van de Poel, W. H. M., Bouwknegt, M., Rutjes, S., Willems, K. A., Moloney, R., D'Agostino, M., de Roda Husman, A. M., von Bonsdorff, C.-H., Rzezutka, A., Pavlik, I., Petrovic, T., & Cook, N. (2013). Tracing enteric viruses in the European berry fruit supply chain. *International Journal of Food Microbiology*, 167(2), 177-185.
- Mejlholm, O., & Dalgaard, P. (2009). Development and validation of an extensive growth and growth boundary model for *Listeria monocytogenes* in lightly preserved and ready-to-eat shrimp. *Journal of Food Protection*, 72(10), 2132-2143.
- Mejlholm, O., Gunvig, A., Borggaard, C., Blom-Hanssen, J., Mellefont, L., Ross, T., Leroi, F., Else, T., Visser, D., & Dalgaard, P. (2010). Predicting growth rates and growth boundary of *Listeria monocytogenes*—An international validation study with focus on processed and ready-to-eat meat and seafood. *International Journal of Food Microbiology*, 141(3), 137-150.

- Minor, T., Lasher, A., Klontz, K., Brown, B., Nardinelli, C., & Zorn, D. (2015). The per case and total annual costs of foodborne illness in the United States. *Risk Analysis*, 35(6), 1125-1139.
- Montville, T. J., & Matthews, K. R. (2013). Physiology, growth, and inhibition of microbes in foods. In M. P. Doyle, & R. L. Buchanan (eds.), *Food Microbiology: Fundamentals and Frontiers* (pp. 3-18). ASM Press, Washington DC.
- Notermans, S., Dufrenne, J., Teunis, P., Beumer, R., Te Giffel, M., & Weem, P. P. (1997). A risk assessment study of *Bacillus cereus* present in pasteurized milk. *Food Microbiology*, 14(2), 143-151.
- Pires, S. M. (2013). Assessing the applicability of currently available methods for attributing foodborne disease to sources, including food and food commodities. *Foodborne Pathogens and Disease*, 10(3), 206-213.
- Pires, S. M., Evers, E. G., van Pelt, W., Ayers, T., Scallan, E., Angulo, F. J., Havelaar, A., Hald, T., & Workgroup (2009). Attributing the human disease burden of foodborne infections to specific sources. *Foodborne Pathogens and Disease*, 6(4), 417-424.
- Posada-Izquierdo, G. D., Pérez-Rodríguez, F., López-Gálvez, F., Allende, A., Selma, M. V., Gil, M. I., & Zurera, G. (2013). Modelling growth of *Escherichia coli* O157:H7 in fresh-cut lettuce submitted to commercial process conditions: Chlorine washing and modified atmosphere packaging. *Food Microbiology*, 33(2), 131-138.
- Posada-Izquierdo, G. D., Pérez-Rodríguez, F., López-Gálvez, F., Allende, A., Gil, M. I., & Zurera, G. (2014). Modeling growth of *Escherichia coli* O157:H7 in fresh-cut lettuce treated with neutral electrolyzed water and under modified atmosphere packaging. *International Journal of Food Microbiology*, 177, 1-8.
- Pouillot, R., Hoelzer, K., Chen, Y., & Dennis, S. B. (2015). *Listeria monocytogenes* dose response revisited—incorporating adjustments for variability in strain virulence and host susceptibility. *Risk Analysis*, 35(1), 90-108.
- Pouillot, R., Lubran, M. B., Cates, S. C., & Dennis, S. (2010). Estimating parametric distributions of storage time and temperature of ready-to-eat foods for U.S. households. *Journal of Food Protection*, 73(2), 312-321.
- Pouillot, R., Goulet, V., Delignette-Muller, M. L., Mahé, A., & Cornu, M. (2009). Quantitative risk assessment of *Listeria monocytogenes* in French cold-smoked salmon: II. Risk characterization. *Risk Analysis*, 29(6), 806-819.
- Pouillot, R., Miconnet, N., Afchain, A. L., Delignette-Muller, M. L., Beaufort, A., Rosso, L., Denis, J.-B., & Cornu, M. (2007). Quantitative risk assessment of *Listeria monocytogenes* in French cold-smoked salmon: I. Quantitative exposure assessment. *Risk Analysis*, 27(3), 683-700.
- Ryser, E.T., 2007. Incidence and behavior of *Listeria monocytogenes* in cheese and other fermented dairy products. In: Ryser, E.T., Marth, E.H. (eds.), *Listeria*, Listeriosis, and Food Safety, 3rd ed. CRC Press, Boca Raton.

- Scallan, E., Hoekstra, R., Mahon, B., Jones, T., & Griffin, P. (2015). An assessment of the human health impact of seven leading foodborne pathogens in the United States using disability adjusted life years. *Epidemiology & Infection*, *143*(13), 2795-2804.
- Scallan, E., Hoekstra, R., Angulo, F., Tauxe, R., Widdowson, M., Roy, S., Jones, J., & Griffin, P. (2011). Foodborne illness acquired in the United States - major pathogens. *Emerging Infectious Diseases*, *17*(1), 7-15. doi:DOI 10.3201/eid1701.P11101
- Scharff, R. L. (2012). Economic burden from health losses due to foodborne illness in the United States. *Journal of Food Protection*, *75*(1), 123-131.
- Schvartzman, M., Maffre, A., Tenenhaus-Aziza, F., Sanaa, M., Butler, F., & Jordan, K. (2011). Modelling the fate of *Listeria monocytogenes* during manufacture and ripening of smeared cheese made with pasteurised or raw milk. *International Journal of Food Microbiology*, *145*, S31-S38.
- te Giffel, M. C., Beumer, R. R., Leijendekkers, S., & Rombouts, F. M. (1996). Incidence of *Bacillus cereus* and *Bacillus subtilis* in foods in the Netherlands. *Food Microbiology*, *13*(1), 53-58. doi:10.1006/fmic.1996.0007
- Tiwari, U., Walsh, D., Rivas, L., Jordan, K., & Duffy, G. (2014). Modelling the interaction of storage temperature, pH, and water activity on the growth behaviour of *Listeria monocytogenes* in raw and pasteurised semi-soft rind washed milk cheese during storage following ripening. *Food Control*, *42*, 248-256.
- Wemmenhove, E., van Valenberg, H. J., Zwietering, M. H., van Hooijdonk, T. C., & Wells-Bennik, M. H. (2016). Minimal inhibitory concentrations of undissociated lactic, acetic, citric and propionic acid for *Listeria monocytogenes* under conditions relevant to cheese. *Food Microbiology*, *58*, 63-67.
- U.S. Department of Agriculture Agricultural Research Service (USDA ARS). (2019). Pathogen Modeling Program (PMP) online. Retrieved from <https://pmp.errc.ars.usda.gov/PMPOnline.aspx>. Accessed: December 4, 2019.
- U.S. Department of Health and Human Services Food and Drug Administration and U.S. Department of Agriculture Food Safety and Inspection Service (US DHHS FDA/USDA FSIS). 2003. Quantitative assessment of relative risk to public health from foodborne *Listeria monocytogenes* among selected categories of ready-to-eat foods Available at: <http://www.fda.gov/Food/FoodScienceResearch/RiskSafetyAssessment/ucm183966.htm>. Accessed October 22, 2015.
- U.S. Food and Drug Administration. (2016a). Arsenic in rice and rice products risk assessment report. Retrieved from <https://www.fda.gov/food/cfsan-risk-safety-assessments/arsenic-rice-and-rice-products-risk-assessment>. Accessed: December 16, 2019.
- U.S. Food and Drug Administration. (2016b). FDA Investigates outbreak of Hepatitis A illnesses linked to frozen strawberries. Retrieved from <https://www.fda.gov/food/outbreaks-foodborne-illness/fda-investigates-outbreak-hepatitis-illnesses-linked-frozen-strawberries>. Accessed: August 4, 2020.
- U.S. Food and Drug Administration. (2017). Draft guidance for industry: Control of *Listeria monocytogenes* in ready-to-eat foods. Retrieved from <https://www.fda.gov/regulatory->

- [information/search-fda-guidance-documents/draft-guidance-industry-control-listeria-monocytogenes-ready-eat-foods](#). Accessed: December 16, 2019.
- U.S. Food and Drug Administration. (2019). *Fish and Fishery Products Hazards and Controls Guidance*. 4th ed. Retrieved from <https://www.fda.gov/food/seafood-guidance-documents-regulatory-information/fish-and-fishery-products-hazards-and-controls> Accessed: December 20, 2019.
- U.S. Food and Drug Administration. (2020). Microbiological surveillance sampling: FY 19-20 frozen berries (strawberries, raspberries and blackberries). Retrieved from <https://www.fda.gov/food/sampling-protect-food-supply/microbiological-surveillance-sampling-fy-19-20-frozen-berries-strawberries-raspberries-and#results>. Accessed: August 4, 2020.
- Zwietering, M., De Wit, J., & Notermans, S. (1996). Application of predictive microbiology to estimate the number of *Bacillus cereus* in pasteurised milk at the point of consumption. *International Journal of Food Microbiology*, 30(1), 55-70.