FSIS Response to Peer Review Comments on: *Quantitative Risk Assessment for Salmonella in Raw Chicken and Raw Chicken Products*

Prepared by:

Food Safety and Inspection Service
United States Department of Agriculture
July 2024

USDA Food Safety and Inspection Service

Contents

1. Introduction

The Food Safety and Inspection Service (FSIS) is the public health agency in the U.S. Department of Agriculture (USDA) responsible for ensuring the safety of the nation's commercial supply of meat, poultry, and egg products. FSIS ensures food safety through the authorities of the Federal Meat Inspection Act, the Poultry Products Inspection Act, and the Egg Products Inspection Act, as well as humane animal handling through the Humane Methods of Slaughter Act. FSIS consists of about 9,000 employees, most of whom work on the frontlines in establishments across the country to ensure the production of food is safe.

Despite FSIS sampling data showing reductions in *Salmonella* contamination in poultry products, the Agency's current approach to *Salmonella* has not led to a demonstrable reduction in *Salmonella* infections. To address these issues, the FSIS Office of Food Safety (OFS) developed a new *Salmonella* Initiative, which is a high priority, multipronged approach to reduce *Salmonella* foodborne illnesses from FSIS-regulated products. One piece of this Initiative is the quantitative risk assessments for *Salmonella* in chicken conducted by the Risk Assessment and Analytics Staff (RAAS) within the FSIS Office of Public Health Science (OPHS). RAAS analysts have extensive experience conducting risk assessments to evaluate intervention strategies to reduce foodborne risks and to guide, support, and enhance the Agency's overall decision-making process and risk management policies.

In a manner consistent with the current Office of Management and Budget (OMB) Peer Review Guidelines (Final Information Quality Bulletin for Peer Review, December 15, 2004), FSIS contracted RTI International to conduct an independent and formal peer review of the quantitative risk assessment for *Salmonella* in chicken. This report summarizes the process RTI used to identify and recruit the five scientific experts who conducted the peer review and includes their responses to the charge questions provided by FSIS. Their biographies are also included.

The peer reviewer comments were prepared for FSIS by:

Dr. Juliana Ruzante
Dr. Donal Bisanzio
RTI International
3040 E. Cornwallis Road
Research Triangle Park, NC 27709

RTI Project Number 0216627.003

2. Peer Review Charge Questions

The selected peer reviewers were asked to address the following questions while conducting their review:

1. **Please evaluate the available data and the underlying assumptions used in this risk assessment.**
    a. To your knowledge, have all key studies and data been identified, correctly analyzed, and properly interpreted? If not, please provide additional data sources and citations (where appropriate) or alternative interpretations or analyses.
    b. Have the strengths and limitations of the data been transparently explained?
    c. Given the differences in the data and sampling methodologies, is the overall modeling approach used for chicken products (carcasses, parts, and comminuted) appropriate?
    d. Are the differences in data for the three chicken products adequately described and addressed?
    e. Are the strengths and limitations of the quantitative PCR data method used to generate data for carcass contamination transparently explained and adequately assessed.

2. **Please identify limitations, weaknesses, or inadequacies of the bioinformatics serotype clustering; please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:**
    a. Was the *Salmonella* genomics data appropriately curated and processed?
    b. Are the databases and methods used to determine virulence factors appropriate? Should any other virulence factors have been considered?
    c. Is the clustering algorithm accurately described, utilized, and appropriate for its intended use?

3. **Please evaluate the two-curve dose-response model used to estimate the probability of illness for a given exposure dose of *Salmonella*, giving specific consideration to the following:**
    a. Was the use and modification of the Teunis beta-Poisson model appropriate to describe probability of illness due to *Salmonella* serotypes that differ in virulence? If not, what other models should be considered? Please provide the reference(s) if applicable.
    b. What (if any) other data sources and methods should have been used in the *Salmonella* dose-response model risk multipliers? If not, what other data sources and/or methods should be used? Please provide the reference(s) if applicable.

c. Is the use of the two-curve dose-response model appropriately used to estimate illness estimates? If not, what other approach could have been used with this dose-response model? Please provide the reference(s) if applicable.

4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:

a. Is the scenario analysis technique accurately described, utilized, and appropriate for its intended use (i.e., evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products)?

b. Are the data analyses appropriate and R language source code accurate for the aims of the study?

c. The definition of product lots is based on the sampling frequency of the data. Are the methods used to describe the contamination of those lot from samples appropriate, and if not, what other approach should have been taken?

d. Is the assumption that multiple serotypes are present within flocks appropriate and how else can the mixture of serotypes (i.e., “serotype scheme”) be described?

e. Were any considerations missing from the development of the attenuation multiplier to adequately describe *Salmonella* growth and die-off after raw chicken product leaves processing?

f. Does the Monte Carlo simulation approach adequately model the scenarios?

g. What approach could be taken to assess uncertainty in these conclusions?

h. Are the conclusions drawn from the analysis appropriate?

5. Please identify limitations, weaknesses, or inadequacies of the process control modeling techniques and data. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:

a. Is the correlation between indicator organism and *Salmonella* fully described and well characterized?

b. Is the process control modeling technique accurately described, utilized, and appropriate for its intended use?

c. Are the data analyses and source code accurate?

d. Is the risk management question adequately answered, given the limitations in assessing public health impacts from indicator organism policies?

6. **Evaluate whether the documentation of the data and modeling, and interpretation of the results is appropriate. If not, please provide an alternative outline, approach, and/or suggested language for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:**

   a. Is the report clearly written and complete?

   b. Does the report follow a logical structure and layout?

   c. Are the conclusions supported by the risk assessment?

   d. Is the documentation of the assumptions clear and complete?

   e. Is the documented dose-response, exposure assessment, and risk characterization modelling transparent and reproducible?

3. Selection of Peer Reviewers

RTI identified 16 potential peer reviewers with overlapping and complementary expertise in the following topic areas:

- Quantitative microbial risk assessment (e.g., Bayesian modeling, Monte Carlo)
- R coding
- Dose response modeling
- Bioinformatics: Machine learning methods for genomic data (e.g., random forest modeling)
- Knowledge of current laboratory methods for enumerating (e.g., qPCR, characterizing *Salmonella* with statistical analysis of test results [e.g., variability])
- Epidemiology and surveillance of salmonellosis
- Knowledge of chicken production and/or slaughter processes
- Knowledge of turkey production and/or slaughter processes

Since RTI was also conducting the peer review for the quantitative risk assessment for *Salmonella* in turkey (RTI Project 0216627.002), in conjunction with FSIS, we decided that given the overlapping expertise needed and the similarities between the two QMRA models, it would be appropriate to recruit four out of the five peer reviewers to evaluate both models.

We then contacted 11 reviewers to determine their availability and interest in participating. They were all asked to provide an up-to-date curriculum vitae and to complete a form ranking their expertise and identifying potential conflicts of interest (see form in **Appendix A**). This step ensured that we recruited reviewers with the appropriate scientific stature and experience with related projects who were also independent from FSIS.

Ten of the reviewers were available during the designated period. As specified in the proposal, RTI prepared a summary table for the 10 experts and identified 5[1] based on their CV, self-reported expertise, and conflict of interest information (see summary table in **Appendix B**). RTI met with FSIS on January 23, 2023, via Zoom to discuss the selection. The Agency agreed with the proposed selection. No names, affiliations, or biographies were provided or discussed with the Agency to ensure the blinded process.

All selected reviewers signed a nondisclosure agreement as part of establishing a consulting contract. RTI provided experts with all material provided by FSIS. That included the quantitative risk assessment document to be reviewed, the charge questions, a template for peer reviewers to use to submit their answers, a CSV file with the raw data used, and a zip file with the code used in the QMRA. We also provided a document providing an overview of each file (see **Appendix C**).

[1] Four experts were also recommended for the peer review of the risk assessment for *Salmonella* in turkey.

Peer reviewers had 3 weeks to complete their reviews using the template provided by RTI. Email reminders were sent each week and our team answered any clarifying questions as needed during the review period.

Upon receiving each review, Dr. Donal Bisanzio, research epidemiologist and modeler, and Dr. Juliana Ruzante, senior food safety and public health scientist, reviewed each report for quality and completeness and communicated as needed with the reviewers to address any gaps or ambiguities in the reviews.[2]

[2] RTI reviewed all answers with the exception of one, that was only submitted to RTI on March 15, 2023.

4. Selected Peer Reviewer's Biographies

The following peer reviewers were selected to address the charge questions provided by FSIS. Experts had overlapping and complementary expertise in the areas identified as relevant by FSIS.

Maarten Nauta is a Senior Scientist at Statens Serum Institut in Denmark and worked at the National Institute for Public Health and the Environment (RIVM) and at the National Food Institute of the Technical University of Denmark (DTU) in the Netherlands where he specialized in quantitative microbiological risk assessment, the development of methods for "farm-to-fork" risk assessments, and risk-benefit assessments. He is a mathematical biologist with a PhD in Evolutionary Genetics. Dr. Nauta has taught microbiological risk assessment and risk-benefit assessment of foods to students and food safety professionals worldwide. He has published more than 100 scientific publications in international peer-reviewed journals on genetics, evolutionary biology, mathematical modelling, statistics, risk analysis, engineering, food microbiology, veterinary science, epidemiology, pharmacy, nutrition, and toxicology. Dr. Nauta also been part of several national and international committees organized by FAO/WHO, EFSA and ILSI. He is currently an Associate Editor of the journal *Microbial Risk Analysis* (Co-Editor in Chief since 2020), member of the International Committee of Predictive Modelling in Foods, and member of the EFSA BIOHAZ panel.

Gregory M. Paoli is the Principal Risk Scientist at Risk Sciences International and has a degree in Electrical and Computer Engineering. He has been providing consulting services in the field of quantitative risk assessment applied to human health, public safety and the environment since 1993. He specializes in formal probabilistic risk assessment methods, the development of risk-based decision-support tools, comparative risk assessment, and risk communication. He has experience in food safety, animal health, plant protection, climate change impacts on dams, medical and engineering devices, consumer products, and chemicals management and transportation, including hazardous materials. Greg has served on many expert committees devoted to the risk sciences and is a member of the U.S. National Research Council Committee that issued the 2009 report, *Science and Decisions: Advancing Risk Assessment*, and was invited as an expert reviewer of the U.S. EPA's Framework for Human Health Risk Assessment to Inform Decision Making. He has served on committees for the Canadian Standards Association, National Roundtable on the Environment and the Economy, U.S. NRC Standing Committee, and World Health Organization. Additionally, he has worked with the World Health Organization and the Food and Agriculture Organization of the United Nations since 2003.

Abani K. Pradhan is a Professor in the Department of Nutrition and Food Science & the Center for Food Safety and Security Systems at the University of Maryland in College Park. His research focuses on the area of food safety and risk assessment, including *Salmonella*. He has been working on developing and utilizing appropriate methods and approaches to integrate microbial genomics with risk assessment as well as advanced data analytics such as artificial intelligence and machine learning techniques to evaluate public health risk. Dr. Pradhan has

published several book chapters in food safety and risk assessment, and has over 90 peer-reviewed publications and more than 2,100 citations.

Nitya Singh is an Assistant Scientist at the Department of Animal Sciences and Emerging Pathogen Institute of the University of Florida. She is a bioinformatician and biostatistician and has expanded her skills to public health, infectious disease modeling, and biomedical informatics. She has a PhD in Information Technology and a master's in Biomedical Sciences. Dr. Singh's current research focuses on molecular epidemiology, phylodynamics, meta/genomics, machine learning, and statistical data analysis to support tracking molecular links for possible outbreaks/illnesses, food safety, and women empowerment. She has experience in R coding, handling large datasets, and solving complex coding problems.

Bing Wang is an Associate Professor at the Institute of Agriculture and Natural Resources at University of Nebraska-Lincoln. She has a PhD in Veterinary Microbiology with a minor in Statistics. Dr. Wang is human health risk analyst specializing in addressing microbial food safety issues. Dr. Wang's research aims to improve public health decision making through data analysis and decision tools, particularly the use of epidemiology, systematic review, meta-analysis, and quantitative microbial risk assessment to optimize the food production and processing conditions and enhance the effectiveness of food safety and quality resources. Dr. Wang has published several book chapters and peer-reviewed papers in food safety and risk assessment. She participated in several Joint FAO/WHO Expert Consultation Meetings on Microbiological Risk Assessment (JEMRA) including a recent one convened on behalf of the Codex to develop guidance to control *Salmonella* and *Campylobacter* spp. in poultry meat and another to update the FAO/WHO guideline to provide practical guidance and a structured framework for carrying out risk assessment of microbiological hazards in foods at a national and international level.

5. Individual Reviewer Comments to FSIS's Charge Questions

Here are the unedited and non-summarized comments from peer reviewers to FSIS. Other than the addition of a column with FSIS responses, the only edit made to the responses as provided by RTI was to merge multiple adjacent comment rows when a single response was warranted.

Chapters, sections, figures and tables mentioned in FSIS responses all refer to the main document this report accompanies: FSIS' *Quantitative Risk Assessment for Salmonella in Raw Chicken and Raw Chicken Products*.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
General Comments					
1	28-29	712-716	A	An uncertainty analysis is basically missing. This is a severe shortcoming of the risk assessment, as there are many sources of uncertainty and the risk managers should have a clue about the impact of these uncertainties on the conclusions of the risk assessment for informed decision making. I do understand that it is challenging to include all uncertainties in the modelling and in the interpretation of the results, but the fact that it is challenging for the risk assessors to characterize the uncertainty means that it is almost impossible for the risk managers to do so. They would need more guidance on this than just a statement that you will get back to it later. See for example https://www.efsa.europa.eu/en/efsajournal/pub/6090	An uncertainty analysis has been added to **Chapter 5 Final Product Standards** to best inform decision making.
2			A	In terms of available data, it is shame that only one sample is collected per flock at rehang and post chill. The assumption that all flocks are contaminated with multiple strains seems to be a very important one, definitely if you then study the impact of diverting flocks based on the types of strains found in a single sample. You can do fancy analyses with the dose response and bioinformatics, but it seems the main shortcoming of the proposed methods is that some essential data in terms of variability in strains contaminating the carcasses and the meat is lacking. Clearly, the risk assessors are not to blame for this lack of data,	This data limitation has been summarized in the added to **Table 9: Risk assessment information and assumptions** in **S1.6 Introductory Tables and Figures**.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				but it would be worthwhile to stress that it is important.	
3			B	It's great that Risk Assessment has been performed by keeping mechanistic aspects as per Codex guidelines, covering all 4 areas, including 1. Hazard identification, of pathogenic *Salmonella* in poultry chicken (based on FSIS' Risk Profile, not disclosed here and no other latest reference mentioned in its absence), 2. Exposure assessment, as estimation of *Salmonella* contamination both at processing and final product level. Utilization of virulence factor genomic information-based clustering approach to identify more virulent pathogenic Serotypes is innovative and useful to emphasize the higher risk from the exposure of more virulent *Salmonella*, helpful in optimization of risk management and impact on public health and illness. 3. Hazard identification, using Beta-Poisson dose-response relationship, in this case, looks most appropriate and consideration of virulence provides additional supportive assessment for risk management. 4. Risk characterization, baseline risk estimates, and calibration to the latest *Salmonella* cases attributed to chicken provide a good basis for assessment of risk management of chicken products on public health. Along with these 4 aspects of QMRA, work has taken care of process control monitoring and has provided the details of modeling, along with the limitations of the adopted approach.	Summary of the scenarios run with product types has been added to the conceptual model. A table of the scenarios that were successfully modeled has been added to **section 1.6: Introductory Tables and Figures**. FSIS' *Salmonella* in Poultry Risk Profile has been externally peer reviewed and is available for view here.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				Overall, the approach is good and follows standard QMRA practices. The conceptual model is also adequate, brief names of scenarios mentioning the product types, for which the final assessment was performed successfully, should be considered to be included in Figure 1.	
4			B	A novel dose-response model to focus on highly virulent serotypes dose-response and the impact of their management on public health is efficient. This model will also serve as a useful method for future proposed more testing and data procurement plans to implement efficient risk management policies for all product type.	No response is necessary.
			C	(No comment from this reviewer)	No response is necessary.
5			D	Overall, given the scope of the risk assessment, data, assumptions, and analyses are reasonable and appropriate. However, some assumptions must be explained further and some missing references must be provided for some statements. Please see below for specific comments and details in Q1 a-e.	Responses provided in the responses to this reviewers comments to Q1 a-e.
6			E	The risk assessment, of necessity and therefore appropriately, relies on a large number of assumptions, some very clearly described and articulated while others are very quickly and minimally described. Some more 'evenness' in the treatment of assumptions (and possibly a summary table, similar to that in the dose-response appendix), would be an important contribution to	A summary table of assumptions has been added to section **1.6: Introductory Tables and Figures**.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				the document such that reviewers, the public and the ultimate risk managers have a sufficiently clear understanding of the foundations and limitations.	
7	34	898	E	"To describe growth and die-off of *Salmonella* in contaminated product lots as product travels from the end of processing, through commerce and preparation and consumption, an attenuation multiplier is used. The full derivation of this multiplier is described in Appendix C and an illustration of its utility has been shown in previous work (E. Ebel & Williams, 2015)." (lines 898-901)	
	71	16271-1630		"An attenuation distribution that encompassed all the effects of partitioning, mixing, growth, and attenuation that can occur between production of raw chicken and consumption of chicken servings was also defined (Log10Normal(-5,1.91)" (E. Ebel & Williams, 2015)). (lines 1627-1630)	
	144	382-386		"FSIS provided our team with a lognormal (LN) distribution (base 10)7 LN(-3.037117, 1.279985) representing the distribution of *Salmonella* in raw poultry, and an attenuation distribution LN(-5.00, 378 1.91)[29]. This distribution adjusts the initial dose distribution by the combined effect of cooking, mixing, partitioning, cross-contamination, growth and consumption while considering variability in cooking practices." (pg. 144, lines 382-386). One of the key assumptions in the model is that the three types of products (carcasses, parts, comminuted) experience the same level of "attenuation" downstream of the production process	We have augmented the explanation about the attenuation distribution in section **1.5 Model Approach** as follows: "We summarize the effects of the myriad of pathways contaminated product may follow from the end of processing through commerce and preparation using an attenuation distribution. This attenuation distribution captures the variability associated with mixing, partitioning, growth, cooking and serving size

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				(everything from post-rehang to the consumption of a serving). This assumption, while very convenient to the algorithm, seems to be very difficult (impossible?) to justify. Given that the three product types differ substantially in the distribution of their initial contamination levels (for reasons which are explained well and described as "consistent with what would be biologically expected" differences, line 1104), it would seem that the same concept of "biologically expected differences", would be applicable to the very different sets of downstream processes associated with these products.	processes between production and consumption. A lognormal attenuation distribution ($\mu = -5.00log10, \sigma = 1.91log10$) was calibrated previously for chicken using a single distribution for *Salmonella* contamination, a general *Salmonella* dose-response function from WHO-FAO (FAO/WHO, 2000), and a prior estimate for total *Salmonella* illnesses attributed to chicken (Ebel, 2015). The log10 mean of this distribution (-5log10) is consistent with a scenario where a raw chicken product is properly handled to avoid growth of *Salmonella*, then subjected to cooking to achieve a minimum internal temperature of 165 ºF, which the Agency recommends as the final cooking temperature and has determined will deliver at least a 7log10 reduction of *Salmonella (FSIS, 2017)*, and consumed in a serving size of 100 g (adding 2log10 to the 7log 10 reduction). Lacking alternative estimates, this default attenuation distribution is used across analyses of carcasses, parts,

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				As the authors are acutely aware from their past experience constructing these models, there are a substantial number of factors that would be included in a typical process model between rehang and consumption. For a number of those process model steps, very different assumptions and associated mathematical models would likely be expected across the three product types. From a microbiological perspective, differences in the extent of further mixing, partitioning, cross-contamination potential, water activity and other growth-relevant parameters, fat and other thermally relevant content, the location of the actual bacterium on the surface versus embedded within the product with important impacts on both growth, thermal inactivation and potential for cross contamination, the partitioning of the products into fresh versus frozen streams, the propensity for the cooking and consumption to be at home versus restaurant, and (as FSIS concedes within this document though this may be the least important among these differences), serving sizes. This list could be readily augmented. Given the number of potentially important biological differences, it would seem that their being equal is highly improbable. The decision to apply an equal attenuation factor has important implications for the relative impact of any risk management actions that act on individual products, particularly when they differ so substantially in initial contamination levels, and may very well differ in the same risk-directional sense	comminuted chicken and for *Salmonella* serotypes. Similarly, this attenuation distribution is used in the derivation of the dose-response models for different virulence-based *Salmonella* serotype clusters (**Appendix A**). Nevertheless, FSIS also conducted a sensitivity analysis to explore the effects of alternative attenuation distribution parameters on the estimated effectiveness of risk management options. " The description of differences in initial contamination distributions is not directly relevant to what happens to the product between its production and its consumption. Our default assumption of the same attenuation distribution for all chicken products is based on 1) a common target internal cooking temperature recommendation for any chicken product (the log10 reduction average should be similar for any product), 2) the serving size is similar across products, 3) the default attenuation distribution was calibrated to

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				(e.g. higher levels of contamination with lower levels of attenuation multipliers). The fact that FSIS cannot currently reliably estimate the differences in the overall attenuation multipliers may be a reasonable reflection of the state of the art and available evidence to characterize downstream processes, but that does not justify an assumption of "sameness" when the opposite is more justified.	all chicken illnesses, and 4) alternatives were not readily available.
				However, presenting the relative risk of these products on a per-serving basis, and then combining with the number of servings to generate overall product stream risk and total risk would be much more transparent. This type of analysis could be done with varying dose-response assumptions to simply make the differences in relative risk apparent. This would also allow simple sensitivity analysis with respect to "what if" the attenuation multipliers were different. The current use of a "composite" initial contaminant distribution (and the provision of only this distribution to the dose-response team) has the potential to obfuscate the reality of these three distinct product streams and their contribution to risk, both on a per-serving basis as well as their overall contribution to risk, and the relative merits of risk management actions aimed at the different product streams.	We agree that the estimated risk of illness per serving is directly influenced by the attenuation distribution parameters. Nevertheless, as we have now shown in our sensitivity analysis, the final model predictions regarding the effectiveness of finished product

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
					standards is not highly sensitive to the attenuation distribution. Ultimately, as is implied by the assignment of illnesses per product (i.e., the lambda ill parameter, that is independent of the attenuation distribution) the risk assessment was not directly developed to estimate the risk per serving for each product type, but to estimate the effect of removing lots that exceed some threshold level at production. Our sensitivity analysis suggests that the change in the average concentration following implementation of finished product standards is reasonably predictive of this effectiveness, regardless of the specific attenuation distribution or dose-response function.
8	34 71 144 18	898 1627-1630 382-386 453	E	The actual strict definition of what is included with the concept of the "attenuation distribution" is ambiguous. This is, at least in part, due to the different ways that it is described in the three passages included above (specifically whether serving size is part of the attenuation multiplier). The ambiguity is also present in the Overall conceptual model diagram (p. 18) in which the "Consumption" icon resides outside of the attenuation multiplier. While this makes sense for Number of Servings, this reviewer infers that the serving size is embedded within the attenuation multiplier.	The overall conceptual model has been improved to remove this ambiguity. The recommended example calculations have been included in **section 1.5 Model Approach** where additional detail on the attenuation distribution have been included.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				Some effort was required to reach the conclusion that the attenuation multiplier would have units of "grams per serving" (and its units are nowhere given that this reviewer could find). These units are assumed since when applying the multiplier to the initial contamination level expressed in cfu/g, the result is a consumed dose (i.e. cfu/serving). An example calculation using scalar-valued initial concentration, attenuation multiplier and a final average dose would be very instructive for the reader. For example, if the process were as simple as an overall 7-log reduction in the process, combined with a 100g serving (2 log grams per serving) would yield the median of the FSIS attenuation distribution (-7 + 2 = -5). The -7 has units of "surviving cfu per initial cfu" (i.e., unitless, as in microbial inactivation), and the +2 would have units of grams per serving.	
9	34	900-901	E	"The full derivation of this multiplier is described in Appendix C and an illustration of its utility has been shown in previous work (E. Ebel & Williams, 2015). …" This is not the case. The appendix describes the idea that the multiplier could reasonably be asymptotically lognormal based on CLT and CLT-like arguments. An example of a simple process model with growth and inactivation steps and a few of the expected scalar adjustments resulting in near-lognormal behavior would be more convincing to augment the theoretical basis. The application of the multiplier (i.e., the ability to readily estimate the	The attenuation distribution derived in Ebel and William (2015) was calibrated such that when it was multiplied with an assumed initial distribution of contamination (estimated from a 2007-2008 baseline survey of chicken) predicted a dose distribution that, when the WHO-FAO dose-response model was integrated across it and multiplied by an estimated annual number of chicken servings consumed, equaled an estimated number of chicken *Salmonella* illnesses. None of the inputs from that

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				first two moments of the resulting lognormal distribution when multiplying two lognormal distributions) is also described. However, the multiplier itself (mean of -5.00, sigma of 1.91) is not derived in Appendix C. It is also not fully described in the referenced Ebel and Williams, 2015. Some insight is provided in that publication, but not enough (and many readers would not be likely to hunt it down). “Using a root-finding algorithm in the R programming language (31), we assumed the values in Table 1 to calculate the lognormal parameters for λatten.” And note (b) to Table 1 in Ebel and Williams, 2025, “These parameters are derived by solving equation 7 such that the probability of illness per serving in that equation (after adjusting for attribution, underdiagnosis, and FoodNet catchment area) exactly equals the reported number of FoodNet illnesses for each pathogen." Given the two parameters, and a single equation one might assume that the root-finding algorithm was assigned to estimate the sigma parameter, with the mean parameter assigned the assumed value of exactly -5.00, but this is not stated and is pure supposition on the part of this reviewer. The point of the comment is to remove such guesswork on the part of the reader. Given the importance of this distribution as the dominant source of variability in the dose	analysis were replicated in the present analysis. Therefore, the circularity description seems less apt. Instead, the WHO-FAO dose-response model determined the attenuation distribution (for a number of illnesses estimated in 2015) while that attenuation distribution is used to derive 2 dose-response models in the present analysis without any calibration to a number of illnesses (attenuation is a simple input to the present analysis). Despite the absence of any calibration with estimated illnesses per year in the present study, the estimated baseline probability of illness per serving for carcasses, parts and comminuted product is about 3 per million, 5 per million and 2.5 per 100,000, respectively. For carcasses and parts, these estimates are very similar to our empirical estimate of about 2 per million. Nevertheless, we agree with the comment that the ""risk” (*Salmonella* cases per year) is assumed in the model". As mentioned above, the lamba ill parameter that describes the expected number of illnesses per year for

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				distribution (sigma of 1.91, leading to a final sigma value of 2.3 for the average dose distribution), this should be more fully explained as central to this assessment, and as input into the dose-response relationships derived.	each product is an input to the model rather than an output.
10			E	The application of the LN(-5,1.91) for the attenuation multiplier presents the issue of a potentially circular argument. These parameters are derived from the application of the FAO/WHO dose-response model (Table 1 of Ebel and Williams). Then, the parameters become input into the derivation of the resulting dose-response curves that form a key output (rather than an input) of this FSIS document. In a sense the "risk" (*Salmonella* cases per year) is assumed in the model, and dose-response curves become the output of the overall assessment (in addition to risk reductions, once this dose-response assessment has been assumed). The circularity is in the form of "dose-response assumption determines attenuation distribution" and then "attenuation distribution determines dose-response assumption."	

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				If this is not the case, for reasons this review may not fully understand, it should be clearly explained.	
11			E	Figure 1 should more clearly indicate (particularly through the direction of arrows) that the number of cases is an input rather than an output of the overall calculation process (though estimates of risk reductions could be considered an output). FSIS has chosen a fit-for-purpose model and characterized it as a "calibrated" model, but that fact is not adequately portrayed as the flow of evidence in Figure 1.	The figure has been updated to portray the fit-for-purpose model more accurately.
a. To your knowledge, have all key studies and data been identified, correctly analyzed, and properly interpreted? If not, please provide additional data sources and citations (where appropriate) or alternative interpretations or analyses.					
1			A	In general, the vast majority of referenced literature is from the US, much previous work from the authors of the risk assessment. This is well understandable (it is a US risk assessment and researchers always know their own work best; besides I assume the best experts are hired for the job), but not necessarily appropriate. There is no reference to any literature review, let alone a systematic one and it would have been advisable to do a literature review into *Salmonella* or poultry QMRA, evaluation of the impact of risk-based microbiological criteria and the relevant evidence, before performing the risk assessment. This is a shortcoming, as important studies, unknown to the authors, may have been missed by not reviewing	FSIS conducted a literature review into *Salmonella* in poultry as part of the *Salmonella Risk Profile* alongside this risk assessment (available here). The team responsible for the development of this risk assessment relied heavily on this work and reference to it has been added to the document.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				the literature, and the independence of the risk assessment is strengthened by it. Below, in answers to other charge questions, some relevant studies are mentioned. I have no overview of the available data in the USA, and cannot judge whether relevant data are missing.	
2	38	942-944	A	Reference missing	The reference was added.
3			B	The current QMRA Risk assessment and suggested intervention utilized all relevant data sets curated from all national databases, managed by FSIS and CDC.	No response required.
4			B	The methodology used for data description and statistical analysis used for QMRA model development is standard, as per established procedures and guidelines, and is well referenced. The novel inclusion of bioinformatics virulent factor-based clustering of serotypes is an added advantage in this work to focus the intervention efforts more on the serotypes that have a more direct impact on public health.	No response required.
5			B	The assumptions made while using and modeling datasets are valid.	No response required.
6			C	To the best of my knowledge, the key studies and data necessary for this analysis have been identified, analyzed and interpreted. Where possible, input variables were parameterized using	The work of Nikki Shariat's group was foundational in our recognition of the diversity of serotypes within flocks. However, the results of the Siceloff 2022

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				national representative data. Furthermore, additional efforts have been made to maximize access to relevant industry data. Majority of input variables in this model use datasets that are representative and provide detailed and specific data through consistent sampling and measurement methods. These datasets are suitable for parameterization of variables. However, the characterization of serotype distribution within a flock, which can be used to determine representative cluster-composition schemes, may potentially be improved through a meta-analysis of relevant published studies restricted to research conducted in the U.S. The reviewer is not aware of existing meta-analyses of 1) serotype distribution within a broiler flock or 2) the changes in serotype distribution following a flock from the pre- to postharvest. However, a recently published study by Siceloff 2022 (Nikki Shariat group, 10.1128/aem.00204-22) reported the serotypes observed within breeder flocks monitored through the Georgia Poultry Lab Network and also compared their *Salmonella* monitoring data of preharvest samples with the USDA-FSIS data from carcasses, parts and comminuted collected at processing over the same period, and revealed an apparent discrepancy of *Salmonella* serotype distributions between pre- and post-harvest samples. In this analysis, the USDA-FSIS postharvest data were used to determine the schemes of cluster composition in a flock for the	paper highlight the difficulty of using geographically restricted research data in a risk assessment that is representative of national chicken production. A specific example is that the pre-harvest data did not identify *Salmonella* Infantis until 2019. The increase in the occurrence of the multi-drug resistant version of this serotype began much earlier and it was concentrated in a subset of corporations. From this observation, it would appear these research data are not necessarily representative of the true state of the broiler industry.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				analysis of the impact of receiving guidance. The reviewer recommends that the assessment can be further improved through 1) using preharvest data to determine cluster composition in a flock, and 2) considering the possible shifts in cluster composition from pre- to postharvest in the analysis.	
7			D	In general, given the limited data availability, this reviewer appreciates the efforts FSIS made in generating useful data over the years and used those in this risk assessment. However, clarifications and justifications are needed for several instances and are identified later in this review.	No response required.
8	40	973-975	D	"This carcass data includes samples consisting of a single carcass randomly chosen from a flock at both rehang and then again at post-chill. No attempt is made to choose the same carcass at each location." Please recognize this limitation. While this may be a limitation, for contamination, it would be nice to track the same bird through the shackle line, post-chill, and if possible, the parts from the carcass, so that there will be more certainty in analyzing the contamination and estimating the adverse health effects risk. If possible, FSIS should plan for collecting this data in the future, to better inform future risk assessments.	This and other data limitations have been summarized in **Table17** in **Chapter 3 *Salmonella* Microbial Profile**.
9	67	1568-1574	D	"It is assumed there are 125,115 chicken-associated *Salmonella* illnesses per year. This value is calculated as the product of the total	The statement was revised and additional detail was added to **section 3.2 *Salmonella* Levels** explaining the

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				number of CDC FoodNet cases per year (7,600), the share of these cases that are foodborne (66 percent) and of domestic origin (89 percent), the under diagnosis multiplier for *Salmonella* (24.3) and dividing by the FoodNet catchment area (15 percent). These total cases are distributed across products by assuming the proportion of servings consumed (0.11, 0.83 and 0.06) of all illnesses result from exposure to carcasses (whole chickens), parts and comminuted (ground) forms of chicken, respectively." First, 125,115 chicken-associated *Salmonella* illnesses per year was estimated or calculated, not assumed. Thus, the statement needs to be revised. Second, how the assumption was made regarding the proportion of servings consumed (0.11, 0.83 and 0.06). This is not clear, please explain	proportion of servings consumed (0.11, 0.83, and 0.06) estimates: "In previous work, these estimated fractions were 81%, 13% and 6% for parts, carcasses and USDA Food Safety and Inspection Service. (2015).But subsequent data National Chicken Council. (2022) suggests the share of product marketed as whole carcasses has decreased. Therefore, we adjusted the carcass share down and increased the parts share, accordingly."
10	88	2029-2032	D	"As expected, our analysis demonstrates that higher levels of *Salmonella* on raw products are associated with higher risk of illness, on average, compared with lower levels. Nevertheless, the frequency of level and magnitude of level are inversely related; lower levels occur much more frequently than higher levels." This is a valid point. However, the concentration of bacteria can significantly increase in foods with low initial levels of bacteria in some situations such as cross-contamination or recontamination, or undercooked samples, which experience improper storage conditions.	See response to Q1 General Comments #8-10

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				The attenuation multiplier used to describe *Salmonella* growth and die-off after raw chicken product leaves processing does not account for these situations. This needs further attention and must be addressed in the modeling approach (please see Q4 (e)).	
11	168	4-6	D	Data and Data Analysis: "Prevalence estimates were based on the most recent calendar year (i.e., 2021) of data available at time of analysis for parts and comminuted chicken products." It is not clear why only one year of data was used instead of last several years, please explain.	The most recent year's data were used for estimating the prevalence because the annual prevalence estimates have been decreased substantially over the last five year which would lead to a bias to standard deviation ratio of the prevalence estimator that is unacceptably high (Cochran, 1977). For example, the prevalence of *Salmonella* contaminated chicken carcasses decreased from 0.058 to 0.033 between 2017 and 2021 while the estimated standard deviation of the chicken carcass prevalence in 2021 was $\sigma = 0.0021$. Thus, maintaining a bias-standard deviation ratio (B/σ) of less than the generally accepted values of 0.1 to 0.2 is unlikely (Cochran, 1977). Cochran, W. G. (1977). Sampling Techniques. (3 ed.). New York: John Wiley and Sons.
12	168	11-19	D	Data and Data Analysis: 1998-2019 FSIS Chicken Carcass Verification Program "The primary use of the data is to illustrate the change in Infantis over	The initial FSIS' Key Performance Indicators (REF) included Infantis on the basis of the best available human illness

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				time". What is the rationale behind focusing on Infantis? Please include this.	data at that time, so FSIS evaluated this serotype clustering further. The most recent available data was used throughout the document to describe the current state of *Salmonella* contamination in U.S. chicken production and the associated health effects. In a few instances including when looking at *Salmonella* Infantis in section 2.2 **subsection Inclusion of Salmonella Infantis in Cluster 2**, it was necessary to do a retrospective trend analysis as part of the genomic clustering of serotypes.
13	180	324-325	D	"For all sampling programs, FSIS collected samples within an establishment at regular intervals, so an assumption of systematic sampling is reasonable". This is a nice approach.	No response required.
14	181	362	D	"Methods for estimating the dose distribution"-Is it the method for dose distribution or distribution for contamination level or concentration? "Dose" implies the number of pathogen ingested after exposure assessment (or dose), which is then integrated in dose-response equation to provide the risk estimate of adverse health effect.	The title of this subsection has been changed to "**Methods for estimating the contamination distribution**"
15	190	600-602	D	"To simplify the development, most of the probabilistic components will be replaced by fixed values." Please provide more information on which	The sentence in **Appendix C** subsection "**Methods for modeling illnesses**" was poorly worded. It has been replaced by the following:

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				probabilistic components were replaced with fixed values.	"For simplicity, the symbols for each variable used in the development of the methodology are treated as fixed values. The parameterization of the probability distributions used to describe variability in factors such as the consumed dose are summarized later."
16	191	631-634	D	"There is no evidence to suggest that any flock of broiler chickens is truly free of *Salmonella* contamination (De Villena et al., 2022), thus the preferred parameterization of Assumes P(ill) that all servings have the potential for some level of contamination, so that the random variable describing dose *D* describes the average number of pathogens in each serving." This is a reasonable assumption. However, it may be possible that not all chickens are contaminated. For example, researchers have found that the prevalence of *Salmonella* within a contaminated flock was ~57%, with within-flock prevalence of *Salmonella* for positive flocks was 17.2, 8.1, and 53.9% for ceca, crops, and neck skins (Mainali et al. 2009. Journal of Food Protection. DOI: 10.4315/0362-028x-72.10.2202). Please consider different proportions of chicken contamination within a flock by using a method such as profiling/uncertainty analysis.	There is no assumption that every chicken is contaminated, only every flock, as has been clarified in **Table 9: Risk assessment information and assumptions**. The model lacks the resolution to make such an uncertainty analysis feasible.
17			E	The risk assessment represents a very substantial and expertly applied treatment of past and more current data on *Salmonella* sampling. Inferences	No response required.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				related to concentration and prevalence seem to be very carefully applied.	
b. Have the strengths and limitations of the data been transparently explained?					
1			A	Strengths and limitations have been explained scattered over the report, and Appendix B describes the data sources and the data. However, a concise overview of the data used to answer each of the risk management questions and the strengths and limitations of these data sources is missing. This would increase the transparency.	In addition to the discussion of strengths and limitations of data sources when they are used throughout the report, a limitations column has been added to Table 17: Description of main sources of data used in the risk assessment in **Chapter 3 *Salmonella* Microbial Profile**.
2			B	All extensive data sources, mainly through government agencies are clearly listed in terms of the number of samples used and time points of data sets. The limitations of availability of data and the associated reasons including less surveillance during the Covid pandemic and less sampling for certain product categories are also mentioned.	No response required.
3			C	The majority of data sources used in this analysis were well-described, with explanations of their limitations. However, an exception was noted regarding the inadequate description of the data used to derive the attenuation distribution (as outlined in Q4-e). Additionally, certain variables were estimated deterministically, leading to an inadequate characterization of the accompanying uncertainty/variability in the predictions. Given the absence of robust data, the approach used is acceptable. However, a sensitivity analysis could	A full description of the attenuation model development has been added to section **1.5 Model Approach** and a sensitivity analysis has been added to **Chapter 5 Final Product Standards**.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				be advantageous to better comprehend the potential impact of these uncertain variables. For example, on Page 87, Table 25, the changes in the estimated mean of illnesses avoided per year can be quantified by varying values of input variables *c*, *h*, *k*.	
4			D	The strengths and limitations of the data have been explained well. Please see below the details and specifics to further support some statements.	No response required.
5	42	1020-1023	D	"For the poultry industry, *Salmonella* and *Campylobacter* occurrence is more frequent on products produced by lower-volume establishments. The opposite phenomenon is observed in the pork and beef industries, where a small number of large establishments account for the majority of the contaminated product reaching consumers." Could you please provide some reference and evidence to support this?	The negative correlation between pathogen occurrence and production volume for broiler chickens is discussed and demonstrated on page 7 of the following open-access publication. Williams, M. S., Ebel, E. D., Golden, N. J., Saini, G., Nyirabahizi, E., & Clinch, N. (2022). Assessing the Effectiveness of Performance Standards for *Salmonella* Contamination of Chicken Parts. International Journal of Food Microbiology, 378, 109801.
6	168	21-30	D	2022 Young Chicken Carcass Sampling Program-It seems this data collection is nice. In the future, when more data come through the carcass sampling, it would help in updating this risk assessment.	We agree. The complete data set was used in the risk assessment covering August-October 2022.
7	168	26-27	D	2022 Young Chicken Carcass Sampling Program-"Additional laboratory testing was performed to quantify indicator organism (APC and	Yes, the sampling locations for carcasses throughout this document are rehang and post-chill.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				Enterobacteriaceae) at both locations." Does both locations mean Rehang and Post-chill?	
8	175	208	D	NHANES Chicken Consumption Data Analysis: "The risk assessment used the GRM50 estimates to derive empirical probability of illness estimates." What are GRM and GRM50? Also, this acronym was not presented in the glossary.	Definitions of these variables are in Appendix B where they were estimated. Since this is not standard terminology, but a variable name, it was not added to the glossary.
9	178	273	D	"The best fit is for a gamma distribution". Was @Risk software (Palisade) used? Please mention the software used for this and figure 44 on Page 178.	The @Risk reference was added.
10			E	Many of the strengths and limitations of the data are well characterized and transparently explained. This reviewer chose to focus on the distribution of doses and what could be characterized as a question of alternate means of estimating risk that would not require the complexity of the numerical integration and the complexity of the dose-response fitting that was undertaken. FSIS explicitly seeks a "fit for purpose" risk assessment, and the analysis that follows contemplates an alternate approach that FSIS may also consider to be fit-for-purpose. This (much simpler approach) could be used to compare with FSIS' results, and may allow for an alternate means of communicating the very abstract nature of the stochastic processes involved.	Comment 10-12 are addressed together: To fully address the details of these comments, a **Techniques for approximation** subsection was added to **Chapter 5**.
11			E	One limitation (or perhaps more accurately, an important consideration) of the overall situation is	In general, we agree with this comment. Nevertheless, our analysis must discriminate between "passing" and

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				not described at all. This relates to the implicit situation associated with the spectrum of average doses assumed to occur in poultry servings. There is a relatively low level of contamination (combining prevalence-concentration concepts) in the raw poultry products, which is then combined with an attenuation multiplier distribution that reduces levels even further for a very high percentage of servings. The dose distribution is described as the product of the two lognormal distributions, and the means of calculating the two moments of the resulting lognormal distribution are provided at several points. The parameters of the final dose distribution may not have been shown (this reviewer couldn't find them), but are f(d) = LN(-8.04, 2.3). This is an extremely wide distribution for which 50% of the average doses are less than 10^{-8} cfu/serving (e.g., 1 cfu per 100 million servings), and 99% of the servings have a value less than 1 cfu per 100 servings). The average dose of 1 cfu per serving (0 $\log_{10}$ cfu) is 3.5 standard deviations above the median (i.e., at the 99.98th percentile). The authors are aware of this, but they should share this insight with the reader. In terms of the actual physical doses experienced by consumers, a huge percentage of servings (>99.9%) will have no *Salmonella* contamination at all, as a result of the Poisson random variable part of the Poisson-Lognormal mixture distribution of actual (as opposed to average) doses. As is	"failing" units at the end of production based on a concentration criterion, then consider the influence of 2 *Salmonella* clusters, and then re-mix the resulting probability of illness estimates. From the simplest conceptual outline of this model, the math is somewhat messy. As part of our sensitivity analysis, we've employed the simplifications suggested here. Our simplified model illustrates that the primary driver of the predicted effectiveness of alternative policies is the change in average initial contamination that results from diverting a share of failing lots and replacing them with average lots. The specific parameters of the attenuation distribution or dose-response function are not strong influencers of the model's predictions about the effect of the policies considered.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				expected for very low average concentrations, the next most probable Poisson outcome is that there will be only one cfu per contaminated serving. Each possible other outcome of the Poisson process (d=2 cfu, 3, 4, etc.) will be less and less likely, since P(n) = P(n-1)*lambda*(n-1)/n, i.e., very rapidly diminishing when lambda is very small, whereas the increased risk for R(n) = R(n-1)*n/(n-1), such that the probability-weighted contribution to risk for each increasing integer dose approaches zero rapidly. This outcome (many uncontaminated servings) is described in the document in a qualitative sense, but it should be stated in a quantitative way since readers may not understand the extremely rare nature of the contamination and the very low doses that correspond to positive servings. It may also foster an improved understanding of the actual range of doses that underlie the rest of the analysis. It also allows for contemplation of much more simplified analyses that may be accessible to a much larger audience, and also simplify the estimation of the impact of risk management activities. The implications of the rare event nature of the serving contamination, and the very low doses (most commonly only 1 cfu) among the contaminated servings is that some simplified calculations are possible. A histogram showing the result of a Poisson-Lognormal mixture distribution would be very instructive showing the probability	

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				mass assigned to 0, 1, 2, 3, >3, etc.) would be instructive to the reader. FSIS estimates that there are 73 billion servings of poultry consumed annually in the US. The arithmetic average of the dose distribution is 0.011 cfu/serving (-1.95 log10 cfu/serving calculated as $10^{(u+0.5*ln(10)*sigma^2)}$ for the parameters of the dose distribution). By extension, there are 818 million cfu consumed per year among these servings, most commonly as 1 cfu/serving when contaminated. FSIS estimates approximately 125,000 cases per year of Salmonellosis associated with this poultry consumption. So on average, there are 150 cases per million cfu, or 0.00015 cases per cfu. This corresponds to an estimate of the mean of the Beta distribution, or the ratio of alpha/(alpha+beta). A similar analysis can be conducted to distinguish the two serovar clusters, using the arithmetic means of the dose distributions and relative prevalence of the serovar clusters and the number of cases per year from each cluster. This is similar in concept to the approach taken with the full dose-response analysis, but does not require such complex mathematical assumptions. A further implication of this analysis is that the impact of risk management measures applying to raw poultry can be estimated to a very reasonable approximation by their impact on the arithmetic average of the concentration distribution. Ultimately, it is the sheer number of pathogens surviving to contaminate servings, largely one at a	

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
12				time per serving, that determines the risk. When the median and average attenuation is so significant (median of 8 log reduction), the exact pattern or shape of the distribution is less important, while the arithmetic average of the concentration distribution ultimately drives the number of potentially surviving organisms downstream.	
			E	The derivation of the dose-response curve spanning from -10 log to +9 log cfu/g, and then parameterizing it with a 9 term polynomial regression, on the basis of 100s of millions of single cfu exposures seems somewhat questionable. That is, although the approach may be found to be mathematically “correct”, the approach seems “overwrought” given the evidence base. In addition, the characterization of the uncertainty in the dose-response assessment using alternate 9 term polynomials, would seem to misstate the true level of uncertainty, given the true, overall level of uncertainty expressed elsewhere in the document, and the fact that the evidence is based on, by an overwhelming majority, single-cfu exposures.	

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
c. Given the differences in the data and sampling methodologies, is the overall modeling approach used for chicken products (carcasses, parts, and comminuted) appropriate?					
1			A	I would say it is fit for purpose. As I note below, the use of the same attenuation distribution for all products is debatable but defendable, I would not expect that another approach would give very different results, among others because the authors work with relative risks where uncertainties partly cancel out. Further, I understand you have to deal with the existing differences in a pragmatic way. So it is appropriate for your context.	The use of a single attenuation distribution was further evaluated in the uncertainty analysis that has been to **Chapter 5 Final Product Standards**.
2			B	The overall modeling approach is valid and efficient and using statistical modeling has shown that qPCR technology is limited by its poor discriminatory power, especially for samples with low *Salmonella* contamination, in correctly classifying the samples as *Salmonella* positive. So, using qPCR quantification-based interventions aiming to lower the *Salmonella* detection threshold will be of limited use in terms of their practical applicability. This is an important prediction.	No response required.
3	53	1253-1260	B	Please rephrase to make it clearer, the first paragraph and the equation for Misclassification at the Post-Chill. Equations and texts don't sync clearly.	This was an error in the equation. LOD should have been LOQ. The error has been corrected.
4	51	1216	B	Please add unit in ??: "high levels (0.75 to 4.43 ??)"	The unit was added.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
5			C	The modelling approaches employed for various products are deemed acceptable. It is recognized that due to inadequate data availability, particularly for parts and comminuted products, certain risk management options, such as the implementation of performance standards associated with indicator bacteria, could not be considered. However, as more data become available, the modelling concepts presented in this analysis seem acceptable.	No response required.
6			D	The overall modeling approach used for chicken products (carcasses, parts, and comminuted) is reasonable and appropriate.	No response required.
7	48	1124-1130	D	"Simulating the shape of log10-transformed realizations drawn from the three individual distributions demonstrates that the resulting distribution ….For these reasons, a lognormal distribution was chosen to model the overall level per gram of *Salmonella* at the end of production." This is a reasonable assumption.	No response required.
8	169	68-72	D	2012 Chicken Parts Baseline-"Samples were collected from January through August of 2012. Although this survey collected samples from many different types of chicken parts, only the data related to breast, wing and leg sampling were used here. We chose these parts because they had similar levels of *Salmonella* and Campylobacter contamination and constituted about 90% of the chicken parts produced in the United States (FSIS,	The percentage breakdown per legs, breasts, wings is not documented as part of FSIS chicken parts sampling. For more information, on random selection of chicken parts see the FSIS sampling instructions in New Sampling Instructions and Testing for Chicken Parts and NRTE Comminuted Poultry | Food Safety and Inspection Service (usda.gov)

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				2015).” Please clarify what percentages are for breast, wing, and leg, separately.	
9	199	833-842	D	Data regarding contamination (prevalence and concentration), serovars, and meta data, at different steps of poultry production and processing (e.g., at the farm (flock), arrival at the primary processing plant, rehang, pre-chill, and further processing) would be helpful for future risk assessments. FSIS mentioned, as a part of cooperative agreement, it is working closely with industry partners to better understand these data gaps and develop data sharing strategies. Data from industry partners would be helpful.	No response required.
10			E	Although it may not be material to any of the results, it is not clear to what extent FSIS considers the Poisson variability associated with product sampling when calculating Prevalence and the probability of a positive sampling result. For example, for a 30 g sample, the Limit of Detection is expressed as 1/30 cfu/g. While this is true on average, product contaminated at a concentration less than 1/30 cfu/g can and will result in a positive sample simply due to Poisson variability in the sampling process (a sample happens to pick up a bacterium, but most of the time does not). The probability of a positive sample at the Limit of Detection is not 0, it is P(Poisson>0) = $1-e^{-C*m}$ which for a 30g sample and a concentration of 1/30 cfu/g is $1-e^{-1}$ or 0.63. Even when the concentration is $1/10^{th}$ of the LOD, the probability of a positive	There were three approaches used to fit the concentration distributions. The use of the Poisson-lognormal for the qPCR data and the MPN method for the remaining product-pathogen pairs explicitly incorporates the Poisson variability. While the AC data are the result of an MPN experiment, the levels are sufficiently high that inclusion or exclusion of Poisson component has little effect on the bias of the parameter estimates.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				sample is approximately 10% ($1-e^{-0.1}$). It is not clear how these factors into FSIS's treatment of sampling results. The use of the Poisson-Lognormal distribution theory is applied clearly in the qPCR section, but it's use elsewhere is less clear to this reviewer.	
d. Are the differences in data for the three chicken products adequately described and addressed?					
1			A	The presented data and the description in the three products are clear to me.	No response required.
2			B	Yes, properly described and tabularized satisfactorily	No response required.
3			C	Yes. The data sources, sampling coverage, sample size, laboratory methods were adequately described for the three chicken products.	No response required.
4			D	The differences in data for the three chicken products are adequately described and addressed. However, some analytical choices and assumptions must be explained further. Please see below for specific comments and details.	No response required. Additional explanations are provided with the specific peer reviewer comment.
5	171	142-143	D	"These missing results were addressed by randomly imputing sample results from samples where the MPN analysis was performed." Please include how imputing was performed.	The text of the risk assessment was revised and this reference was added to address this comment: van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software, 45*(3), 1 - 67.

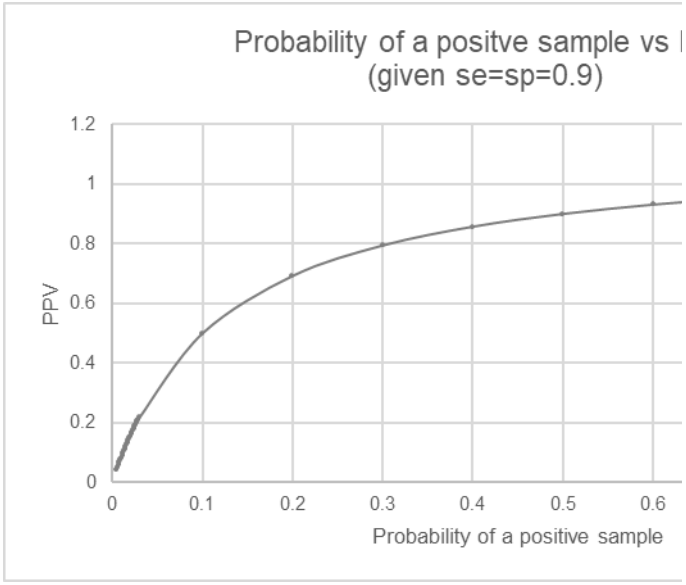
Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
					https://doi.org/10.18637/jss.v045.i03
6	184-185	457-469	D	Bias adjustment for whole carcass levels-"While the precision of the qPCR level data is low…..has an 80% chance of being declared statistically significant." Did the authors perform the power calculation/analysis? The example is good, please provide more information.	This result, and other details related to the statistical properties of the FSIS commodity surveys, will be published in the future.
7	185	475-479	D	"A value of -0.4 was chosen because the bias is expected to lie somewhere between -0.35 and -0.65. The lower end of the range was chosen because the Poisson component of the Poisson lognormal distribution is likely to account for some of the measurement error in the underlying level." It is not clear the lower end is -0.35 or -0.4, please explain.	The bias of an unbiased estimator is 0. In a previous study, the range of observed biases associated with ignoring the measurement error component in this application fell in the interval (-0.65, -0.35). Given that the Poisson component of the likelihood function is expected to address a portion of measurement error, the chosen bias correction was on the lower end of the range of observed biases.
8	187	553-554	D	"These missing results were addressed by randomly imputing sample results from those samples with results." Please mention how random imputing was performed.	The text of the risk assessment was revised and this reference was added to address this comment. van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software, 45*(3), 1 - 67.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
					https://doi.org/10.18637/jss.v045.i03
9	185-186	486-498	D	Methods that scale rinsate levels to whole carcass levels- "…Attachment characteristics of *Salmonella* are assumed to be more consistent with Enterobacteriaceae, so the 14% removal rate was chosen with uncertainty in this estimate characterized r ~ beta(14,86)". The assumption was based on *Enterobacteriaceae* data. While this is an adequate approach, please provide more information and reference to substantiate this assumption that attachment characteristics of *Salmonella* are consistent with *Enterobacteriaceae*.	Enterobacteriaceae are a family of facultative anaerobic bacteria including *Salmonella*, *Shigella* and *Escherichia coli*, so *Salmonella* are Enterobacteriaceae. The text was edited to clarify that we are assuming the attachment characteristics of *Salmonella* are assumed to equivalent to that of all bacteria in the class.
10	186	500-507	D	Methods that scale whole carcass levels to a per serving level-The implicit assumption is that all servings will have equal level of bacteria. While this assumption is valid and adequate, given the lack of such data, this was not mentioned. Please mention this in the report. In addition, it is possible that different servings may have different levels of pathogens and this needs to be recognized.	FSIS subject matter experts respectfully do not agree with this assertion. The lines of text in question describe the mean of the distribution. The variance term still ensures that the concentration in all servings differs.
11	186	508	D	"The carcass levels can also be converted to a per gram basis." Please show the calculation, as an example, as the analysis was made in mL basis for carcass and parts.	The formulas and parameters for all values used in the models are now summarized in **Table 11: model parameters and variables for final product standards and receiving guidelines estimates**.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
12	187	533-535	D	"The same adjustment factor can be applied to the carcass level distribution to convert that level distribution to a per gram basis." It is not clear why the analysis was done on mL basis not gram basis. Please show the calculation, as an example, as the analysis was made in mL basis for carcass and parts.	The sampling and laboratory methods for carcasses and parts use a unit of measure based on milliliters of rinsate. These values are then converted to a per gram basis for servings. The formulas for the conversion are provided in **Table 11: model parameters and variables for final product standards and receiving guidelines estimates**.
13			E	With respect to the data, the differences are adequately described and addressed.	No response required.
e. Are the strengths and limitations of the quantitative PCR data method used to generate data for carcass contamination transparently explained and adequately assessed.					
1			A	The qPCR data method is well-described, apart from the issues mentioned below.	No response required.
2	50	1175	A	Provide a proper reference to the NACMCF report	A reference to the NACMCF 2023 report has been added.
3	51	1209	A	What is a crossing point value? I assume it is the CT value?	The crossing point is BioMerieux GENE-UP version of other qPCR instruments' CT (crossing time) value.
4	51	1216	A	y is not explained	The definition of variable y has been added parenthetically.
5	51	1218	A	When Figure 10 is introduced, I am not yet able to understand it. It only gets clear on page 52. That is confusing, please explain it sooner	The order of the figure and its explanation were adjusted to improve clarity.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
	Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.				
6	52	1235-1237	A	So the results given here are for the equation in line 1233. I had to read it a few times to be sure, please clarify it in the text.	The correction to the equation addressed in comment 3B to Q1.c.
7	52	1249	A	The analysis concludes that the PPV obtained are too low to be acceptable. This is justified, because qPCR results above the LoQ are most likely when the true level is between LoD and LoQ. This is well explained in lines 1263-1270. The conclusion (line 1284) that nearly all finished product carcass and parts samples that are assigned enumerated values above the LOQ of 10 cfu/ml using qPCR would be incorrectly classified is adequate.	No response required.
8	53	1257-1270	A	Where do the values 135, 505 etc. come from?	The correction to the equation addressed in comment 3B to Q1.c clarifies this point.
9	55-56	--	A	How do you explain that 1 and 2 logs and 3 and 4 logs overlap, but 2 and 3 do not?	This lack of overlap is an unexplained feature of qPCR technology that has been observed in two different datasets (Chaney (2022) and a confidential dataset shared with FSIS). Theoretically, this clustering of the results shouldn't exist. At this time, the technology does not have sufficient discriminatory power to conclusively enumerate individual samples with a high degree of accuracy. It is a technology whose power lies in repeated trials and trend studies (De Villena, 2022).

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
10			B	The modeling approach has clearly depicted the limitations of qPCR and highlights its non-reliability as an efficient sampling method, especially for low levels of *Salmonella*. The model has shown with statistical proof that there is no significant difference in removing lots based on qPCR enumeration than removing a portion randomly. This finding has set stage with statistical support, for leveraging alternative intervention approaches like virulent serotype-based DR modeling and risk assessment for preventing public illnesses more effectively.	The results of our modeling cannot be extrapolated to all methods that utilize qPCR in whole or in part. No response required.
11	64	1449-1453	B	Please rephrase the language to explain the third finding more clearly. Not very comprehendible.	This finding has been rephrased.
12	64	1443	B	Is there a typo in the table number, should it be table 17? “Comparing to Table 16 observe”	Caption and cross-reference errors have been corrected throughout.
13	52	1249	C	The potential misclassification associated with the qPCR method was transparently explained. However, additional clarification is needed regarding the statement about the low positive predictive value (PPV). Below is a curve of the changes in PPV as a result of varying probability of a positive sample, which was generated based on the equation of PPV calculated with p, se, and sp. Even for a method with relatively high Se and Sp, such as 0.9 used to generate the curve, the PPV can be very low when p is very low. According to	The low PPV is highly attributed to the small proportion of samples that contain *Salmonella* levels above the LOQ, as has already been noted in the report, and this limitation will be relevant to other enumeration methods with high LOQs. For further description of the PPV see this informative article:

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				Figure 10, the p(>=LOQ, 10 CFU/ml) among chicken carcass is fairly low. Hence, it may be worth noting in the report that the low PPV is primarily attributable to the fact that only a small proportion of samples contain *Salmonella* at levels above the LOQ, which is an unavoidable limitation. This limitation might also be relevant to other enumeration methods. 	Spann, P. (1988, March 7). False Positive the AIDS Test Nightmare. *Washington Post*. https://www.washingtonpost.com/archive/lifestyle/1988/03/07/false-positive-the-aids-test-nightmare/f343d4a3-f399-4b0c-9db9-d9b7c74d29cf/
14			D	The strengths and limitations of the quantitative PCR data method used to generate data for carcass contamination were reasonably explained and assessed.	No response required.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
15	54	1290-1292	D	"Furthermore, current and future risk assessments assessing the public health impact of *Salmonella* level that use data from an assay with a low PPV should be evaluated closely before use." Has such an evaluation been performed before? Could you please elaborate further?	FSIS is not aware of statistical treatments of the topic other than that presented in this document.
16	182	400-403	D	"Poisson process assumption most likely still underestimate the variability at low levels because the qPCR estimates have estimated standard deviations on the order of 0.5 log10 at levels of 2 log10 cfu/ml in chicken rinse samples". Although a valid assumption, please explain the rationale for using this process.	Poisson process assumption was used as a conservative estimate in line with other treatments of enumeration data. The underestimation is explained by the numbers from the reference.
17	53	12463-1270	D	The poor performance of the current qPCR technologies has been described in this section.	No response required.
18			E	Yes. The treatment of qPCR data is transparently explained with respect to limitations and adequately assessed.	No response required.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q2. Please identify limitations, weaknesses, or inadequacies of the bioinformatics serotype clustering; please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				General Comments	
1	135-145		A	In principle it is a clever approach to define clusters of serotypes to obtain different classes of virulence, and differentiate the DR relations for the different clusters. I am not experienced in bioinformatics, so it is hard for me to comment on the details of the approach.	No response required.
2	141	248-250	A	As a relative outsider, I lack an explanation of why this approach is actually needed. The objective of the exercise seems to be to define virulence clusters, which have a relative risk as defined in lines 248-250, and is calculated by dividing the relative frequency of a cluster among ill people with that among poultry samples. You can use any type of clustering for that, you can also do it at serotype level (without using any bioinformatics). I lack an explanation of the added value of the bioinformatics clustering, please make sure to include one.	The objective of the work was to use genomics to classify serovars into clusters based on similarities of Virulence Factors (VFs), and to assign appropriate dose-response models to the underlying serovar clusters. Clusters segregated in this way had distinct and robust epidemiological characteristics (i.e., the risk multiplier). Additional epidemiological outcomes such as hospitalization and invasive illness were examined in previous work conducted by EpiX Analytics in 36781801 (medrxiv.org) which exhibited differences by cluster (see Figure 4 and Supplemental Table S4). Although the work could have begun with serotypes, FSIS sought to avoid ignoring any underlying genetic variability present within many serotypes. Other approaches including high resolution genomic analysis are promising, but because they are so computationally intensive, they have only been applied to a limited number of strains or have

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q2. Please identify limitations, weaknesses, or inadequacies of the bioinformatics serotype clustering; please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
					focused on a single subtype, which was not appropriate for this risk assessment.
3	Appendix A		A	It is quite hard to read the report in Appendix A, as it is difficult to differentiate between headings and subheadings and the terminology is not always consistent. For example, equation 1 on p. 140 describes the probability that strain s belongs to cluster Ci as Pr(s ϵ Ci) but then this terminology is not used in the section of the occurrence of *Salmonella* in poultry, which I think refers directly to that. Further, I would be much helped if it was clarified why you do what you do, and not only what you do.	The headings and subheadings have been more clearly noted and are now consistent with the table of contents. FSIS is also providing an expanded description of the clustering process in new *Bioinformatics Supplemental Materials*.
4			A	Last point: I am not sure this whole exercise is particularly relevant for the risk assessment. In the end, two clusters of *Salmonella* are defined based on well acceptable criteria (i.e. we see a difference in virulence between those clusters) and that may help us at some point in the risk management (if we find a cluster 1 sample there is more reason to do something about it than if we find cluster 2). I doubt whether more detailed analyses are needed here. In, for example, the analysis of the paired cluster results (section 3.6) we see that it is challenging to identify meat samples as being contaminated with a specific cluster, I would	The paired cluster results (**section 3.5**) suggests that both clusters can occur in sampled flocks/lots although a general concordance rate (60%) of *Salmonella* serotypes was observed at rehang and post-chill. Nevertheless, serotype mixtures within flocks/lots can be described/tested via different schemas (section 5.1); however, this is limited to carcass data. Multiple point per lot sampling data is required for comminuted chicken and chicken parts, as exists for carcasses sampled at rehang (receiving) and final post-chill, to potentially appreciate further use of the different clusters in terms of risk management.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q2. Please identify limitations, weaknesses, or inadequacies of the bioinformatics serotype clustering; please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				guess this uncertainty is much larger than what improved bioinformatics cluster analysis could solve.	However, this is an objective, data-based, and repeatable approach that defined stable high-risk and low-risk clusters, and therefore is fit for the purpose of designing supportable risk management options.
5			B	Overall, very nice approach to consider the virulence of the hazard in the model estimation of the risk assessment thresholds, suitable from the perspective of public health as well as the product industry. Minor textual errors detailed below.	No response required.
6			C	The report describes approaches used for clustering serotypes based on their virulence characteristics and modelling dose-response relationships for each identified cluster in a clear and thorough manner. The methods employed were unique and innovative, which provide valuable insights for the further improvement in the use of omics information for hazard characterization and other components in risk assessment.	No response required.
7			D	The authors have employed current best practices in the analysis of genomic data. However, the rationale behind certain instances of data selection should be justified and in some cases, reanalysis and revision are needed. Specifically, inclusion of beef data (see 2(a) below), exclusion of serovars with	These comments are fully addressed in 2(a) and 2(c) below.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q2. Please identify limitations, weaknesses, or inadequacies of the bioinformatics serotype clustering; please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				less than 50 assemblies/isolates in the initial machine learning dataset/virulence loci matrix (see 2(c) below), and exclusion of FoodNet outbreak data (see 2(c) below).	
			E	(no comment from this reviewer)	No response required.
a. Was the Salmonella genomics data appropriately curated and processed?					
1			A	I am not able to judge that	No response required.
2			B	All big repositories for compiling the data were used. 1. For the bovine-, chicken-, turkey- associated or beef, chicken, and Turkey origin: NCBI (PRJNA242847)USDA Food Safety and Inspection Service CDC isolated strains, FSCIS HACCCP FSCIS_regular) NCBI (PRJNA292666) WGS for Foodborne Pathogens of US Department of Agriculture surveillance project for NARMS fsis_narms) NCBI(PRJNA292661) WGS for Foodborne Pathogens of US Food and Drug Administration surveillance project for NARMS fda_narms 2. For Human clinical cases: (as mentioned in the text) Used NCBI (230403) CDC PulseNet USA surveillance for food-borne disease for sporadic and domestic acquired cases.	The dataset was intended to serve as an example to parse the NCBI metadata related to human clinical cases.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q2. Please identify limitations, weaknesses, or inadequacies of the bioinformatics serotype clustering; please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				Outbreak cases attributed to beef, chicken, and turkey came from National Outbreak reporting system. # Comment: in the code 'NCBI_parse.RMD' file, In the first section of data import, I only see a file input as 'norm' without many annotations about its sources (only says List of SeqID provided by JP). Please clarify, is this file compiled with both clinical (sporadic, domestic acquired cases) and NORS-reported outbreak cases?	
3			C	Yes. Stringent quality control measures were employed to eliminate low-quality data from further analysis.	No response required.
4	136-138	96 156, 163–168	D	The *Salmonella* genomics data appears to be appropriately processed (using prevalent analytical tools and techniques, and best practices for quality control). However, the *Salmonella* genomic data curation looks problematic because of the inclusion of genomic assemblies isolated from beef.. This risk assessment was focused on chicken and chicken parts. Previous research suggested that there was a great genomic diversity in *Salmonella* isolated from different sources/species. For example, differences in virulence gene expression were observed in *Salmonella* isolates from different sources –	The EpiX Analytics team assumed that the clustering results would not depend on the species where the isolates originated from (see, Table 55. To test this assumption, EpiX Analytics performed a prior analysis which included a variety of isolates originating from multiple species and the isolates categorized in the same clusters regardless of origin(Fenske, 2022). By including beef-related isolates, clustering was accomplished with over 40,000 *S. enterica* isolates from human, poultry and beef sources, which resulted in robust and stable cluster designations (k=2, 3, 4) with more isolate stability for less common

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q2. Please identify limitations, weaknesses, or inadequacies of the bioinformatics serotype clustering; please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				pigs, chicken, cattle, (Table 2; Pavon et al. 2022. BMC Microbiology, DOI: https://doi.org/10.1186/s12866-022-02697-6). Accordingly, please remove beef data and reanalyze the remaining data for clustering analysis.	isolates. Moreover, the risk multipliers were estimated using poultry associated outbreaks and food isolates. Therefore, serotypes rarely encountered in poultry (and more often in beef such as Newport and Dublin) do not significantly contribute to the risk multiplier estimate, while serotypes that are common in poultry (e.g., Typhimurium, Enteritidis, and Kentucky) do. For these reasons, FSIS respectfully disagrees with the need to conduct a new clustering analysis.
5			E	I do not have sufficient expertise or experience in the curation and processing of genomics data to scrutinize this part of the risk assessment.	No response is necessary.
b. Are the databases and methods used to determine virulence factors appropriate? Should any other virulence factors have been considered?					
1			A	I am not able to judge that	No response required.
2	139	187-201	A	I believe the authors do not determine any virulence factors, they use the ones that are available from databases. I cannot judge whether these databases are appropriate, but have no reason to doubt that they are.	No response required.
3	139	188-201	B	Overall steps and methodology implementation for annotation for virulent factors are good.	No response required. Response follows in the next comment.

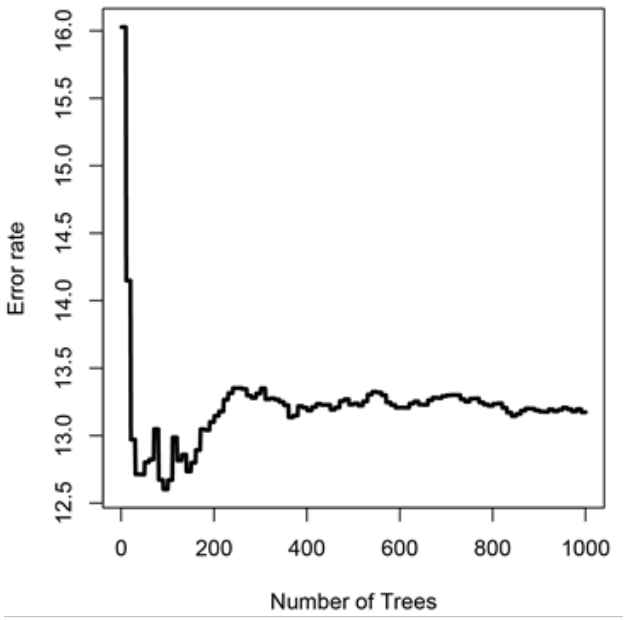
Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q2. Please identify limitations, weaknesses, or inadequacies of the bioinformatics serotype clustering; please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				However, the language to implement the steps should be rephrased to make it clearer, specifically for the unclear statements as mentioned in the following 3 rows:	
4	139	193	B	Not a proper rationale or reference is provided to support the choice of the reference genome of Typhimurium serotypes, for this task.	Open reading frames (ORFs) from *Salmonella* Typhimurium reference strain LT2 were used to help identify and annotate Virulence Factors from the non-redundant database. The reference strain was not used to exclude any ORFs already present in the database and therefore was not expected to significantly affect the VFs used for clustering. Strain LT2 was chosen primarily because it is derived from a complete genome sequence and therefore full coordinates (ORF start and stop) for any matches would be readily available. This reference strain is also commonly used in *Salmonella* genomics. FSIS has developed a *Bioinformatics Supplemental Materials* (available here) to clarify steps of the process.
5	139	192-194	B	Not a very clear description, were the ORFs of virulent factors from the custom database and ORFs of Reference combined? Please rephrase and make it a clear description of the process implementation.	
6	139	196-197	B	Not very clear description of the step performed, was the parsing step performed after PROKKA annotation or before? Please rephrase	Prokka gene annotation pipeline can run several processes simultaneously, which FSIS has described in the *Bioinformatics Supplemental Materials* (available here) to provide additional clarity. Once the VF

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q2. Please identify limitations, weaknesses, or inadequacies of the bioinformatics serotype clustering; please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
					database was reduced into a non-redundant dataset composed of representative sequences, these results were used to define the primary annotation database for consistent gene naming in the isolate assemblies. The parsing mentioned in these lines could imply: (1) identifying how Prokka annotated VF factors or (2) processing the annotation on the isolate assemblies to determine presence/absence of each VF to identify the gene profiles. Text in **Appendix A** has been amended to clarify this process.
7			C	This project relied on two primary databases, PATRIC and VFDB, which are recognized as the most important resources for *Salmonella* virulence genes. There are other databases that could potentially provide additional information. >*Salmonella* Genome Database (SGD), a comprehensive database of genomic and functional information on *Salmonella* including virulence genes. >CARD, comprehensive antibiotic resistance database, also includes information on virulence genes for various bacterial pathogens, in addition to ARGs The reviewer suggests comparing the virulence genes included in PATRIC/VFDB	We thank the reviewer for this information. The purpose of using the VF database was to distinguish clusters based on presence/absence among the strains/assemblies. For that purpose, only VF represented on at least 10 assemblies and no more than 95% of assemblies were included. Furthermore, inherently virulence factor databases continuously evolve and as new information through research is produced, virulence factors should be assessed, incorporated into the model, and the clustering re-analyzed. FSIS is also exploring how to utilize other database such as the VirulenceDB tool developed in a collaboratively work by a group of scientists from NCTR, CVM and other institutes, hosted by at Division of

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q2. Please identify limitations, weaknesses, or inadequacies of the bioinformatics serotype clustering; please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				and SGD/CARD to ensure that no major virulence genes have been overlooked.	Microbiology, National Center for Toxicological Research, US FDA.
8			D	The databases and methods used to determine the virulence factors to form the genetic matrix are appropriate.	No response is necessary.
9			E	I do not have sufficient expertise or experience in the use of such databases and the determination of virulence factors to scrutinize this part of the risk assessment.	No response is necessary.
c. Is the clustering algorithm accurately described, utilized, and appropriate for its intended use?					
1			A	The term “clustering algorithm” only occurs in Table 41. I guess you refer to the machine learning algorithm?	Yes, it refers to the machine learning algorithm, namely unsupervised random forest, used for classification.
2			A	I get the idea of what has been done, but do not have sufficient experience in it to comment on the details.	No response is necessary.
3	135	74-83	A	The description of the methodology is not very detailed. There is a reference to some publications that are not peer reviewed. The summary graph (Figure 36) is not explained (the caption is not at all informative). In the graph, the clustering seems to be done at the gene level, but in the risk assessment, it is done at the serotype (subspecies) level. This is confusing, please explain it better.	Clustering is accomplished using Virulence Factors (VF) and ultimately, the gene profile (presence/absence) of isolates. Once clusters are determined, serotype information is identified for each isolate (posthoc), and so, clusters can and do contain multiple serotypes (Table 48). Most serotypes were assigned to a single cluster, but some (e.g., Infantis) were assigned to two clusters. In the case of Infantis when k=4 clusters, 88% of the Infantis isolates

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q2. Please identify limitations, weaknesses, or inadequacies of the bioinformatics serotype clustering; please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
					were assigned to cluster 2 and 12% were assigned to cluster 3. This, in particular, corresponded with different patterns of VF presence/absence associated with the pESI plasmid. However, this was not an issue when k=2, as all Infantis isolates resided in cluster 2. In the case of a serovar's isolates being split between clusters, the serovar was ultimately assigned to the cluster with the most isolates for that serovar (i.e., 'best' cluster by majority). FSIS has developed *Bioinformatics Supplemental Materials* (available here) to provide additional clarity and understanding of the methodology. EpiX Analytics has also updated the caption of Figure 36 to clarify the process.
4			B	Overall concept and implantation is appropriate and useful.	To further elucidate the process, we have included a comprehensive *Bioinformatics Supplemental Materials* (available here).
5			B	The Conceptual implementation of the Random Forest and clustering is good, assuming pieces of example code provided to review, represents the overall implementation. Despite a few gaps in code (not provided Nimble implementation), the later added sample codes for serotype multipliers and uncertainty-variability draws run fine and the corresponding textual concept and algorithm explanation is clear and adequate.	

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q2. Please identify limitations, weaknesses, or inadequacies of the bioinformatics serotype clustering; please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
6	142-143	315-347	B	The latest sample code provided for multiplier estimations and outbreak definitions runs OK. However, the overall summary description of the approach used for attribution to cluster and Comparison with FoodNEt data is good, however, to understand and review the implementation piece of the mentioned text, no code was no code was provided (line 315-347)	
7			C	Random forest is a powerful machine algorithm approach that can be used for cluster analysis. The number of trees is one of the main hyperparameters in a random forest model influencing the accuracy and computational efficiency. In general, a higher number of trees can lead to better performance at the expense of efficiency. In this analysis, random forest model was conducted with 10,000 trees. There was no justification why this number of trees was selected. There is no universal rule for determining the optimal number of trees. However, it can be considered to include plots or description of findings from plotting changes in out-of-bag error rate by increasing the number of trees so that the sufficiency/insufficiency of the current choice can be demonstrated. Below is an example plot of OOB vs number of trees.	Thank you for the comment regarding random forest model hyperparameters. Undoubtedly, it is true that the number of trees plays an impactful role in the performance and predictive power of the analysis. We have added a *Bioinformatics Supplemental Materials* (available here) expanding on the rationale behind the key parameters: the number of trees and the number of features to consider at each split. The random forest was run on a supercomputer to analyze 40,000 isolates and 193 virulence factors, which completed only 10,000 trees within a 7-day allocated wall time. Rather than trialing numerous hyperparameters, given the extent of resources available, results were instead compared with clustering analysis conducted on a smaller subset of the data (50k trees and fewer isolates (12k) and VFs (182)); although,

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q2. Please identify limitations, weaknesses, or inadequacies of the bioinformatics serotype clustering; please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				 In addition, a random forest model can rank the importance of evaluated traits on classification, or virulence factors in this case. Were all 193 virulence factors used for classification? Can any virulence factors be potentially excluded to form a parsimonious model without significant changes in OOB errors from the full model? It would be helpful to share the top-ranked virulence factors in this report, as this might shed light on the potential	the project was beef-focused a similar high virulence cluster 1 arose. Furthermore, the resulting groups were also compared with other genomics-based clustering approaches (Table 13). Additionally, out-of-bag (OOB) error is not necessarily relevant with unsupervised random forests (URF) as each data row left out in the decision tree (DT) needs to have an outcome in order to calculate the overall OOB. Unlike supervised random forest (SRF) models, which consider training and test data, synthetic data for the URF case would need to be created to evaluate the performance of the model. When conducting the clustering, there was more concern with using bagged DTs with random feature selection to generate a similarity matrix. For further clarification, FSIS developed *Bioinformatics Supplemental Materials* (available here), which includes A "Virulence Factor" section. We also make a note here that as stated in the Risk Assessment, EpiX Analytics excluded putative virulence loci present in more than 95% of assemblies or which were found in fewer than 10 assemblies, which resulted in a final database of 193 loci to be used in the URF. A sensitivity analysis was not performed on the clustering algorithm

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q2. Please identify limitations, weaknesses, or inadequacies of the bioinformatics serotype clustering; please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				use of multiplex PCR methods for cluster identification in a high throughput way without linking to serovars.	which may have resulted in a more parsimonious model. The requested ranking will not be included in this report*. As a starting point, the complete rankings of VFs from an earlier analysis focused on beef and human isolates (12k isolates and 182 VFs) by EpiX Analytics is available here: https://www.medrxiv.org/content/10.1101/2022.12.13.22283417v1 *Peer reviewers were provided access to the data and underlying information for this risk assessment in accordance with the Office of Management and Budget (OMB) information quality peer review guidelines. OMB guidelines exempt the sharing of risk assessment information in circumstances where there are compelling interests, including privacy concerns, trade secrets, intellectual property rights, or other confidentiality protections)(Guidelines, Section V(3)(b)(ii)(B, 67 FR at 8460). For this reason, a small part of the work conducted in partnership with EpiX Analytics (external private sector collaborators) was not made available to the peer reviewers. Nonetheless, all the methods were fully

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q2. Please identify limitations, weaknesses, or inadequacies of the bioinformatics serotype clustering; please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
					documented in Appendix A of this FSIS risk assessment report.
8			D	Although the clustering algorithm was described and utilized well, the authors should justify some of the choices made for method selection for transparency purposes. However, please reanalyze the data by including serovars with less than 50 assemblies/isolates in the initial machine learning dataset/virulence loci matrix. Also, please check for any missing data in FoodNet and NORS that can be useful in the development of virulence matrix. Please see below for the details and specifics:	Responses to the reviewer's questions are provided below.
9	139	177–178	D	*Sample selection:* The authors mentioned that serovars with less than 50 assemblies/isolates were not included in the formation of the initial machine learning dataset/virulence loci matrix. Moreover, the matrix of virulence loci was constructed *excluding* these serovars (It is mentioned in Page 139, lines 203–209 that the final database consists of 36,647 samples and contains 193 virulence loci); however, the final supervised random forest clustering was performed *including* these serovars (Page 140, lines 231–236: This ultimately brought the number of isolates allocated to clusters to 40,038; ~4000 newly added	It is possible that by not including these "rare" isolate assemblies, influential or important virulence factors could be missed or overlooked. At the same time, however, there are imperfections in the data due to the probabilistic approach in gene annotation; namely, true genes could be missed in general and false genes could be annotated. Generally speaking, there should be less uncertainty regarding the abundant isolates compared with rare isolates. Perhaps more important than missing differences in virulence by excluding rare serovars is limiting VFs to those only coming from *Salmonella*, since some VFs are

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q2. Please identify limitations, weaknesses, or inadequacies of the bioinformatics serotype clustering; please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				assemblies/isolates). Since this accounts for approximately a tenth of the newly formed dataset, this should have been included in the construction of initial virulence matrix. It has been reported that there are genetic differences in the virulence and antimicrobial susceptibility of different serovars that genomic data can identify (Xu et al. 2021. BMC Infectious Diseases. DOI: https://doi.org/10.1186/s12879-021-06340-z; Suez et al. 2013. PloS One. DOI: https://doi.org/10.1371/journal.pone.0058449; Tsai and Coombes. 2019. Trends in Microbiology. DOI: https://doi.org/10.1016/j.tim.2019.01.004). By not including these assemblies/isolates, there is a chance that the important virulence genes might have been missed. Please reanalyze the initial virulence matrix by including the serovars with less than 50 assemblies/isolates.	commonly passed through horizontal gene transmission. The analysis performed by EpiX Analytics circumvented this by including selected *E. coli*, *Shigella* and *Yersinia* VFs in addition to *Salmonella* VFs. Future iterations should investigate modifying the lower threshold requirement of 50 isolates per serotype and other potentially informative genomic and VF data.
10	138-139	169-173	D	*Sample selection:* "We identified enterica isolates associated with human clinical cases from BioProject PRJNA230403 (CDC PulseNet). We included sporadic, domestically acquired enterica isolates from the FoodNet active surveillance network. However, we did not consider outbreak cases from FoodNet in the initial unsupervised random forest. Rather, beef-,	Unlike NORS outbreak cases, FoodNet cases are considered sporadic, although some are associated with an outbreak, as FoodNet is an active laboratory and population-based surveillance system. Thus, it is difficult to determine what specific exposure (e.g., poultry, beef, or others caused a person with a sporadic infection to become ill. Risk

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q2. Please identify limitations, weaknesses, or inadequacies of the bioinformatics serotype clustering; please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				chicken-, and turkey-attributed outbreak isolates instead came from the National Outbreak Reporting System (NORS) dataset." FoodNet was used for the sporadic cases, whereas NORS was used for the outbreak cases. Why were these two different sources used? Were the authors concerned that FoodNet did not report all of the outbreaks occur in the US? While the approach used was appropriate, please check both FoodNet and NORS for any missing data that can be useful in the development of virulence matrix.	multipliers were estimated using poultry-associated outbreaks from NORS and poultry food/food commodity isolates from FSIS regulatory sampling programs. FoodNet cases were used to corroborate that similar proportions of sporadic and outbreak cases were assigned to each cluster.
11			D	*Selection of method:* Why have the authors employed the random forest method for clustering? Please justify this choice. Random forest method is a powerful classification strategy that is commonly used for classifying labeled microbial data. This is an appropriate approach as long as labels (i.e., output variables) are available. However, in cases where, such labels are not available, other methods such as *k*-means could have also worked well to distinguish inherent patterns in the data. [Wen Nies et al. 2019, *Processes*, DOI: 10.3390/pr7090550 (https://www.mdpi.com/2227-9717/7/9/550); Lupolova et al, 2019, Microbial Genomics. DOI: 10.1099/mgen.0.000317]	Although other clustering methods may be considered, here are reasons for EpiX Analytics choice of Unsupervised Random Forest (URF): 1) Many unsupervised learning algorithms (including k-means) rely on a metric to evaluate the pairwise distance between samples, the choice of a metric may strongly impact the quality of the resulting clustering. URF computes distances between instances in unsupervised settings where the prediction task is performed by a majority vote; 2) Given that Virulence Factors (VFs) are evolving, random forests are generally

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q2. Please identify limitations, weaknesses, or inadequacies of the <u>bioinformatics serotype clustering</u>; please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				The final selection of 2 clusters (over the tested 4 clusters) is scientifically sound. However, as noted by the authors, this may need to be revised in the future based on the changes in seroprevalence in outbreaks and isolated from food and environmental samples over time (as an example, serovar Infantis may prove to be the dominant serotype, as seen over the last decade).	computationally efficient and scalable to big data, due to trees being trained independently which allows for parallelization of the algorithm; 3) URF is invariant to monotonic transformations of the input variables; 4) URF is robust to outliers due to the well-known robust property of trees. Feature selection has been shown to be an important part of high-dimensional clustering, otherwise feature noise can greatly influence the clustering result away from the desired result. We agree that clustering may need to be updated in the future due to changes in seroprevalence.
12	70 145	1620-1623 419-426	D	"The first cluster consists, generally, of the more virulent *Salmonella* serotypes; we call this grouping C1. The second cluster consists, generally, of the less virulent serotypes, although some serotypes commonly observed among human illnesses (e.g., Heidelberg, Infantis) are included in this grouping called C2."	EpiX Analytics has added clarification in the text (**Appendix A** subsection *Serovar assignment*) and now more clearly reflects the information in **Table 48.** Please note that although the cluster ordering from 1 to 4 does observe a decrease in virulence, this was not the method that assigned the cluster labels. The clustering algorithm assigns the cluster labels 1 to 4

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q2. Please identify limitations, weaknesses, or inadequacies of the <u>bioinformatics serotype clustering</u>; please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
	147-148	Table 34		"Serovar assignment for k=2, 3, and 4 clusters are provided in Table 34. The serovars composing Cluster 1 remained consistent at the three levels of k (Figure 37). When k was increased from 2 to 3, the majority (98%) of Kentucky isolates separated into their own cluster (Cluster 3). Kentucky remained on its own when k was increased to 4 and most Infantis isolates (88%) formed their own cluster. The remaining serovars comprising Cluster 2 in the k=2 designation continued to cluster together as k increased to 3 and 4. Isolates (i.e., non-serotyped) which were not assigned a serovar due to missing "O" or "H" antigens (n=26) may comprise a group of diverse serovars, which split between Cluster 1 and 2 for all levels of k based on supervised random forest." Serovar cluster assignments for k= 2, 3 and 4. There is a discrepancy between the text (page 145, lines 419-426) and the Table 34. The text suggests that in 4-cluster assignment, Kentucky remains in cluster 3 while Infantis moves to cluster 4. However, this is not reflected in the table (Table 34). The table shows that Kentucky moves to cluster 4 while Infantis moves to 3. Please clearly mention in the text that in the 4-cluster assignment, Infantis comprises cluster 3 while Kentucky	(when k=4) and risk multipliers are calculated thereafter. In the results presented here (k=2, 3, and 4), it is merely a coincidence that virulence decreases as the cluster goes from 1 to 4.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q2. Please identify limitations, weaknesses, or inadequacies of the <u>bioinformatics serotype clustering</u>; please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				moves to cluster 4. Also, please clearly mention as the cluster goes from 1 to 4, the virulence decreases.	
13			E	The clustering algorithm is relatively completely described, and should be reproducible for those with sufficient technical knowledge of the software tools available. The transparency of the impact of multiple numbers of clusters (and the fact that the ultimate choice was made by FSIS), is welcome.	No response is required.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q3. Please evaluate the two-curve dose-response model used to estimate the probability of illness for a given exposure dose of *Salmonella*, giving specific consideration to the following:					
				General Comments	
1	153	516-529	A	It is not clear to me whether you use Teunis et al data (caption fig 39 and line 517-518) or the outbreak data from cluster 1 (line 517) or both. It is stated that you fit a model to data from one data set using data from another. Do you mean you fit a model based on one data set to another data set? There are dots in the figure that are not blue. What data did you use to fit the curves? Besides, the "dose" is the mean dose, as the mean is the Poisson parameter. The term "mean" is well defined, also in the context of the dose, so I would use that instead of "intensity".	The dose-response model was derived from Teunis (2022) data on higher virulence cluster 1 strains: Enteritidis and Typhimurium. Teunis (2022) provided new and corrected data from Teunis (2010). All dots are asterisks positioned in the centroid of a blue circle; however, the radius is very small in some cases. The asterisks represent illnesses from exposures to Enteritidis and Typhimurium strains whereas the blue circles symbolically represent the overall outbreak size. The caption has been updated to clarify the graphic (i.e., stars vs blue circles).
2	25	575	A	I like the approach to use this fitted equation. Very pragmatic!	No response required.
3			A	One thing that strikes me is that much attention is given to the uncertainty dimension in the DR relation, but I could not find anywhere where it is applied. Why complicate the model if that is not fit for purpose? I would recommend to exclude the uncertainty dimension or at least explicitly clarify from the start the fact that this complicated part of the analysis is not used in the risk assessment, so the reader	The dose-response uncertainty analysis was leveraged in the final product standard model uncertainty analysis that has been added to **Chapter 5**. Also, the incorporation of uncertainty was needed to derive the most reasonable dose-response model, such as the median estimate in the uncertainty dimension.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q3. Please evaluate the two-curve dose-response model used to estimate the probability of illness for a given exposure dose of *Salmonella*, giving specific consideration to the following:					
				can decide whether (s)he thinks it is worthwhile to focus on it.	
4	144-145	389-393	A	It is informative to read that the parameter v is assumed to be constant. I am not able to judge whether that makes sense, because the reason behind it is not given. I looked up the Teunis et al. reference that is given, but could not find it there. So please explain.	This follows the approach and assumptions described in: (1) Teunis (2010) where they assumed a random offset from the transformed parameter omega = logit(u) only, but not on log(v), and (2) Thebault (2013) where a single parameter was used for log(v). This has been updated in the Assumptions Table in EpiX Analytics' report, **Appendix A**.
5			B	The idea of creating two dose-response models with the support of genomic information about virulent factors to cluster more and low virulent is great to decide the multipliers for dose-response and is very logical in the view of stricter risk management for public health with more infectious serotypes.	No response required.
6			B	Proper methods to account for all details of modeling like the Bayesian approach, bootstrapping, and sensitivity and specificity analysis have been used in the modeling process and is appropriate.	No response required.
			C	(no comment from this reviewer)	No response required.
7			D	The two-curve dose-response model used to estimate the probability of illness for a given exposure dose of *Salmonella* is an appropriate choice. However, some of the	Responses to the reviewer's questions are provided below.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q3. Please evaluate the two-curve dose-response model used to estimate the probability of illness for a given exposure dose of *Salmonella*, giving specific consideration to the following:					
				assumptions and analytical choices made while using and modifying this model need further justifications and reevaluations (e.g., scaling factor for dose response model for Cluster 2 based on Cluster 1, illness given infection, serotype switching, assumption that all products have same growth and inactivation kinetics, and not considering the geographic differences while developing dose-response model with genomic data). For details and specifics, please see below Q3 a, b, and c.	
8			E	While there is relatively complete characterization of the overall approach to estimating the dose-response, there is insufficient transparency as to the distribution of doses that will be ultimately used for the ultimate risk characterization. For example, the results of the exposure assessment component (i.e., the distribution of average doses associated with servings) should be shown as a distribution, overlaid by the resulting dose-response curve. This will clearly show which portion of the dose-response curve is critically important to the ultimate conclusions of the risk assessment for these products.	By definition, all microbial food-safety applications fall in the category of rare events. Therefore, the exposure distribution will be such that the majority of the mass of the distribution falls in the visually linear (in log space) portion of the dose-response function. Similarly, the visually linear portion of the exposure distribution will coincide with the region of the dose-response model below the first inflection point. FSIS believes the only meaningful visual comparison for food safety applications would be a comparison of different pathogens. For example, a comparison of *Listeria monocytogenes* and *Salmonella* would be visually different because the exposure distribution would be left-shifted, relative to *Salmonella*, and the dose-response would be right shifted.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q3. Please evaluate the two-curve dose-response model used to estimate the probability of illness for a given exposure dose of *Salmonella*, giving specific consideration to the following:					
					The requested figure was not added to the risk assessment because such a comparison would be beyond the scope of the risk assessment.
a. Was the use and modification of the Teunis beta-Poisson model appropriate to describe probability of illness due to *Salmonella* serotypes that differ in virulence? If not, what other models should be considered? Please provide the reference(s) if applicable.					
1	143-145	348-407	A	The use and modification of the Teunis model seems appropriate. The same approach is used as in the referenced peer reviewed papers, for which, to my knowledge, no good alternatives are available. I support the approach taken	No response required.
2			A	Still, some consideration should be given to the fact that the DR model is based on outbreak data. Outbreaks typically occur for more virulent strains, so strains that do not cause any outbreaks, are not taken into account in the analysis. This might lead to an overestimation of the risk of illness. For Campylobacter, for example, Teunis et al. (2018 Epidemics 24, 1-20.) show a big difference between data from outbreaks and challenge studies. One can wonder whether that implies that DR models based on outbreaks overestimate the risks? This should be discussed in the report	This is a valid point. If outbreak strains are more infective, one would expect a difference between outbreak and challenge data. Teunis (2022) found a lower infectivity challenge studies, but challenge studies typically involve young, healthy subjects and the same strain, reducing the variability in the results. We believe that outbreak data provides more realistic infectivity information compared to challenge data as it better represents the variability in strains, individuals, and the consumed dose. Ultimately, for this risk assessment the relative difference in the dose-response functions between the higher virulence cluster 1 and

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q3. Please evaluate the two-curve dose-response model used to estimate the probability of illness for a given exposure dose of *Salmonella*, giving specific consideration to the following:					
					lower virulence cluster 2 serotypes is more important than the slope of the dose-response. Nevertheless, additional details on the strains comprising the outbreak data considered is presented in Chapter 2 to improve transparency in the outbreak proportion derivation of the risk multiplier.
3			B	The Teunis beta-Poisson model the 2010 and the latest update in 2022 are the most efficient methods to model the dose response and predict the probability of illness.	No response required.
4			B	All relevant references are included, and methods are implemented appropriately.	No response required.
5	158		B	All assumptions made during the modeling and use of parameters are very logically enumerated in Table 41	No response required.
6	143	355	C	Yes, the reviewer agrees that Teunis beta-Poisson model is appropriate for describing the DR relationship, with some modifications. The description stating that the probability of illness given infection is 1 should be removed as it can be confusing. This assumption is not valid as the morbidity to infection ratio was estimated much lower than 1 based on the estimated P(ill) and P(inf) in Teunis' studies (Teunis 2022, 10.1016/j.epidem.2022.100653).	As suggested by the reviewer, text clarifying that a model was used with illness as the outcome, rather than assuming an infection:illness ratio of 1, has been added to the text (**Appendix A**).

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q3. Please evaluate the two-curve dose-response model used to estimate the probability of illness for a given exposure dose of *Salmonella*, giving specific consideration to the following:					
				Additionally, it appears that the same morbidity to infection ratio of 1 was used for both cluster 1 and 2, which is not a valid assumption. Instead, it is recommended clarifying that a model was used with illness as the outcome since there was no data available on the number infected individuals in outbreak data.	
7			D	The use of the Teunis beta-Poisson model, which is the comprehensive published model for *Salmonella* dose-response has been well-justified for the purpose of this assessment. As the analysis and inclusion of genomic data in risk assessment are gaining attention, other researchers are also putting efforts in this area. We inform the authors of a different method, for example, Karanth & Pradhan (Risk Analysis, 2022; DOI: 10.1111/risa.13924), also attempted to include genomic data in a dose-response framework (without clustering according to serotype). However, for the current scope of the risk assessment, the use and modification of the Teunis beta-Poisson model is appropriate for serotypes clustered according to virulence.	The approach implemented by EpiX Analytics aims to construct reasonable dose-response models comprising a wider range of serotypes in poultry; however, improving the resolution is an important step for future risk analyses.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q3. Please evaluate the two-curve dose-response model used to estimate the probability of illness for a given exposure dose of *Salmonella*, giving specific consideration to the following:					
8			E	The use of a very recent and peer-reviewed dose-response assessment from a top scientific team in this domain (i.e., the Teunis model) is a very reasonable foundation for the overall dose-response method. However, as discussed above under Q1 “Overall Comment”, for the purposes of this risk assessment an overall dose-response curve (estimating Pill for average doses up to 106 cfu/g, for example) is strictly not required, it is ultimately not applicable.	No response required. The comment is addressed above in Q1 “Overall Comment”.
b. What (if any) other data sources and methods should have been used in the *Salmonella* dose-response model risk multipliers? If not, what other data sources and/or methods should be used? Please provide the reference(s) if applicable.					
1			A	I would not know any other data sources or methods that should have been used as alternatives.	No response required.
2			B	The risk multiplier calculation includes the latest usable, clean, and extensive data available from the FSIS sampling program and NORS databases, to feed into the calculation of risk multipliers.	No response required.
3			B	The methodology used is conceptually correct, and the provided example code for multiplier estimation is in agreement with the conceptual description of the model and runs fine.	No response required.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q3. Please evaluate the two-curve dose-response model used to estimate the probability of illness for a given exposure dose of *Salmonella*, giving specific consideration to the following:					
4			C	Although the present method assumed the same risk multiplier shared between serovars categorized in the same cluster, it is worth estimating the risk multipliers for the most frequent serovars in each cluster separately. This could allow for a relative comparison between serovars within each cluster regarding their infectivity/pathogenicity, which could be used to evaluate the validity of the risk multiplier approach by comparing with the relative infectivity/pathogenicity ranking learned from other empirical/simulation studies.	Epidemiological data does not exist at present that would allow resolution down to the level of a risk multiplier for each serotype.
5			D	The data sources employed for the development of the analytical dataset (2 clusters; 193 virulence loci) are complete, barring the concerns cited in the previous question # 2, parts a & c.	No response required.
6			E	See comments in response to Q1 which suggest approaches to exploit the rare-event nature of the contamination and the possibility of much more simplified approaches. This would also appear to be potentially (and possibly more so) applicable to many "usually well-cooked" (e.g., substantially attenuated) products in FSIS's mandate. This is reinforced by Ebel and Williams,	FSIS has included a simplified approximation method to compare with the initial more complex model in **section 5.3 subsection Techniques for Approximation**.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q3. Please evaluate the two-curve dose-response model used to estimate the probability of illness for a given exposure dose of *Salmonella*, giving specific consideration to the following:					
				2015, Table 1, which demonstrates that the median of the average dose distribution for E. coli/Beef is 10-11.7 cfu/g and for Campylobacter/Chicken is 10-9.3 cfu/g (based on the sum of utest and uatten). The values for the median of the average dose distribution are based on adding μtest = –7.69 and μatten = –4.05 to get μserving = –11.74 (i.e., the mean and median on the log-scale) for E.coli/Beef, and μtest = –1.68 and μatten = –7.70 to get μserving = –9.38 for Chicken/Campylobacter. The values for μtest and μatten come from the third and last rows of Table 1 of Ebel and Williams, 2015, respectively.	
c. Is the use of the two-curve dose-response model appropriately used to estimate illness estimates? If not, what other approach could have been used with this dose-response model? Please provide the reference(s) if applicable.					
1			A	As explained below, I struggle with the question, as the dose response model is only used for illness estimates that are used to calculate the risk reduction obtained by control measures, where cluster 1 and cluster 2 serotypes are differentiated. In principle, this is a good approach, as, when it works, it gives the opportunity to introduce more efficient standards, targeted at the most virulent strains. Technically, my judgement is that the two-curve DR model	A sensitivity analysis has been conducted to evaluate the role of adjusting c (proportion of higher virulence C1 *Salmonella*) across the range [0,1] as well as utilizing the lower and upper bounds of the two-curve dose-response models.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q3. Please evaluate the two-curve dose-response model used to estimate the probability of illness for a given exposure dose of *Salmonella*, giving specific consideration to the following:					
				is well developed and appropriate to use. But given the uncertainties in the identification of the strains that are actually present on the meat, I am not sure the usage is actually appropriate. This should be discussed.	
2	190	615	A	When you search the document for “illness estimates”, there is reference to the CDC estimate I/N. This does not use the dose response relation. This may seem a weird comment, but it illustrates that it is not particularly well explained how (and for what purpose) the DR model is actually used. (And, by the way, the term “two-curve” does not occur in the document either.)	We have enhanced the *dose-response modeling* subsection in **Chapter 5** in the report to improve the clarity on how the dose-response model is used.
3	67-68		A	The dose response model is not used for the baseline probability of illness estimate. This implies either that the risk assessment is done without hazard characterization, or that what is written here is an alternative hazard characterization, without using a dose-response model. This should be made explicitly clear and additionally it should be made explicitly clear why you do derive a DR model then. I understand this is for the purpose of the evaluation of the control measures, but I am only able to read that	A section describing the "**Descriptive Estimates of Risk per Serving**" as described by the dose-response model was added to the document (**Section 4.4**) making it clear that the derivation of the dose-response model was for the purposes of evaluation of control measures. The baseline probability of illness estimates serves as the empirical, rather derived, hazard characterization.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q3. Please evaluate the two-curve dose-response model used to estimate the probability of illness for a given exposure dose of *Salmonella*, giving specific consideration to the following:					
				between the lines. It should be clarified and discussed	
4	25	578-586 601-640	A	The dose response model is used for the serotype-based final product standards, and for intervention at the receiving step	A sensitivity analysis (**section 5.3**) has been added to the document outlining the effect of alternate dose-response models on the illnesses avoided estimates.
5			B	The conceptual and R code implementation of the two-curve dose-response model is appropriate and reproducible.	No response required.
6			B	The variability and uncertainty estimation script is adequate in terms of the conceptual implementation of the proposed model in the report. The steps followed in the script to implement the textual concepts read OK. However, the script could not be tested by running because of the unavailability of an input file containing uncertainty estimates generated from the nimble script provided by the external collaborator, due to some copyright issues.	No response required. Peer reviewers were provided access to the data and underlying information for this risk assessment in accordance with the Office of Management and Budget (OMB) information quality peer review guidelines. OMB guidelines exempt the sharing of risk assessment information in circumstances where there are compelling interests, including privacy concerns, trade secrets, intellectual property rights, or other confidentiality protections)(Guidelines, Section V(3)(b)(ii)(B, 67 FR at 8460). For this reason, a small part of the work conducted in partnership with EpiX Analytics (external private sector collaborators) was not made available to the peer reviewers. Nonetheless, all the methods were fully

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q3. Please evaluate the two-curve dose-response model used to estimate the probability of illness for a given exposure dose of *Salmonella*, giving specific consideration to the following:					
					documented in Appendix A of this FSIS risk assessment report.
7			C	Yes.	No response required.
8			D	The two-curve dose-response model is an appropriate choice. However, some of the assumptions and analytical choices made while developing this model should be justified further. For specific comments, please see below.	Responses to the reviewer's questions are provided below.
9	139	191	D	The Pathosystems Resource Integration Center (PATRIC) database has been recently renamed to Bacterial and Viral Bioinformatics Resource Center (BV-BRC).	Thank you for this correction.
10	140 143	Footnote 4 355–356	D	Illness and illness given infection cannot be interchangeably used, since not all infections with *Salmonella* lead to illnesses (Teunis et al. (2010), *IJFM*, DOI: 10.1016/j.ijfoodmicro.2010.09.026). The authors' rationale has not been sufficiently justified. Since the authors have access to exposure estimates, this reviewer suggests the development of separate infection and illness given infection curves consistent with Teunis et al. (2010) rather than assuming illness is equal to infection.	Indeed, not all infections lead to illnesses. EpiX Analytics' report now includes text clarifying that the beta-Poisson model directly links *Salmonella* exposure to illnesses, in contrast to Teunis (2008). Extensions to include separate infection and illness given infection curves should be considered in the future, although preliminary tests by EpiX Analytics resulted in an overparameterized model due to the lack of sufficient data regarding the number of infections associated with each Enteritidis and Typhimurium outbreak.

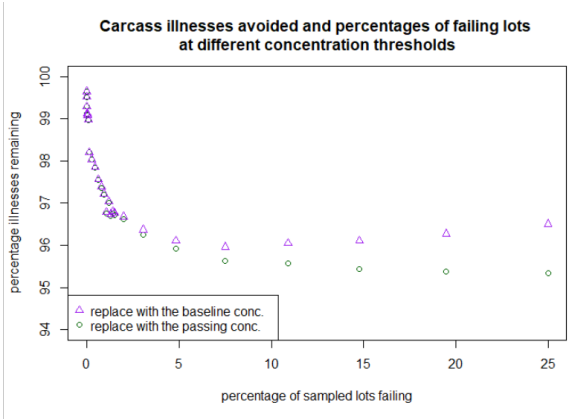
Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q3. Please evaluate the two-curve dose-response model used to estimate the probability of illness for a given exposure dose of *Salmonella*, giving specific consideration to the following:					
11	141 142	264–273 309–314	D	The authors mentioned testing two modes of assignment of cluster to isolates/assemblies that have not been annotated – best cluster and proportion cluster. Both methods are scientifically sound. However, the authors did not clearly state whether they used both or one versus the other and justification for the choice. Please provide more clarification.	The baseline risk multipliers were estimated using the proportion cluster approach. However, since the majority of serotypes clustered together in the k=2 clustering scenario (**Table 17**), the difference between using proportion cluster and best cluster was negligible (**Table 20**). A description of the baseline scenario is provided as a footnote for **Table 44** (Sensitivity of risk multipliers to different modeling and data transformation options). Text (lines 472-475) describing the baseline scenario has also been added immediately prior to the multiplier tables.
12	142	288	D	The non-parametric bootstrap model to account for uncertainty should be described in further detail for clarity and transparency.	Uncertainty was incorporated into the *Salmonella* in poultry and outbreak case estimation by cluster. The non-parametric bootstrap approach briefly mentioned in the report considers randomly sampling with replacement the FSIS poultry samples (for the proportion in poultry case; i.e., denominator of the risk multiplier) or the curated list of poultry-attributed outbreaks from NORS (with additional components such as underreporting factors, recency, and various allocations of outbreaks to foods for the proportion in outbreak case; i.e., numerator of the risk multiplier). For each bootstrap sample, the random selections were summed to calculate the proportions by cluster and determine the 95% confidence intervals described in **Tables 18-19**.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q3. Please evaluate the two-curve dose-response model used to estimate the probability of illness for a given exposure dose of *Salmonella*, giving specific consideration to the following:					
13	144	374	D	Equation - In extrapolating the dose-response of Cluster 2 from that of Cluster 1, the authors proposed using a factor where the relative risk (RR) for cluster 2 is being divided by that of cluster 1 (RR1).It is just the use of a scaling factor (i.e., RR2/RR1). It is mentioned "The DR model for Cluster 1 (including Enteritidis and Typhimurium) was developed from outbreak data associated to these serovars"- Page 143, lines 350-351. Why not the same procedure that was used to develop the dose-response (DR) model for Cluster 1 was used in the development of the DR model for Cluster 2 (e.g., Infantis, Kentucky)? The DR model for Cluster 2 would have been developed using outbreak data associated with these serovars for Cluster 2 (e.g., Infantis, Kentucky) rather than using a scaling factor to the DR model for Cluster 1. Please provide explanation for this choice and compare the analysis and results from both methods (1) currently used scaling factor, and (2) using outbreak data associated with Cluster 2.	Epix Analytics did not use the same procedure for the dose-response model for cluster 2 as for cluster 1 as there is more robust data on Enteritidis and Typhimurium. In addition, cluster 1 is comprised of a small, select group of serovars. Cluster 2 has a wide-range of serotypes and would be more heavily skewed to Infantis and Kentucky by deriving the dose-response model in the same fashion. Furthermore, given the lower virulence of serovars in cluster 2, outbreak data is less abundant (**Chapter 2**), and often do not report dose consumed. For example, no poultry-attributed outbreaks of Kentucky appear in the CDC outbreak data.
14	145	406–407	D	The authors mentioned fitting a polynomial model on the initial DR model. This reviewer agrees with the choice made, as the linear models are easy to understand and	A polynomial was fit to the dose-response models with 95% credible intervals. This generalization assists with portability and

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q3. Please evaluate the two-curve dose-response model used to estimate the probability of illness for a given exposure dose of *Salmonella*, giving specific consideration to the following:					
				interpret. But the explanation and rationale provided here for this choice is unclear and incomplete. Please provide the explanation.	efficiency of implementation without loss of detail/information.
15	148	Table	D	The virulence of the 4 clusters relative to each other is not clearly delineated. Although the reader can eventually infer that cluster 4 is less infectious than 3 and so on, this needs to be clearly and prominently mentioned.	FSIS has added an additional explanation highlighting cluster construction and notating of decreasing risk in the **Chapter 2**. Please note that in the results presented (k=2, 3, and 4), it is merely a coincidence that virulence decreases as the cluster goes from 1 to 4.
16	149	440–445	D	The authors mentioned that two serotypes (Berta and Saintpaul) switched serotypes during bootstrapping. Table 34, Pages 147-148 indicated that Berta and Saintpaul were retained in Cluster 1 irrespective of the results of bootstrapping, which indicated the switch to Cluster 2. Was any change made to the cluster assignment of these serotypes to account for this? Please compare the results with and without the switch.	Serotype switching analysis was conducted to assist in assessing the stability of the clusters generated from the random forest algorithm. Isolate switching was rare except in these two serotypes. To account for situations such as these, best cluster and proportion cluster weights were explored in the subsequent risk multiplier estimation, which did not yield any significant differences overall.
17	159	Table	D	Assumption 7: *Salmonella* inactivation and growth are not product-specific. This is not appropriate, as *Salmonella* inactivation and growth can be product-specific (for example, see Table 1 in Silva & Gibbs (2012), *FRI,* DOI: 10.1016/j.foodres.2011.06.018). Also, another risk assessment for *Listeria monocytogenes* used different kinetic parameters for different types/sub-	The clustering and dose-response models were developed on the aggregated product and commodities (chicken, turkey, carcasses, parts, and comminuted). This is a simplifying assumption to capture the overarching *Salmonella* inactivation/growth. FSIS agrees that product-specific inactivation and growth would be ideal. However, reducing to more product-specific behavior at this stage could

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q3. Please evaluate the two-curve dose-response model used to estimate the probability of illness for a given exposure dose of *Salmonella*, giving specific consideration to the following:					
				categories of deli meats (e.g., ham, turkey, and roast beef) (Pradhan et al. 2009. Journal of Food Protection, DOI: 10.4315/0362-028x-72.5.978). Please consider the use of different kinetic parameters for different products.	potentially result in contradictory dose-response models for each cluster by product.
18	159	Table	D	Assumption 8: Although the Teunis models are the comprehensive dose-response models currently available, they are based on primarily European data. While these can be extrapolated to the United States, the resultant curves may be marginally different when considering the dose-response models with genomic data. Please recognize this difference and the uncertainty associated with it. Studies have shown that the genomic signatures of *Salmonella*, particularly in antibiotic resistance patterns, differ with the geographic regions both among different countries (US, Europe, Africa, and China), (Cao et al. 2023. Scientific Reports. https://doi.org/10.1038/s41598-022-24150-4) and within a country (Carroll et al. 2017. Applied and Environmental Microbiology. https://doi.org/10.1128/AEM.00140-17).	Thank you for the comment and supporting materials. Text has been incorporated into Assumption 8 of the Assumptions Table in the EpiX Analytics' report, **Appendix A.**
19			E	While it is not possible to conclude that the dose-response models are not "appropriately used", the level of complexity	A *techniques for approximation* section was added to **Chapter 5 Final Product Standards** exploring the effect of this simplification.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q3. Please evaluate the two-curve dose-response model used to estimate the probability of illness for a given exposure dose of *Salmonella*, giving specific consideration to the following:					
				of the analysis may simply not be proportionate to the level of data on which it is based. As discussed above, the sheer number of microbes in raw products, before and after intervention, and the mean probability of illness for single cfu exposures might be sufficient to answer the risk management questions. This would need to be confirmed with far more analysis than is allowed for within the time/resources available for peer review.	

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				General Comments	
1	70	1601 -	A	The approach is selected because FSIS believes in it. Please explain the origin of that belief and provide appropriate references, describe clearly why it is chosen.	The rational for the approach has been added to the intro of **Chapter 5 Final Product Standards**, including appropriate references.
2			A	I support the approach used, but I miss references to similar studies, such as Nauta et al. 2012 (http://dx.doi.org/10.1016/j.ijfoodmicro.2012.07.018) Please explain how the approach used (dis-)agrees with theirs ?, to clarify the choices made in selection of the methodology.  Nauta et al.'s Figure 4 is giving the relation between the fraction illnesses avoided and the	Thank you for taking the time to validate our results.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				fraction non-compliant lots for several thresholds and is comparable with Figures 16 and 17. I used the R model provided to make a new graph similar to Nauta et al. 2012, which shows a kind of ROC curve, informing the risk manager on the % of failing lots (that will give some monetary loss) vs. the relative health benefit, shown below. It can be helpful for decision makers and if you want to compare approaches. An interesting difference between this graph and the one provided by Nauta et al. 2012 is that never more than 5% of illnesses is prevented here, whereas there is goes up to 100%. That is due to the assumption of Nauta et al. that all lots are tested. That is not realistic and the approach used here is more informative.	
3	92-92	2146-2184	A	I do not understand what you are doing here. Table 27 is not explained well. What are the numbers in parentheses? How do I derive the extent of Cluster 1 serotypes from the table? What do you mean by the concordance rates? What are you actually doing in lines 2157-2164? And in lines 2180-2184?	This section was rewritten based on peer review comments to improve clarity and more clearly define and outline the approach.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
4	95	Table 29	A	What is in the table? All I understand from the caption is that this is a contingency table. What are the numbers?	Further detail was added to explain Table 29 in the section that was rewritten, including variable definitions and details of the table construction.
5	98	Figure 22	A	The figure caption is not sufficiently informative. What is the boxplot (caption says average rate)? Why is the order on the x-axis not the same as in Table 30? How does this Figure relate to Table 30?	The table has been corrected with additional details to correspond to figure. The figure axis was re-ordered as well.
6	101	Table 30	A	Please provide a more informative table heading: what does the table present? What is F_C1 without diversion?	More informative table headings were added.
7	101-102	2307-2313	A	Here the risk assessors advise on risk management. I would say that is inappropriate. Clearly, risk managers should consider the financial impacts of any intervention measure, and be aware of the uncertainty in the analysis, but the conclusions drawn from this are up to the risk managers. As no cost analysis is done (and probably is not part of the mandate), you should not draw conclusions related to cost.	This sentence was removed.
8			B	Overall, the modeling approach and assumptions made to model the data are appropriate. The use of the Bayes method to model the mixture serotypes identified to be clustered in C1 (high-virulent) and C2 (less virulent) is	No response required.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				adequate, to depict the impact of this serotype-based dose-response modeling and intervention to minimize the impact on public health due to contamination of high-virulent serotypes in the chicken products.	
9			B	The current analysis only models the public health impact of this intervention as illnesses avoided, as future step this work can be extended to include the Direct intervention costs or DALY estimates can be used to better represent the public health impacts (Havelaar AH, Mangen MJ, de Koeijer AA, Bogaardt MJ, Evers EG, Jacobs-Reitsma WF, van Pelt W, Wagenaar JA, de Wit GA, van der Zee H, Nauta MJ. Effectiveness and efficiency of controlling Campylobacter on broiler chicken meat. Risk Anal. 2007 Aug;27(4):831-44. doi: 10.1111/j.1539-6924.2007.00926.x. PMID: 17958495.	The costs of these interventions will be outlined in the FSIS cost-benefit analysis, also conducted in support of rulemaking.
10			B	Some discrepancies observed in the description of the report, that need attention are enumerated as follows:	Responses to the reviewer's questions are provided below.
11	75	1770	B	Sentence: "alpha=n/L = (i.e., if all units are tested, then all failing units will be diverted)", does not seem right, should it be (if all tested units are failed, then all failing units will be diverted)??	The parenthetical was corrected by adding the word "failing" in the definition of n: "is the total number of failing units tested per year".

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
12	76	1778	B	In the sentence "h is the share of C1 *Salmonella* among failing units in which C1 was detected", should it be …C1 was detected at *post-chill?* And similar to the next sentence in the definition of k? Please add the word 'post-chill' at the end where C1 / C2 is detected in corresponding definitions of h and k, for clarity.	The word "post-chill" was added for clarity.
13	77	1782	B	The use of alpha in the sentence "criterion and C1 is detected (alpha) and" is not very clear here given the definition of the alpha in line 1768. Please clarify.	The definition of alpha was modified in line 1768 to improve clarity.
14		1774, 1778, 1999, 2102-2103, 2106	B	Please clarify the difference between terminology: two seroclusters (C1 and C2) - clear, C1 *Salmonella*- not clear, C1 unit- not clear. Especially, the term C1 *Salmonella* has been loosely used making it difficult to be distinguished between C1 *Salmonella* and C1	A C1 *Salmonella* is a Salmonella of any serotype that is clustered in C1. A C1 unit, as described in the document, is a unit (or lot) that test positive for a C1 *Salmonella*. C1 describes the entire cluster. C1 *Salmonella* describe members of C1.
			C	(no comment from this reviewer)	No response required.
15			D	The scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final	No response required. The attenuation multiplier is discussed in Q4 (e) below.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				chicken products are reasonable. However there are some concerns regarding the use of attenuation multiplier (please see below Q4 (e)).	
			E	(no comment from this reviewer)	No response required.
a. Is the scenario analysis technique accurately described, utilized, and appropriate for its intended use (i.e., evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products)?					
1			A	Yes, the scenario analysis technique is well described and appropriately applied	No response required.
2	75	1722-1736	A	The model parameters are only explained from line 1749 onwards. It is easier to follow if they are defined immediately.	A table of model parameters has been added.
3	75	1741	A	Explain/Justify why c=0.2	Explanation was added to the line question, summarizing this finding from Chapter 3 *Salmonella* Microbial Profile: "Data from the 2022 FSIS Exploratory Sampling program indicate post-chill serocluster proportions as roughly 0.2 in cluster 1, and correspondingly, 0.8 in cluster 2."
4	75	1742	A	How is the lognormal distribution defined? It can be done in different ways.. So please give a clear definition	Full definitions of method details are provided in Appendix C: Theory. The standard definition was used.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
5	75	1742-1743	A	This is just the exposure distribution. Why write it differently here? That is confusing. Please refer to exposure assessment here.	The sentence was rephrased to make explicit reference to the exposure assessment.
6			B	Scenario analysis using the probability framework and exposure dose estimation with monte Carlo simulation is adequate to model the illness via high virulent cluster C1 and through scheme 1.	No response required.
7			B	Overall result as estimates of diverted lots for C1 serotypes and its protecting impact on reducing the illness % is legitimate.	No response required.
8	24 75	560	C	The executive summary mentioned that "the public health impact of chicken carcass final product standards encompasses the illness estimates for all secondary chicken products." However, it is not entirely clear how the illnesses associated with parts and comminuted products were included, given that these secondary products are primarily fabricated or further processed from carcasses. To comprehensively consider the secondary effect, it is crucial to describe the connection between the reduction of *Salmonella* on carcasses and the reduction in parts and comminuted products. Unfortunately, this relationship was not outlined.	A description of the relationship between carcasses and parts and comminuted products was added to **Table 7:** Risk assessment information and assumptions, the conceptual diagram, and a clear description of it was added to the **Chapter 5 Final Product Standards**.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
9			D	The scenario analysis was appropriately described and utilized within the current scope of work.	No response required.
10	70	1599-1601	D	"A major assumption of this modeling approach is that consumer demand for raw chicken products will be met by the industry, so every removed lot will be replaced by another lot in the aggregate." While this is a reasonable and valid assumption, please provide evidence to support this.	Evidence was added to the introduction of **Chapter 5 Final Product Standards**. The text that begins: "This assumption is considered reasonable because of the high consumer demand for prepared chicken products" outlines the evidence that supports this assumption.
11	70 78	1610-1611 1809-1811	D	"The serotype model is based on the available FSIS two-point chicken carcass data and, as such, cannot be used for parts and comminuted product where two-point data is not available." "Because of the absence of paired sampling data and complications in theoretic comparisons, this serotype model cannot be used for chicken parts or comminuted chicken performance standards." Please emphasize this point that serotype model is only for carcass and not for parts and comminuted product.	This point was added to the table of assumptions, highlighted in the executive summary, and discussed in **Chapter 8 Discussion**.
12			E	The overall approach seems appropriate to its intended use in answering the risk management questions.	No response required.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
b. Are the data analyses appropriate and R language source code accurate for the aims of the study?					
1			A	I could not identify anything inappropriate here, the R program is well structured and seems to work well.	No response required. Thank you.
2			B	The provided sample script 'finProdStds_Jan25_23.R' correctly implements the proposed model, as described on pages 71-81.	No response required.
3			B	Also, additionally provided, supplementary sample codes and data sets, complements the support for the model accuracy and detailed model implementation.	No response required.
4			B	The conceptual approach in the 'Receiving guidelines' module, using the detailed Bayesian posteriors is appropriate that has been implemented via Monte Carlo simulations, with distribution mixtures accounting for uncertainty. Overall approach and modeling as described in the text is adequate, however, no modeling script or any data was shared for reviewing the implementation of modeling and results presented for the 'Receiving guidelines' section.	Additional script has been provided describing the implementation of the modeling for the Receiving Guidelines Chapter.
5			C	The R codes were developed in line with the method description.	No response required.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
6			D	Data analyses and R language code are appropriate given the aims and scope of the study.	No response required.
7			E	The R code is clear, and relatively well documented.	No response required.
8			E	As otherwise noted, the Poisson-Lognormal nature of microbiological sampling is not clearly evident. As example, the following line of code: pass.init <- 10^rtruncnorm(iter,a=-Inf,b=log10(conc.thresh[k]),mu.init,sig.init) does not include the possibility that a lot may pass or fail randomly due to the Poisson nature of a random sample. This line of code applies only the lognormal component of the sampling process. While missing, It is not clear whether this is an important part of the risk assessment. The use of both classical numerical integration (e.g., R function “integrate”) and Monte Carlo simulation is noteworthy. Some explanation of why these two techniques were used, for the purposes that they are used, would be helpful to understand the overall approach. In addition, some indication that 1 million iterations (“iter <- 1000000”) are sufficient to estimate the impact of threshold concentrations would be	As stated in the intro to **Chapter 5 Final Product Standards**: “All public health outcome predictions presented in this chapter are based on a determination of pass/fail status of each lot using a test with high accuracy.” The public health analysis is not intended to evaluate test performance characteristics. Although the Poisson nature of sampling results was addressed when fitting the contamination distributions, it is not considered in this model. In addition, misclassification of enumeration from sampling is not part of this model. Application of a Poisson distribution in this part of the model – especially at the lower limit of detections – would only serve to introduce noise that would necessitate more Monte Carlo iterations (we ran 100 million to get reasonably stable results across the full range of concentration thresholds) without providing any additional insight (i.e., some lots just below the LOD might fail and some lots just above the LOD might pass). Furthermore, as

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				appropriate, given the conclusions rely on this estimate.	explained in the "Accuracy of quantitative PCR methods" section, the current method for accurately quantifying samples is a much more relevant concern than the Poisson variability associated with sampling. As identified in the introduction, this model intends to assess the direct effects of the risk management options after applying highly accurate testing techniques.
c. The definition of product lots is based on the sampling frequency of the data. Are the methods used to describe the contamination of those lot from samples appropriate, and if not, what other approach should have been taken?					
1	Section 3.2		A	I am not sure I understand the charge question (c), but I assume it refers to this section.	Yes, the question refers to **section 3.2** and the full methods outlined in Appendix C.
2	46	Table 11	A	Please clarify how the implied prevalence is calculated.	Specific text to define the implied prevalence was added in the text preceding the **Table 19**: "implied prevalence, which is defined as the mass of the lognormal distribution above the limit of detection of the assay."
3	48	Figure 9	A	Please explain in the figure caption what exactly I see in this graph.	Figure caption was updated.
4			A	The methods described in Section 3.2 seem appropriate.	No response required.
5			B	The definition of lots for different chicken types is appropriate and different assumptions made for the modeling approach to describe the lot	No response required.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				contamination, pass/fail status, and the corresponding probability of Serotypes from C1 or C1 cluster or from Scheme 1 or Scheme 2 is adequate.	
6			C	Yes, the lot definition appears to be aligned with the sampling structure where data were collected.	No response required.
7	171-172	167-173	D	Data Analysis-Flock Size: "While it is not possible to correct the data or remove erroneous entries, the mean flock size, and its variability, were estimated using the following logic. We assume that the majority of entries for an establishment are accurate and developed a list of all flocks and their size. The influence of outliers is mitigated by determining the median reported flock size as well as the flock sizes representing the 40th and 60th quantiles of the distribution. A log10 transform was applied to these values and the parameters of a normal 171 distribution are estimated using rriskDistributions library in R." While it seems a nice approach, what is the rationale of using 40th and 60th quantiles? Why not 25th and 75th or other percentiles? The later may be more appropriate as it captures more range compared to the one used in the analysis.	The 40th and 60th percentiles were chosen because at these cut-off values there were no observations in the dataset that appeared to be outliers.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
8			E	The treatment of sample information appears to be appropriate. The only exception may be the treatment of the Poisson part of the Poisson-Lognormal nature of the sampling process, consistently throughout the analysis.	This has been addressed in the reviewer's other comments.
d. Is the assumption that multiple serotypes are present within flocks appropriate and how else can the mixture of serotypes (i.e., "serotype scheme") be described?					
1			A	There is no direct evidence that this assumption is correct, although the circumstantial evidence indirectly shows that it is very likely (uncertainty dimension) that many flocks (relative frequency, variability dimension) have varying ratios of serotypes in a mixture (another variability dimension). The data is insufficient to fully characterize this uncertainty and variability, so it makes sense that a simplifying assumption is made, and the one that is actually made seems appropriate to me. But the impact of the assumption could be analyzed in more detail, e.g. by a scenario analysis involving different feasible assumptions.	A sensitivity analysis was added that models different flock serotype mixture scenarios. The finding that there is very little change in the proportional reduction in illnesses across the range of concentration thresholds further supports the use of this assumption as realistic.
2	73	1680-1682	A	I do not understand what is meant here, although it seems there is much relevant information in this sentence as it deals with the actual applicability of the overall approach. Please expand and rephrase.	The sentence was rephrased and expanded to emphasize the importance of the serotype approach.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
3	Section 5.1		A	I understand the approach taken and think that, in terms of the methodology chosen, it is fit for purpose. But the assumption of two schemes is questionable: it would be quite likely that there are more. I think it should be discussed how probable it is that there are more than two, and how that would impact the results? I would recommend to do a scenario analysis involving different feasible assumptions.	While it may be likely that there are more than two schemes, it is beyond the descriptive power of a two-point sampling by two serocluster data set. For example, our current approach requires estimation of C1 frequency in each of two schemes and the frequency of the schemes. In **Table 25**, we have 4 equations and 3 unknowns. But, adding just one more scheme converts the algebra to 4 equations with 5 unknowns. Given the sensitivity analysis concluded that changes to the cluster scheme make little impact on illness estimates, to extend the assumption past two schemes is unwarranted and lacks an empirical grounding.
4			A	I would be helpful to see a table with the final estimates of P(S1), P(S2), P(C1|S1) etc., preferably with the attending uncertainties. It is not easy to follow the reasoning, which partly goes in strict mathematical notation without and numbers, and partly in number, without any overview of numerical results.	A table providing the final estimates has been provided in **section 5.2** to clarify.
5	72-75	1670-1721	A	When contemplating about the approach, I developed the following line of thought. (It may be that you are actually reasoning likewise, but then I missed that, which may suggest you could communicate it more clearly):	The calculation provided by the reviewer in their comment is correct. The relative risk value on which the dose-response model is derived (5.66) describes the difference in pathogenicity of the two serotypes

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				There are two types of (sampled) flocks defined: S1 and S2. (line 1683-1687) These flocks are all contaminated with strains from both clusters, and so are the product lots derived from them. I assume the final exposure is just to a strain (or strains) from one cluster. The relative frequency of S1 is 0.146, that of S2 is 0.854. These numbers do not change from rehang to exposure. Within S1, a fraction 0.875 is C1 and a fraction 0.125 is C2. Within S2, a fraction 0.09 is C1 and a fraction 0.91 is C2. (line 1692-1694) The probability that we have S1 if we find C1 is 0.62 (line 1715). A calculation for the probability that we have S1 when we find C2 is not given, but using the same calculation, P(S1|C2) = 0.023. At line 1721 the section stops without concluding anything. It would be very helpful to get a summary here. Now my line of thought is: The relative risk S1/S2 = (0.875*2.15+0.125*0.38)/(0.09*2.15 + 0.91*0.38) = 1.93/0.54 = 3.57, so lower than 5.66. If we find C1 at post chill, we have 62% change of S1, i.e.: having a lot with 87.5% C1, and 38%	without any of the additional weighting factors associated with the differences in occurrence. We have added three figures to describe the decomposition of the population and the related probabilistic statements for the differing in risk of illness. The value calculated by the reviewer of 2.46 is correct, but this does not directly relate to the 5.66 values used in the dose-response development. The reasoning is correct, but it does not require any modification. We hope the addition of **Figures 18-21** helps clarify the issue.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				of S2, i.e.: having one with 9% C1, so overall 0.62*0.875 + 0.38*0.09 = 0.58 probability of actually being exposed to C1. This is h in table 25. If we find C2 at post chill, we have 2.3 % change of S1, i.e.: having a lot with 87.5% C1, and 97.7% of having one with 9% C1, so overall 0.023*0.875 + 0.997*0.09 =0.11 probability of actually being exposed to C1 when you find C2. This is k in Table 25. So the relative risk when finding C1 at post chill, as compared to C2 is (0.58*2.15 + 0.42*0.38)/(0.11*2.15+0.89*0.38) = 2.46 To me this implies that, if you believe in the assumptions made in the analysis, the relative risk of finding a cluster 1 serotype as compared to finding a cluster 2 serotype post chill is 2.46, not 5.66. It may be that you actually do this in the calculations that follow, but it is very challenging to follow the details. Please check my reasoning, I may have missed something somewhere. However, if I am correct, the analysis should be modified accordingly.	
6	91-92	2127-2134	A	I would challenge both assumptions. As for the first, I am quite surprised that all flocks are contaminated by *Salmonella*. In	The recent study by Obe (2023) found that all farms tested had *Salmonella* detected in at least

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				Europe only a minority of flocks is found to be positive. https://www.efsa.europa.eu/en/microstrategy/salmonella-dashboard. If then you assume that all flocks are contaminated by serotypes from both clusters, the contamination level must the extraordinary large. As for the second: if all flocks are contaminated with cluster1 *Salmonella*, I would guess that also all end products may be contaminated with cluster 1 *Salmonella* (just as, when only one serotype would be present in a lot, I would assume that all products in the lot can be assumed to be contaminated). I cannot directly foresee what that will do to the probability of illness per serving, but assuming that each tested carcass represents the only serotype on the carcass does not seem justified to me. The observation that this "will lead to estimates of the potential reductions in illness that are likely larger than what would occur in practice" could be a serious problem for the assessment. I would recommend to perform a scenario analysis to explore the impact of this source of uncertainty on the evaluation of health impacts.	one house. This reference was added to the document to further enforce this point. Further, a sensitivity analysis was performed to explore the impact of these sources of uncertainty.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
7			B	The serotype scheme is based on the WGS virulent factor-based clustering, so has a lot of support from the genomic data. The use of a conditional probability-based model fitting with consideration of different schemes and cluster behaviors for virulent /less virulent serotypes is appropriate.	No response required.
8			B	Also properly referenced and common distributions have been used to model the serotype mixture and support the assumptions	No response required.
9			C	The reviewer agrees with this assumption that it is not uncommon to detect multiple serotypes within flocks. However, the reviewer recommended including another scheme (scheme 3) when there are no obviously dominant clusters in a flock. In both Scheme 1 and 2, either cluster 1 or 2 were considered as the dominant one. The positive predictive value of identifying a flock with *Salmonella* in the high-virulence cluster in a scheme where high-virulence cluster is dominant or the negative predictive value of identifying a flock with *Salmonella* in the low-virulence cluster in a scheme where low-virulence cluster is dominant is expected to be higher than a scheme without an obviously dominant cluster (e.g., a flock with both clusters close to 50%). As a result, in a	While it may be likely that there are more than two schemes, it is beyond the descriptive power of a two-point sampling by two serocluster data set. For example, our current approach requires estimation of C1 frequency in each of two schemes and the frequency of the schemes. In **Table 25**, we have 4 equations and 3 unknowns. But, adding just one more scheme converts the algebra to 4 equations with 5 unknowns. Given the sensitivity analysis concluded that changes to the cluster scheme make little impact on illness estimates, to extend the assumption past two schemes is unwarranted and lacks an empirical grounding.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				scheme without a dominant cluster, false classification can be higher, which may lead to a larger disparity in predicted versus true avoidable cases. Hence, it is important to also evaluate a scheme without a dominant cluster for the impact assessment of receiving guidelines.	
10			D	The assumption that multiple serotypes are present within flocks is appropriate and needs further references to substantiate this.	See next two comments.
11	62	1396-1397	D	"Recent research has found that samples of individual chicken carcasses almost always contain multiple *Salmonella* serotypes (C. P. Thompson et al., 2018)." In addition to the reference cited, are there any other research reported or citations available to corroborate this, please see below.	Yes. Two additional studies have been added. Both present similar finding of *Salmonella* being present on all farms and the substantial mixing of serotypes within houses and/or farms.
12	62	1399-1403	D	"A limitation of this risk assessment is the lack of data to characterize the degree to which the serotype identified in the sample represents a dominant serotype within the flock, or if the serotype assignment is a poor predictor of serotype composition of the flock because there isn't an overwhelmingly dominant serotype within the flock." Although identifying this limitation is well appreciated, it would be nice to add some references to substantiate this.	Two new references documenting recent research on the high degree of serotype mixing were added.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
13	91-92	2127-2134	D	"The first assumption of this analysis is that all flocks contain a mixture of serotypes, some of which belong to the high virulence cluster denoted ….. The second simplifying assumption is that the serotype observed for each tested carcass represents the only serotype on the carcass (or at least that the observed serotype is sufficiently dominant as to explain the majority of the probability of illness given exposure). While the first assumption is reasonable, the validity of the second assumption is questionable (Cameron P Thompson et al., 2018) and will lead to estimates of the potential reductions in illness that are likely larger than what would occur in practice." Two assumptions are important and the authors have identified those.	Yes. Two additional studies have been added. Both present similar finding of *Salmonella* being present on all farms and the substantial mixing of serotypes within houses/farms
14			E	I am not familiar with the literature on the expectation of single or multiple serotypes within a flock. As a basic assumption, it would be difficult to "prove" that only a single serotype is to be expected since there is nothing preventing the presence of multiple serotypes (even if one is dominant), so multiple serotypes would seem to be the safer assumption.	No response required.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
e. Were any considerations missing from the development of the attenuation multiplier to adequately describe Salmonella growth and die-off after raw chicken product leaves processing?					
1			A	Some considerations are missing. This attenuation distribution is obtained from a referenced paper, without any reference to assumptions and shortcomings of the method. But it would be helpful with some background information, as it is crucial for all the calculations done. Still, it is an accepted peer reviewed paper, so we have to assume it is suitable.	We have added more description in our introduction of the attenuation distribution. Our referenced paper provides support for use of the lognormal distribution (Ebel, 2015): **Models for describing consumption dose distribution.** Data to directly estimate the parameters of the dose distribution $f(\theta_{consump})$ are rarely available. In this study, data to estimate the parameter vector of a distribution describing the average contamination of a food product were assumed to have been collected during or immediately after production and are represented as $\lambda_{test} \sim f(\theta_{test})$. The lognormal distribution is appealing for describing microbial data from different locations in the food chain *(7, 9, 16, 29, 49)* and when the observations are not integer valued (e.g., a most-probable-number [MPN] test with an LOD of 0.03 CFU/g or the averaging of multiple plate counts). The lognormal distribution also is mathematically convenient for scaling the pathogen level to account for variation in sample J. Food Prot., Vol. 78, No. 8 volume and sampling efficiency *(50, 51)*, modeling the effects of cooking *(4)*, growth *(34)*, and cross-contamination *(7)*. A lognormal distribution is obtained asymptotically even when intermediate processes that modify a lognormal distribution are not themselves lognormally distributed *(26)*. This result is important because even when some intermediate processes are not lognormally distributed, it is reasonable to assume that $f(\theta_{consump})$ follows a lognormal distribution.
2			A	Having said that, I don't manage to retrieve the origin of the numbers used (mean effect and sd), the reference is not very clear. It makes sense that the overall effect of all processes between food leaving the industry to the mouth of the consumer is a distribution, but this distribution is not necessarily lognormal, and undoubtedly uncertain. If the processes involved are just growth and inactivation, the lognormal distribution may be OK, but I have strong doubts when mixing, partitioning and bacterial transfer are involved. It would be nice to put this is a broader perspective by for example referencing Chapman et al (Microbial Risk Analysis 2–3 (2016) 3–15), Nauta and Christensen 2011 DOI: 10.1111/j.1539-	

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				6924.2010.01481.x and Neves et al, https://doi.org/10.1016/j.mran.2017.09.001. These models are mainly for Campylobacter and not for *Salmonella*, but still clearly illustrate that alternatives are feasible and may be more appropriate. I realize that using a different approach for the exposure assessment, as in the referenced papers, would much affect the overall modelling approach, may complicate the analyses done and does not necessarily reduce the uncertainty. I recommend to do some sensitivity analyses with a different model (such as the model presented by Nauta et al 2012) to explore how it affects the results. In a discussion, the authors should compare their approach to this approach or other approaches (such as those referenced above), address the uncertainties and explain the basis of their assumption that the use of a lognormal distribution is appropriate. An interesting difference between approaches used in these papers and the one used here, is that they are not anchored in the observed number of cases. The advantage of that is that the exposure assessment is probably more realistic (based on evidence on exposure), the disadvantage is that you usually get much higher estimates of the number of cases than	We have now also conducted sensitivity analysis on the attenuation distribution and included this variable in our examination of uncertainty. As it relates to absolute risk, we agree the attenuation distribution is important. Nevertheless, the ultimate output of the finished product standards assessment is the proportional reduction in illnesses. Therefore, the change in risk (before and after implementation of a risk management decision) is less affected by alternative assumptions about the attenuation distribution. This effect is explored in the sensitivity analysis. As we explain elsewhere, our default assumption of the same attenuation distribution for all chicken products is based on 1) a common target internal cooking temperature recommendation for any chicken product (the log10 reduction average should be similar for any product), 2) the serving size is similar across products, 3) the default attenuation distribution

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				what we actually observe in epidemiological data. (see https://doi.org/10.1111/risa.12153, https://doi.org/10.1111/risa.12538). In the approach used here all the “error” is basically put into the attenuation distribution, and it seems everything is OK. That is a practical approach, but not necessarily correct. There is much more to say about this than is done here. I realize that it will be a bit cumbersome to do that in this risk assessment, but compared to the extraordinary focus on dose response, this part of the exposure assessment would deserve some extra attention, definitely when shortcomings are discussed. We expect the attenuation distribution is the same for all serotypes (line 1640-1641), but are we sure? Are there no differences in growth, inactivation and persistence? All this deserves more discussion, and the potential impact of the associated uncertainties has to be addressed in this discussion. Another debatable assumption is that the attenuation distribution is the same for all three chicken products. As they are prepared differently, this is probably not the case. You could however benefit from the conclusion of Nauta and Christensen 2011 that the effects on	was calibrated to all chicken illnesses, and 4) alternatives were not readily available.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				relative risks from differences between products/models are usually not that large.	
3			B	For representing the attenuation factor, the direct use of fitted parameters of lognormal distribution for all the processes (mixing, partitioning, microbial growth and die off, etc.) involved in various steps of farm-to-fork processing of the product, is a shortcut method. There should be a justification like lack of data or assumption to use an average distribution to model all intermediate steps. Not enough support textual or R code has been given in the report for justifying this assumption.	We agree and have augmented the report with more description of the attenuation distribution.
4	71 192	1630 683	C	The inclusion of attenuation multiplier to consider the bacterial population change through different biological (growth, inactivation) and physical processes (partition, mixing, removal, cross contamination) is necessary, as the present model is not mechanism oriented. When assuming the independence between initial testing distribution and attenuation distribution, the utilized approaches of calculating composite mean and standard variation are reasonable. However, the assumption of independence was left untested. Evidence has shown that pathogenic bacteria on chicken carcasses can be influenced by the	We agree and have augmented the report with more description of the attenuation distribution. Furthermore, we have explored its influence on the effectiveness of risk management decisions in a sensitivity analysis.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				contamination of incoming birds, i.e., both the initial external carcasses contamination and the level in caeca (Seliwiorstow 2016, example of *Campylobacter* in broiler, 10.1016/j.ijfoodmicro.2016.03.010). It is recommended to address the potential correlation between the initial contamination and attenuation factor in the report, which has not been given sufficient consideration in the current analysis. In addition, although estimated parameters of the attenuation distribution were presented, limited information was provided regarding the mathematical process or data sources used to derive the attenuation distribution. The methods for deriving the attenuation distribution were referred to a prior publication that didn't include sufficient details either. The reviewer suggests including more details about the data and methods used for the parameterization of the attenuation distribution. The utilized modeling approach assumes that the attenuation factor is the same for both cluster 1 and cluster 2 bacterial populations. However, research has shown the great impact of between-strain variation (den Besten 2017, 10.1016/j.ijfoodmicro.2016.04.025). As shown in the study by Siceloff (2022,	We acknowledge that it might be possible that attenuation is (somewhat) different between cluster 1 and cluster 2. Nevertheless, this effect could already explain some of the difference between the dose-response functions for these clusters and the relative risk of observed illnesses used to develop these functions. Whether we incorporate uncertainty about different attenuation distributions – and their corresponding effect on the derived dose-response function – or simply accept that the current dose-response functions may reflect some difference in attenuation, neither relates directly to the effectiveness of

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				10.1128/aem.00204-22), the apparent discrepancy in serotype distribution between pre- and postharvest indicated the influence of processing steps on *Salmonella* are different between serotypes. To capture the variation to some extent amendable for risk assessment, cocktail cultures are commonly used for the development of predictive models. Although incorporating predictive model or following a mechanism approach is beyond the scope of this analysis, a discussion of the equal attenuation assumption is recommended. It could be challenging to create separate attenuation distributions for high-virulence and low-virulence clusters. It is worth mentioning this limitation transparently and discussing its potential impact on the estimation of risk.	detecting and diverting non-compliant lots at the end of production.
5			D	There are many steps from raw chicken leaving production/processing facilities to final consumption. There have several steps been overlooked. Sub-lethal cooking temperatures, improper storage, and cross-contamination or recontamination events can cause increase in pathogen load. This has been missing in the modeling approach, and needs to be considered.	We agree and have augmented the report with more description of the attenuation distribution. Furthermore, we have explored its influence on

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
6	34 191-192	898-901 655-658	D	"To describe growth and die-off of *Salmonella* in contaminated product lots as product travels from the end of processing, through commerce and preparation and consumption, an attenuation multiplier is used. The full derivation of this multiplier is described in Appendix C and an illustration of its utility has been shown in previous work (E. Ebel & Williams, 2015)." "Note that the dose-dependent probability of illness per serving has some inherent limitations, with the most obvious one being that the dose at the point of consumption is unknown. The second limitation is that it is difficult to model the changes between the last point at which the product is sampled." Although the limitation has been identified, it is an important one. For example, how the effect of cooking and other methods that reduce the pathogen level or cross-contamination and recontamination that increase the level would be modeled? Reduction in pathogen was represented through an attenuation distribution or multiplier. However, sub-lethal cooking temperatures, improper storage, and cross-contamination or recontamination events can cause increase in pathogen load. This needs further attention and must be addressed in the modeling approach. In general, in QMRA	the effectiveness of risk management decisions in a sensitivity analysis.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				studies, these steps are considered and modeled. For examples, please see below some studies. Jeong et al. 2019. Journal of Food Protection. DOI: https://doi.org/10.4315/0362-028X.JFP-18-113 Dan-Xuan et al. 2018. MDPI. DOI: https://doi.org/10.3390/ijerph15102324	
7			E	See comment above in Overall comments for Q1 related to the assumption of the same level of overall attenuation for the different chicken products.	No response required.
f. Does the Monte Carlo simulation approach adequately model the scenarios?					
1			A	Yes	No response required.
2			B	Implementation of MC simulation with the given virulence-based clusters' adjusted (adjusted with C1 and C2 multipliers) dose samples, drawn from a combined lognormal distribution of prevalence and attenuation factor, to estimate the illness responses is appropriate	No response required.
3			C	Based on the description on Monte Carlo simulation approach provided in the current version, the stability of the simulation results could not be sufficiently evaluated. The reviewer recommends including information	The model for illness reduction (i.e., 1-P(ill,new)/P(ill. Baseline)) is expected to be a smooth function. The model also is most unstable for the performance standards that allow the highest concentrations. One of the reasons for

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				regarding the criteria and testing methods used for assessing convergence in the report.	running the models for intact product (i.e., carcasses and parts) concentration thresholds higher than 10 cfu/mL was to assess convergence of the model. The number of simulations was increased from an initial value of 10 million Monte Carlo simulations until the estimated number of illnesses avoided in the upper tail (e.g., diversion of lots with greater than 10cfu/mL) agreed well with a 4th degree polynomial, with the definition of good agreement being a difference between the polynomial fit and the average of the Monte Carlo estimates being roughly 10% or less for the 50 and 100 cfu/mL threshold values. This approach resulted in there being differences of less than about 3% for remaining concentration thresholds.
4			D	Monte Carlo simulation technique adequately modeled the scenarios.	No response required.
5	80	1855-1857	D	"The baseline probability of illness is determined using numerical integration and the probability of illness among passing lots are estimates from 10 million Monte Carlo iterations." The number of iterations are sufficient.	No response required.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
6			E	There is no reason to believe that Monte Carlo simulation would not adequately model the scenarios. However, it is worth exploring whether more simplified approaches can replicate the estimate of impact of risk management measures, such as by considering the impact on the arithmetic mean of the distribution of the raw products, before and after risk management actions are implemented. In addition, the Monte Carlo simulation may benefit from the use of Importance Sampling of the right tail of the distribution (e.g., doing the risk calculations with a higher concentration distribution g(d) with good coverage of the right tail and then reweighting the samples by multiplying the resulting values by f(d)/g(d), where f(d) is the original target distribution). Given the simplicity, simple numerical Integration (non-Monte Carlo, and therefore not subject to reliance on random number generation) due to the rare event nature and the use of lognormal distributions that span 12 orders of magnitude (+/- 3 standard deviations with sigma=2 log10 units).	A subsection exploring the simplified approach was added to **Chapter 5 Final Product Standards**.
g. What approach could be taken to assess uncertainty in these conclusions?					

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
1	71	1651-1652	A	What is meant by "substantial uncertainty"? How important is that?	An uncertainty analysis was added to **Chapter 5 Final Product Standards** to quantify "substantial."
2			A	I realize it is challenging to characterize the uncertainties in the conclusions, but it could be done by adding a "layer" of expert knowledge elicitation, as for example described in the EFSA uncertainty guidelines. (EFSA Journal 2018;16(1):5123, 39 pp. https://doi.org/10.2903/j.efsa.2018.5123). This could be a very useful approach to address the overall uncertainty for conclusions like this, once the assessment questions are well defined. It can be done, though maybe not feasible anymore at this stage of the risk assessment process. Alternatively, scenario analyses addressing the uncertainties can provide insight on the impact of different uncertainties in the data and modelling assumption. This has, among others, the advantage that the risk assessor is forced to list those uncertainties.	An uncertainty analysis was added to address the overall uncertainty in the conclusions and a sensitivity analysis was added to provide insight on the contributions to uncertainty of different data and assumptions. Those uncertainties were all outlined in the original version of the report, but a summary table has been added to **section 1.6 Introductory Tables and Figures**.
3	87-88	2016-2021	A	Again, the uncertainty is "substantial". I wonder where that falls in the set of qualifications "considerable", "large", "quite large" and "very large". (I try to clarify my point with some irony here, what I mean is that "substantial" is	An uncertainty analysis was added to **Chapter 5 Final Product Standards** to quantify "substantial."

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				undefined.) You could give the analysis a try by evaluating scenarios. Yet, it seems this will/can be done in the future (line 2019). Having said that, I agree with the conclusion that this 25% reduction will be hard to achieve, I would also be surprised if the uncertainty was that substantial.	
4			B	Illness responses estimated from the MC simulated cluster-wise log doses were only integrated to calculate the mean value. The integration function 'f' coded in lines 50-52 in the given script 'FinProdStds_Jan25_23.R" with integrated 'DRPolyForInteg' and 'DRPoly' functions also estimates lower and upper limits of illness, using the polynomial solution of clusters multipliers which were imported as summary values from 1000 bootstrapped samples, considering the uncertainty of the data.	Thank you for the advice. FSIS carefully evaluated these expert suggestions and integrated them into the sensitivity and uncertainty analyses that have been added to the Final Product Standards Chapter.
5			B	Mostly posteriors from Monte Carlo simulations from Bayes implementations are used for estimating the uncertainty which were already implemented in serotype modeling by EpiX. Other method to account uncertainty can be the use of mixture distributions like gamma or beta distribution can be used as per the types of measure used like count/concentration or	FSIS carefully evaluated these expert suggestions (comments 4-8) and integrated them into the sensitivity and uncertainty analyses that have been added to **Chapter 5 Final Product Standards**.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				presence/absence. A good reference can be: Zheng, J. and Frey, H.C. (2004), Quantification of Variability and Uncertainty Using Mixture Distributions: Evaluation of Sample Size, Mixing Weights, and Separation Between Components. Risk Analysis, 24: 553-571. https://doi.org/10.1111/j.0272-4332.2004.00459.x	
6			C	Predictions made in this risk assessment primarily reflect average effects, and do not account for many variability factors, such as variations within and between flock/lot. Moreover, some variables that represent uncertainty due to lack of knowledge were parameterized using deterministic values. Sensitivity analysis can be considered to evaluate the change in predictions by varying input variables in a feasible range (outlined in Q1-b). Furthermore, second-order Monte Carlo simulation can help distinguish between the sources of total indeterminability, once variability and uncertainty can be better characterized.	
7			D	There are methods the authors may refer to assess uncertainty. For example, the knowledge uncertainty and stochastic variability for multiple simulations could be better	

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				presented with complementary cumulative distribution functions (CCDFs) graphs and the summary can be well-represented using 5th, 50th, and 95th percentiles. See this reference as an example: Complementary cumulative distribution function (CCDF), Treatment of Uncertainty in Performance Assessments for Complex Systems, https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1539-6924.1994.tb00266.x Also, Monte Carlo simulation can take into account the input uncertainties (correlated or uncorrelated inputs); See example: Uncertainty estimation and Monte Carlo simulation method https://www.sciencedirect.com/science/article/abs/pii/S0955598601000152	
8			E	Modeling uncertainty with respect to the conclusions from an entirely bottom-up approach (propagating uncertainty in each input toward an overall characterization of uncertainty) is likely to be very challenging, and certain characterizations of uncertainty will be nearly impossible to quantify. The uncertainty analysis may be dramatically simplified if the overall estimation process could	

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				be dramatically simplified. However, this remains to be confirmed. If allowable, a robust sensitivity analysis may provide sufficient evidence of the role of uncertainty to inform risk management decisions.	
h. Are the conclusions drawn from the analysis appropriate?					
1			A	Overall, the conclusions seem appropriate (fit for purpose), but they are not clearly communicated (see first comment Q6.)	This comment is addressed in comment 2 below.
2	88	2029-2032	A	I challenge that your analysis demonstrates that higher levels of *Salmonella* on raw products are associated with higher risk of illness. I do believe it is true, but it is something you put in the models by the definition of the DR model, you don't need the analyses for that.	FSIS clarified in the document that the usage of the dose-response model is an implicit assumption, and that the analysis attempts to quantify the degree of this assumption, rather than demonstrate it.
3			B	The scenario analysis and integrated Dose-response models accounting for the genetic virulence-based cluster are very extensive and appropriate, employing the latest and most established methods used for QMRA modeling. The conclusion drawn to avoid illnesses/year from *Salmonella* contamination in different product types looks legitimate.	No response required.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on chicken at receiving and in final chicken products. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
4			C	Yes, the conclusions are supported by results obtained.	No response required.
5			D	The conclusions drawn from the analysis are reasonable and appropriate.	No response required.
6			E	I cannot detect any way in which the conclusions drawn could be characterized as inappropriate.	No response required.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Please identify limitations, weaknesses, or inadequacies of the process control modeling techniques and data. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				General Comments	
1			A	The reduction approach compares a reduction in the level of APC with prevalence of *Salmonella*. There is definitely logic in that approach. I wonder however, why you do not consider the level of APC per se. I can imagine that the levels are somehow correlated as well. The APC level at the start of the process may therefore be correlated with the *Salmonella* level or the prevalence at the start. I imagine that more hygienic companies have lower levels at the start. Once the performance standard is defined as a reduction, it may be beneficial for a company to be non-hygienic, as in that case it is easier to achieve a reduction. However, in that case the level of *Salmonella* will not be lower in the end, so the objective is not achieved. I suggest that you show the correlation of the levels as well and discuss the usefulness of a performance standard based on the level instead of reduction.	The idea of linking levels of indicator organisms to either the prevalence or levels of pathogenic bacteria has been proposed and studied on multiple occasions. In these previous studies there has either been no significant relationships, or the correlation was too low to develop effective performance standards. A brief discussion and additional references were added to address this point.
2	104	2379-2381	A	Missing figure numbers.	The figure numbers were corrected.

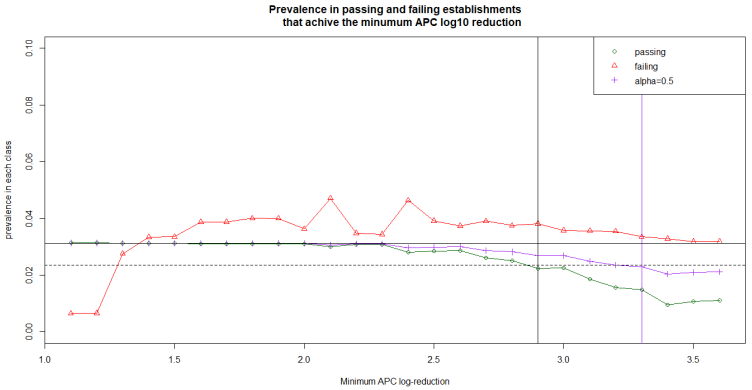
Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Please identify limitations, weaknesses, or inadequacies of the process control modeling techniques and data. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
3	105	Figure 25	A	In the caption, also explain the bars and the lines.	Figure captions were corrected throughout the document.
4	112	2546-2547	A	Specify it is *Salmonella* prevalence	The prevalence was specified as *Salmonella*.
5	112	Figure 29	A	I do not understand the red line in the figure. Why does it have this shape? I assume failing = non-compliant and passing = compliant? Please be consistent in the terminology	The terminology was clarified.
6	114-115	2581-2600	A	This paragraph requires some knowledge on the current approach, that I am not familiar with. Please define exactly what the sensitivity and specificity stand for. I get the idea the reasoning is OK, but it is hard to check.	Clarification was added.
7	118-119	Summary	A	The summary is a (good and relevant) discussion, not a summary. Please summarize the conclusions.	A summary of conclusions was added to the discussion chapter.
8			B	Process control modeling is good with the support of appropriate plots of fitted APC data (baseline vs current data). The approach is suggesting the two well-reasoned approaches, including APC log-reduction ~ *Salmonella* and APC fraction present~ *Salmonella* for managing the performance standards, enforcing	No response required.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Please identify limitations, weaknesses, or inadequacies of the process control modeling techniques and data. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				compliant/non-compliant status ('two-strata model') on establishments, irrespective of their size/nature, in the lieu of reducing *Salmonella* related illnesses.	
9	103 107		B	After introductory paragraphs in lines 2323-2334, in sections 7.1 to 7.4 methods and results are presented combined, however, overall interpretations and takeaways from the used methods are not very clearly explained for all sections. The figures' captions and their explanations are almost similar or merged together with very few added inferences. Statement in line 2347-2350 in the Data description section suggest that "there was no further analysis of the data collected for this (or these? unclear) organisms". This contrasts with next sections 7.2-7.4, In lines 2436-2438 and 2441-2442, suggesting interest and appeal of using the indicator organisms-based performance standards, please clarify	Text was added clarifying indicator use.
10	104	2373	B	The equation has typos (,) and unexplained variables named y_{pc} and y_{rh}	The explanation of the (,) notation was clarified.
11	104	2384	B	Was a censored version of the distribution was used? Please clarify.	A censored version of the distribution was not used. These data are censored, because microbiological data are generally collected by methods with upper and/or lower

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Please identify limitations, weaknesses, or inadequacies of the process control modeling techniques and data. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
					limits of detection. The distribution accounts for this censoring using best statistical practices (Helsel, D. R. 2005).
12	105	2391-2394	B	What is taken away from Figure 26 is, an unclear explanation.	Additional text was added explaining.
13	113	2566-2568	B	Reference?? for the statement stating the inverse relationship between volume ~ and contamination.	This is demonstrated in **Figure 39** and the attending discussion.
			C	(No comments from this reviewer)	No response required.
14			D	Overall, the process control modeling techniques and data are reasonable and appropriate. However, the risk management question number 4 was not addressed (please see Q5 (d) below).	Discussion of risk management question #4 was added to **Chapter 8 Discussion**.
15	104	2359	E	"where the average APC levels were 4.50 at rehang and 2.46 at post-chill, for an average log reduction of 2.04" This type of change in the log-scale average is used to then say "on average only 1 aerobic bacterium out of every 1000 is surviving between rehang and post-chill". The use of the differences in the average on the log-scale can be misleading. The actual average of 1 bacteria surviving out of every X would be best represented by	The use of the differences in the average on the log-scale can be misleading. The actual average of 1 bacteria surviving out of every X would be best represented by the ratio of the incoming and re-hang arithmetic average concentrations. Means on the log-scale do not predict the overall average fraction of survivors. This consideration applies to the risk management question itself, as to whether a difference in the arithmetic mean, rather than the geometric mean, should be the appropriate basis for the risk management intervention. It is possible to have no change in the average log10 value, while having a

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Please identify limitations, weaknesses, or inadequacies of the process control modeling techniques and data. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				the ratio of the incoming and re-hang arithmetic average concentrations. Means on the log-scale do not predict the overall average fraction of survivors. This consideration applies to the risk management question itself, as to whether a difference in the arithmetic mean, rather than the geometric mean, should be the appropriate basis for the risk management intervention. It is possible to have no change in the average log10 value, while having a substantial reduction in overall contamination levels (when using the appropriate, i.e., non-logarithmic) scale.	substantial reduction in overall contamination levels (when using the appropriate, i.e., non-logarithmic) scale.
16			E	I agree with the use of the maximum likelihood techniques to deal with below LOD results as far better than substitutions with ½ the LOD or similar simplifications.	No response required.
a. Is the correlation between indicator organism and *Salmonella* fully described and well characterized?					
1			A	Correlation is not formally studied. What is shown is the relationship between APC reduction and *Salmonella* prevalence (fig 27), and between APC < LOD and *Salmonella* prevalence. There is reference to a weak correlation that was previously found. (lines 649-379 in the executive summary)	Correlation between indicator organisms and pathogens has been studied extensively and the only observed correlation is that outlined in this approach and Ebel 2015.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Please identify limitations, weaknesses, or inadequacies of the process control modeling techniques and data. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
2	104	2367	A	Correlation is only mentioned when there is reference to a previous study.	While correlation between indicator organisms and pathogen rates has been the subject of long study, the only correlation that has been established is in the referenced study (Ebel, 2015).
3			A	As an answer to the question, I would say the correlation is not fully described. I do think however it is sufficiently characterized for the purpose of the assessment and no action is needed	No response required.
4			B	Yes, well described and plotted clearly.	No response required.
5	109	Figure 27	C	The correlation between APC reduction and the proportion of *Salmonella*-positive samples might have been overestimated. The relationship was built using data from two primary groups of establishments characterized based on production volume and business model: high-volume corporate and low-volume independent. However, the data distribution between the two groups is highly imbalanced between the range of low and high APC reduction. In the range of above 2.5 log reduction in APC, most data points are based on high-volume corporates, appearing more or less randomly distributed across the range of 2.5-3.5 log reduction, which is expected to show no correlation between log reduction and proportion of *Salmonella*-positive	That process control guidelines do account for establishment volume. This can be seen in **Figure 35** where the prevalence of failing establishments is not monotonic and the overall prevalence is essentially unchanged until the average log reduction exceeds about 2.5. These phenomena occur because the extremely small contribution of small establishments to the overall production volume.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Please identify limitations, weaknesses, or inadequacies of the process control modeling techniques and data. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				samples. However, a decreasing curve was estimated in this range, which could be the result of integrating data points in the low log reduction range originating from a different type of establishments. The reviewer recommends the development of separate relationships for different establishment types, which can be used to predict avoidable cases. Based on the findings presented in Figure 27, it is reasonable to expect a lower value of predicted avoidable cases once separate relationships are developed and used. This is because the reduction in prevalence of *Salmonella* would be predicted less for large establishments that contribute to most of the chicken production in the U.S. Developing separate relationships for different establishment types could lead to more accurate predictions of avoidable cases, which should be lower than the predictions in the current analysis.	
6			D	The correlation between indicator organism and *Salmonella* are reasonable.	No response required.
7			E	The relationships are well described, including limitations. More emphasis on the concept of the fraction of surviving organisms (and to what extent FSIS	Discussion of the “fraction of surviving organisms” has been added above in Q5 General Comments E#15.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Please identify limitations, weaknesses, or inadequacies of the process control modeling techniques and data. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				believes that *Salmonella* would share a similar "fraction of survival" with the indicator organism) might make the relationships more easy to understand (particularly for those skeptical of the use of indicator organisms), and may help to understand why it is difficult to see correlations for censured data as concentrations get quite low, despite the expectation that concentrations of the target and the indicator organisms are both being reduced by the same interventions within the processing environment.	
b. Is the process control modeling technique accurately described, utilized, and appropriate for its intended use?					
1			A	The modeling approach is clearly described. I think it is utilized well and appropriately.	No response required.
2	111	2519-2521	A	A concern I have is the alfa = 0.5. This holds the bold assumption, that if the standards are not mandatory, half of them will introduce the standards. Is there any evidence that this is realistic? Would it not be predominantly those that are already paying attention to performing well, that would apply the standards on a voluntary basis, so the actual effect will be less?	Relying on past FSIS risk assessment experience, it was determined that alpha = 0.5 was a sufficiently descriptive assumption for answering the charge questions.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Please identify limitations, weaknesses, or inadequacies of the process control modeling techniques and data. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
3			B	Well explained and correctly modeled. The suggestion of mandatory compliance for APC eliminations looks statistically supported.	No response required.
4			C	The description of the modelling technique theory in the report appears to be adequate. The approach and summarized data presented in the report have the potential to provide valuable insights for further research. However, some model's quantification was not provided, such as the relationship between the log APC reduction and proportion of *Salmonella* positive (outlined in Q5-a) and the parameterization of the attenuation distribution (outlined in Q4-e), which can be a drawback, limiting the applicability of the model's findings/summarized data in other contexts.	These comments were addressed in the questions in the reviewer mentioned Q5-a and Q4-e.
5			D	The attempt and efforts made to describe and use of process control modeling techniques are appropriate.	No response required.
6			E	The description is sufficiently described and techniques appear to be appropriate	No response required.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Please identify limitations, weaknesses, or inadequacies of the process control modeling techniques and data. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
c. Are the data analyses and source code accurate?					
1			A	I have not been able to identify any inaccuracies, but it is not well explained what data are actually present in the data file "Chicken_carcass_exp_APC_review.csv". Column names are abbreviations that are not explained.	Explanations of column name abbreviations have been provided in the code book. [Part of OPARM Data Posting Draft Document]
2			B	Yes, the R code runs fine and matches the reported estimates.	No response required.
3			C	The R codes were developed in line with the method description.	No response required.
4			D	Data analyses and source code given the scope of the work are appropriate.	No response required.
5			E	I was not able to scrutinize the data and source code for this part of the analysis	FSIS provided all relevant data and source code to the external peer reviewers, but we appreciate how large the volume of written documentation and code provided to the peer reviewers was, and appreciate how carefully the reviewers assessed the document in a short timeframe.
d. Is the risk management question adequately answered, given the limitations in assessing public health impacts from indicator organism policies?					
1			A	The risk management question reads: "What is the public health impact of monitoring/enforcing process control from rehang to post-chill? Monitoring could include analytes such as	Interpretation and responses to each risk management question are now included in **Chapter 8 Discussion.**

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Please identify limitations, weaknesses, or inadequacies of the process control modeling techniques and data. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				Enterobacteriaceae, Aerobic Plate Count, or other indicator organisms, analysis could include presence/absence or levels and the monitoring could also include variability of actual result versus expected result, log reduction, absolute sample result, or other individual establishment specific criteria." As the question is not perfectly clear, it is hard to judge whether the answer is adequate. The answers are not spelled out in a conclusion section or in the summary, which makes it even harder to judge whether they are adequate. I recommend a dedicated Section in the report where each risk management question is answered.	
2			A	As I interpret it correctly, the answer for the APC reduction approach is that the model indicates that a standard of 2.9 log reduction would achieve the 25% objective if all establishments achieve it, and 3.3 if half of them do. This is useful information, but not an answer to the question as formulated.	The broad question formulation, and the limits of what could be analytical assessed, are discussed in the executive summary, process control, and discussion chapters.
3	27-28	670-679	A	I find a conclusion formulated in the Executive summary. This answers the question how the HP2030 targets can be achieved, but that is not the formal question.	As outlined in the text, the formal risk management question related to process control could not be fully answered. In an effort to provide useful and meaningful information related to this risk management question, FSIS chose to focus on the Healthy People 2030 targets as they

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Please identify limitations, weaknesses, or inadequacies of the process control modeling techniques and data. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
					were both of interest to FSIS risk managers and could be answered by the available data.
4			A	One can wonder whether I do think that the answer given is adequate given the limitations in assessing public health impacts from indicator organism policies. The analyses performed seem sound, but require an assessment of the uncertainty to be sufficiently informative for the risk managers and that needs to be clearly communicated.	Given the prevailing focus of FSIS risk managers and stakeholders on enforceable final product standards and the timeframe in which these analyses were conducted, an uncertainty analysis could not developed at this time. FSIS is considering future efforts to more fully assess uncertainty in process control.
5			B	Yes, all limitations and feasibilities of implementing the new compliance standards are presented with the support of a statistical model and observed/fitted data descriptive plots.	No response required.
6			C	This analysis assumes that prevalence reduction in chicken products is proportional to the risk reduction, and therefore avoidable cases were not estimated to address the risk management questions. This approach seems reasonable for the intended purpose. General, risk management questions in this section were adequately answered. However, the means used to quantify the association between APC-based standards and *Salmonella* prevalence may have led	Detailed response is provided in Q5-a, as outlined by the reviewer.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Please identify limitations, weaknesses, or inadequacies of the process control modeling techniques and data. Please provide alternative data, data analysis, and/or modeling approaches if the FSIS approach is deemed inappropriate or inadequate. Specific consideration should be given to the following:					
				to an overestimation of the predicted effect (outlined in Q5-a).	
7			D	Given the limitations in assessing public health impacts from indicator organism policies, risk management questions were appropriately answered except risk management question number 4. The analyses for question number 4 were not performed. Please include the analyses for this question in the report or delete this question from the report if the analyses have not been done.	FSIS has taken additional steps to more fully respond to risk management question #4 in **Chapter 8: Discussion**.
8	103	2323-2324	D	"Multipoint sampling (e.g., rehang and post-chill) is required in the evaluation of process control, and thus, not assessed for parts and comminuted product." Please mention that this evaluation of process control is for carcass.	It was noted in the **Chapter 7 Process Control** that the evaluation was only for carcasses, but additional clarification was added to the introductory material, particularly in **Table 10: Interpretation of risk management questions and table of scenarios**.
9	16	385-386	D	Risk Management Question # 4: What is the public health impact of implementing combinations of the risk management options listed above? This reviewer could not find this analyses and results in the report related to the combined efforts of risk management questions 1, 2, and/or 3.	FSIS has taken additional steps to more fully respond to risk management question #4 in **Chapter 8: Discussion**.
10			E	The answers to the question are appropriate to the analysis undertaken.	No response required.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q6. Evaluate whether the documentation of the data and modeling, and interpretation of the results is appropriate. If not, please provide an alternative outline, approach, and/or suggested language for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
				General Comments	
1			A	A main concern about the risk assessment report is that it lacks a specific section summarizing the conclusions. I have been looking for the answers to the four risk questions, but cannot easily find them. How can risk managers use this risk assessment if there are no clear and well formulated answers to their questions? The conclusions in the executive summary state that "The estimates provided in this risk assessment can be used by risk managers" (line 699) but does not summarize these estimates. A clear overview of risk questions and answers should be provided in the executive summary. There is more attention for the novel DR model and the bioinformatics (also in these charge questions, by the way), than on how the risk questions are answered and the quality of the answers. This surprised me, as the answers to the questions are usually by far the most important message to the risk managers that required this risk assessment. Clearly, if new methodologies are introduced, they should be critically evaluated, but it would be helpful to also explicitly address	While the executive summary does provide a complete overview of all findings and estimates in the risk assessment, FSIS has taken additional steps to more fully respond to risk management question #4 in **Chapter 8 Discussion** to provide additional clarity. **Chapter 8 Discussion** outlining the response to the risk management questions has been added to address this imbalance. This concern boils down to the fraction of higher virulence cluster 1 serotypes among *Salmonella.* We've added a sensitivity analysis exploring this effect. As outlined in previous comments: yes, it matters with respect to the calculated probability of illness, but it doesn't change conclusions which estimate the effectiveness of the policy.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q6. Evaluate whether the documentation of the data and modeling, and interpretation of the results is appropriate. If not, please provide an alternative outline, approach, and/or suggested language for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
				what they actually add to the answers of the risk assessment. As I understand it, the two-curved DR model in principle allows to divert "higher risk" lots, which may give more efficient risk reduction. The bioinformatics is used to identify these "higher risk" flocks, so it is a tool that is used to improve a new tool. This seems a sound approach to me. But I get the impression that the identification of the lots as C1 or C2 based on one sample per lot is the main source of uncertainty when this approach is implemented, much more than the details in the new approaches. This issue is not addressed and it should be.	
2			B	The compilation of the document in terms of describing the data collection, assumption, and usage for data modeling is extensive	No response required.
3			B	Also, the accompanying R code supports the implementation of the theoretical foundation explains in the report.	No response required.
4			B	There were a few typos, unclear statements, or discrepancies, noticed during through revision of the report, as enumerated in the response to the next question, that need attention.	Typos, unclear statements, discrepancies were addressed throughout the report and directly responded to in this document as appropriate.
			C	(no comments from this reviewer)	No response required.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q6. Evaluate whether the documentation of the data and modeling, and interpretation of the results is appropriate. If not, please provide an alternative outline, approach, and/or suggested language for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
5			D	The report is not well written, as there are many redundancies, in some cases lack justifications (as mentioned in comments to questions 1-4), missing conclusions and limitations sections in the main report, and several typographical and grammatical errors. This report needs thorough proof reading and improvement in writing. For details and specifics, please see below.	The report was thoroughly proofread, and the mentioned details and specifics were corrected. Responses to the reviewer's questions are provided below.
			E	(no comment from this reviewer)	No response required.
a. Is the report clearly written and complete?					
1			A	As a non-native English speaker, I do now and then struggle with the terminology. As the foreseen readership of this report will in the United States, that is probably OK, but please be aware of it.	No response required.
2			A	The figure captions are generally very poor. When a graph is presented, it is a good habit to explain what the different lines and dots represent, what is on the axes and what the reader should read from the graph. The readability of the report would be improved by clearer figure captions.	Figure captions were improved throughout the document.
3			A	Missing Table 6, table numbering and table of contents need to be updated.	Table numberings were corrected.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q6. Evaluate whether the documentation of the data and modeling, and interpretation of the results is appropriate. If not, please provide an alternative outline, approach, and/or suggested language for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
4	19	481-485	A	Where is the answer to risk management question #4? I find it in line 590. An overview of the answers to the questions would help.	While the executive summary does provide a complete overview of all findings and estimates in the risk assessment, FSIS has taken additional steps to more fully respond to risk management question #4 in the Chapter 8 Discussion of this document.
5	26	Table 5	A	The number of flocks diverted seems to be identical to the number of annual salmonellosis cases that could be avoided. I don't think this Table is correct. I think it should be the same as table 30. Please correct the Table	This table was corrected.
6	34	910	A	The use of the word "safe" is questionable as there is still a risk.	The word "safe" was modified to "safer."
7	36		A	Figure 4 is misleading, as "rehang" is after the actual slaughter of the birds, the process after rehang is not slaughter and processing, but just processing (it took me a while to understand that). Then, for Growth and Die Off, you don't use a multiplier (which to me is a number) but an attenuation distribution (line 1628), which makes much more sense. So don't use the term "multiplier" in the figure and be consistent in the terminology used. Finally, the directions of the arrows in the figure suggests that you estimate the annual number of illnesses from the exposure + dose response relation, but you don't, the annual	Given the paucity of preharvest data, rehang was used as the nearest available proxy for slaughter. Therefore, the title "slaughter and processing" is an appropriate descriptive term for a schematic that describes the broad strokes of the model. The direction of the arrows and other figure revisions have been made.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q6. Evaluate whether the documentation of the data and modeling, and interpretation of the results is appropriate. If not, please provide an alternative outline, approach, and/or suggested language for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
				number of illnesses is estimated from the epidemiological data (lines 1568 – 1574 and 1834 -1840). Please revise the figure.	
8	37	921	A	Part I does not include an exposure assessment as defined in the glossary. The word exposure occurs only once in this part.	The heading was removed. However, the definition of exposure assessment is accurate. See Comment Q6e #3 for more clarification.
9	38	935-937	A	It looks as if you use a minimum infectious dose principle here, which conflict with the DR models used later on.	The conflicting sentence was removed.
10	38	939-945	A	References are required here.	The missing references were inserted.
11	40	975	A	To me this suggest that sampled and rinsed carcasses are re-entering the food chain(?)	FSIS uses a food-safe broth for carcass and parts rinsing: New Sampling Instructions and Testing for Chicken Parts and NRTE Comminuted Poultry | Food Safety and Inspection Service (usda.gov)
12	41	Table 10	A	What is var[P]?	The notation “var[p]” was replaced with the symbol $\sqrt{var[P]}$.
13	42	Figure 5	A	Are these annual numbers?	Yes Figure 5 represents annual numbers. This has been clarified in the figure caption.
14	43	Figure 6	A	Explain better what is shown in the graph, the axes legends are not clear.	The axes legends have been edited to improve clarity.
15	46	Table 11	A	What are the units of the numbers?	The units have been further clarified.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q6. Evaluate whether the documentation of the data and modeling, and interpretation of the results is appropriate. If not, please provide an alternative outline, approach, and/or suggested language for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
16	46	1109	A	Please explain what you mean by a more homogeneous distribution	Explanation was added to the text clarifying "homogenous distribution."
17	47	Figure 8	A	Again, the figure caption is not precise enough.	Figure captions were updated for Figure 8 and throughout the document as appropriate.
18	48	1128-1131	A	I would be very helpful if you explain here what you use this distribution for (and what not). It has taken me quite some time to discover that you use this in combination with the attenuation distribution to fit the DR relation to the observed number of cases. At least, that is what I think you do in the end (line 1624-1631).	Upfront clarification was added to express that the mixture distribution was developed to fit the dose-response relationship used in the risk assessment.
19	48	Figure 9	A	Again, the figure caption is not precise enough.	Figure captions were updated for Figure 9 and throughout the document as appropriate.
20	Chapter 4	1480-1516	A	I get the strong impression that this chapter is written by different authors. Some parts of it (e.g.: Lines 1487-1506) read more like a commercial or a grant proposal than a summary of scientific evidence. The only relevant information of this chapter for the risk assessment seems to be in lines 1512-1516. I recommend to shorten this section.	We respectfully disagree with the reviewer and note that other reviewers did not identify this as an issue. No changes to the document were made.
21	66	1528-1539	A	This describes the background of an important estimate, 66% attribution to food for *Salmonella* infection. This is derived from experts with a holistic look (this is unclear,	In this instance, the "holistic look" refers to the consideration of all transmission pathways including potential subpathways to formulate/estimate comprehensive attribution rates in this expert elicitation proceedings.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q6. Evaluate whether the documentation of the data and modeling, and interpretation of the results is appropriate. If not, please provide an alternative outline, approach, and/or suggested language for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
				please explain what is meant), not a word about uncertainty. In comparison with the very detailed statistical analyses at other points in the risk assessment, this surprises me. I would expect a better basis. Please discuss the uncertainty.	This approach estimated the attribution rate to foodborne *Salmonella* was approximately 66% with a 95% uncertainty interval of 48%-81%. Other similar studies assessing foodborne pathways of *Salmonella* fall well within the range of uncertainty, largely overlapping confidence intervals and hovering near the estimated mean (Netherlands 55% (95% CI 32-88%); Canada, median 63% (90% CI 32-80%); Australia 71% (min-max 65-75%). Further, these estimates are generally much lower than that original derived from CDC-reported outbreak data and a case study of sporadic illnesses (Scallan (2011), 94%). Uncertainty remains in this estimation as there is a lack of reliable and robust data on foodborne illness attribution, a dynamic process with underreporting and coverage issues. Nevertheless, these are the best available estimates and the uncertainty does not impact the proportional reduction in illness, only the illnesses prevented estimates.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q6. Evaluate whether the documentation of the data and modeling, and interpretation of the results is appropriate. If not, please provide an alternative outline, approach, and/or suggested language for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
22	70	1601-1602	A	Provide references if you refer to academic literature	References were added to some recent examples of risk assessments that take the described approach.
23	88	2031	A	Magnitude of level = level	The sentence has been rephrased to “the level and its frequency.”
24	103	2327 and 2338	A	Here a sentence is repeated.	The repetition was removed.
25	104	2356-2357	A	If the 1.39 is for the positives only, the actual log reduction is larger. This is not clear.	Censored data methods were used to account for the samples below the limit of detection of the FSIS APC assay. The 1.39 log reduction is not for “positive” samples only and is an accurate representation of the APC log reduction.
26	104	2379 and 2381	A	Figures do not have a number, please add.	The figure number was corrected.
27	111	2506	A	What does exp stand for?	The notation is defined in the before it is used text.
28			B	The overall report is well written, some textual issues, like: Throughout the used fonts inconsistency was observed, especially in equations and notations)	Font inconsistencies were corrected throughout the document. Responses to the reviewer’s questions are provided below.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q6. Evaluate whether the documentation of the data and modeling, and interpretation of the results is appropriate. If not, please provide an alternative outline, approach, and/or suggested language for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
				In addition, there were some technical issues observed, which are enumerated as follows:	
29	22	513-515	B	The units in Table 2 should be reorganized such that they can specifically represent the carcasses and parts, as measured in CFU/ml, while comminuted products in CFU/g (mentioned on page21, lines 506-507)	The layout of Table 5 was retained to best communicate the findings with the risk managers. Additional unit labeling was added to improve clarity.
	24	553-554		A similar problem is in Table 4, units for carcasses, parts, and comminuted chicken are all mixed up, it's an effort to understand what is in what unit, please reorganize, probably using units in the column names of corresponding chicken products can be a possible solution to address this units' discrepancy.	Units were added to the column names for Table 4 as recommended.
	24	569-570		This confusion in the units has passed into the text also, in line 570 comminuted chicken performance is mentioned in CFU/ml, which contrasts with the measuring unit CFU/g for comminuted chicken as mentioned on page 21, lines 506-507. Please be consistent.	The units were corrected.
30	42	1016	B	Figure 5, specify the time scale for the x-axis: 'number of samples collected" in what time, per month? or mention it in the figure description.	Figure 5 was updated to clarify it is the annual number of samples collected.
31	45	1084	B	Description "The bottom row of Table 11 provides the parameters for the population"	**Table 11** was corrected.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q6. Evaluate whether the documentation of the data and modeling, and interpretation of the results is appropriate. If not, please provide an alternative outline, approach, and/or suggested language for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
				does not match the values in the last bottom row of table 11. Seemingly, the given description implies to second last row. Please correct. The actual data in the last row of the table has the value 'Concentration/gram' for the column 'Sample Location' and values given under columns $\overline{\mu}$ and $\overline{\sigma}$ does not make sense. Either clarify that to fit in the table layout of columns or take that information out of the table and mention it in the text	
32	46	1101-1103	B	Please make sure the units in the first column for table 12 are consistent, given the earlier unit mismatches, especially parts/ml and parts/gram.	The units in **Table 12** were correct.
33	53	1253-1259	B	The equation in line 1256 and the contextual description in these lines and on page 52 seem a little mismatched and unclear, please clarify and rectify, if there is any type in the notations of the equation in line 1256.	The equation was corrected. The typo was LOD should have been LOQ.
34	53	1271	B	Repeated words are in the line "These issues can be further summarized"	Repeated words were corrected.
35	53	1277-1280	B	Please rephrase to make it more comprehensible the textual description to explain the contrast between the two plots in figure 11 (data from 1990?) and figure 10 (current data? Duration/time point?). In	The text was rephrased.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q6. Evaluate whether the documentation of the data and modeling, and interpretation of the results is appropriate. If not, please provide an alternative outline, approach, and/or suggested language for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
				addition, the figures' description and the corresponding text description can be made more in sync, and the important takeaway message will be clearer.	
36	66	1533-1539	B	Add a correct reference to this summary of the mentioned SEJ study.	The reference was corrected.
37			C	In general, this is a well-written report. However, some errors were noted during the review. The following are examples, but not an exhaustive list.	No response required. Responses to the reviewer's questions are provided below.
38	3	33	C	Remove "use of"	The suggested words were removed.
39	10		C	Table 6 is missing from the report.	Table and Figure captions were corrected throughout this document.
40	20, 41	Table 1, 10	C	The notation "var [p]" denotes variation, but it seems it actually refers to the sqrt of variation according to the estimated 95% CI.	The notation "var[p]" was replaced with the symbol $\sqrt{var[P]}$.
41	24	570	C	cfu/g? since the description here is about the comminuted chicken	The unit was corrected.
42	46	Table 11	C	When parameter estimates were presented, recommend to consider including "n". Table 11 and 12 are examples.	Sample size was included in parameter estimate tables when appropriate.
43	48	1123	C	delta_all → delta_carcass	The formula was corrected.
44			D	Please check for these errors and correct accordingly.	No response required. Responses to the reviewer's questions are provided below.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q6. Evaluate whether the documentation of the data and modeling, and interpretation of the results is appropriate. If not, please provide an alternative outline, approach, and/or suggested language for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
45	3	33	D	"….for their work to advance the use of use of whole genome sequence data for…". Please delete one "use of". It was used twice.	The typo was corrected.
46	13		D	Hazard Identification-The identification (of) biological agents capable…missing "of".	The missing word was added.
47	13		D	Limit of quantification/quantitation (LOQ) LoQ is the lowest level of microbial cells that can be quantified based on predefined goals for of confidence in the estimation. LoQ is typically higher than the LoD as estimating a numerical value requires more information than requiring a positive/negative result. Please correct "for of" by deleting "for".	The correction was made.
48	13		D	Pathogen Reduction; Hazard analysis and Critical Control Point (PR;HACCP): Please change to "Analysis" instead of "analysis".	The correction was made.
49	79	1847	D	Table 23: Not sure why "Comminuted turkey" is there in the last column of this table, in this chicken risk assessment.	Un-related results were removed.
50	101	2297	D	Please insert "to" in between "due" and "limited serotype data".	The insertion was made.
51	126	2965-2969	D	The same reference Thompson et al. in the reference list has been written twice back to back. Please delete one.	The list of reference was corrected.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q6. Evaluate whether the documentation of the data and modeling, and interpretation of the results is appropriate. If not, please provide an alternative outline, approach, and/or suggested language for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
52	170	117	D	"…..and used in their development of a dose-response mode." Please correct "mode" to "model".	The correction was made.
53	175	223-224	D	"….or 100% of servings of chicken is chicken equal to…."Please delete the extra "chicken".	The extra "chicken" has been removed.
54	176	246	D	Table 44: Daily chicken consumption in grams of chicken commodity. What is GRM4? What is superscript 4?	Thes superscript was corrected to be an "a".
55	177	260	D	Table 46: Average daily consumption for commodity domain. Same as above. What is GRM4? What is superscript 4?	Thes superscript was corrected to be an "a".
56	181	355	D	It seems there is a missing symbol (vertical rectangle).	The symbol was not missing and that has been double-checked
57	187	557	D	It seems the second "mu" and "sigma" are for turkey and not for chicken. Accordingly, the subscripts need correction.	The second "mu" and "sigma" for turkey was removed.
58	187	559 and 561	D	Please add "comma" before "respectively."	The correction was made.
59	189	588	D	"…remains contest…." It seems "contest" needs to be replaced with "constant".	The replacement was made.
60	191	644	D	"EpiX will supply beta-Poisson dose-response…….". It seems the dose-response model was already supplied.	No response required.
61	192	668-669	D	Why is the reference in bold?	Typographical errors were corrected.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q6. Evaluate whether the documentation of the data and modeling, and interpretation of the results is appropriate. If not, please provide an alternative outline, approach, and/or suggested language for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
62			E	I found the report to clearly written, with few exceptions. It is incomplete in some respects as described above (characterization of the basis for and treatment of attenuation factors). Another source of incompleteness would be the degree to which uncertainty is characterized (including simple sensitivity analyses, where possible).	A full description of the attenuation model development has been added to the "model approach" **subsection 1.5** and a sensitivity analysis has been added to **Chapter 5 Final Product Standards.**
63			E	There were a number of instances in which sample calculations might have been helpful, and for which intermediate results could have been demonstrated (e.g., overlaying the actual average dose distribution with the dose-response curves) to enhance clarity, accessibility (for the less mathematically inclined), and to provide important additional perspectives and insights into the nature of the consumer exposures.	This comment was addressed above where the suggestions were made.
b. Does the report follow a logical structure and layout?					
1			A	It largely does, but I do miss conclusions (I was quite surprised the report ended after chapter 7), you have to search for the overall approach and conclusions, even in the executive summary. It should be explained why you develop a new DR model and where and how you use it, and then evaluate whether it was worth the	A Discussion Chapter was added to the document. Further, a full description of the dose-response model development has been added to the "model approach" subsection and a sensitivity analysis has been added to **Chapter 5 Final Product Standards**.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q6. Evaluate whether the documentation of the data and modeling, and interpretation of the results is appropriate. If not, please provide an alternative outline, approach, and/or suggested language for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
				effort (would we get very different results using the "old" one-size-fits-all DR relationship?), there seems to be no place for that in the structure.	
2			B	Description of data, method, and implementation is appropriate and well documented individually in each chapter, however, the order of current chapters can be reorganized to match the proposed Risk management questions: Q1. About Receiving ~ Receiving Guidelines (Ch6), Q2. About Final product ~(Ch5) Q3. About the process control ~ (Ch7)	Organization of the report was structured to best convey the a) model development and b) the needs of the risk managers and FSIS stakeholders.
3			B	Ordering them in the order will make the structure more logically connected: Receiving Guidelines (Ch6) → Ch5 Final product (Ch5)→ Ch6 Process control (Ch7)àCh7	
4			C	Yes.	No response required.
5			D	There are many redundancies in this report. Please check this throughout the report and revise accordingly. For example:	
6	38 57	948-950	D	"As part of a cooperative agreement between FSIS and EpiX Analytics, *Salmonella*	Given the length of the document, certain key points are intentionally repeated for clarity.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q6. Evaluate whether the documentation of the data and modeling, and interpretation of the results is appropriate. If not, please provide an alternative outline, approach, and/or suggested language for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
		1334-1336		serotypes were categorized into two clusters derived from a machine learning algorithm using virulence factors to estimate the genetic similarity between serotypes."	
7	67 79	1568-1574 1834-1840	D	"It is assumed there are 125,115 chicken-associated *Salmonella* illnesses per year. This value is calculated as the product of the total number of CDC FoodNet cases per year (7,600), the share of these cases that are foodborne (66 percent) and of domestic origin (89 percent), the underdiagnosis multiplier for *Salmonella* (24.3) and dividing by the FoodNet catchment area (15 percent). These total cases are distributed across products by assuming the proportion of servings consumed (0.11, 0.83 and 0.06) of all illnesses result from exposure to carcasses (whole chickens), parts and comminuted (ground) forms of chicken, respectively."	
8			E	The overall structure of the report was appropriate. A global table of key assumptions would be a useful addition.	A table of key assumptions was added to **section 1.6 Introductory Tables and Figures**.
c. Are the conclusions supported by the risk assessment?					
1			A	Yes. But see my general comment above about the fact that the conclusions are not listed, and all my other comments.	No response required.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q6. Evaluate whether the documentation of the data and modeling, and interpretation of the results is appropriate. If not, please provide an alternative outline, approach, and/or suggested language for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
2			B	Overall findings of the current risk assessment, mainly suggesting interventions to update the final products' thresholds for all three investigated products will help avoid a considerable % of foodborne *Salmonella* illnesses as summarized in table 4.	No response required.
3			B	To properly present the effectiveness and efficiency of controlling *Salmonella*, the work can be extended to consider the uncertainty estimates in the DR models. Also, the Direct intervention costs or DALY estimates can be used to better represent the public health impacts.	A sensitivity analysis concluded that the dose-response models do not impact the final estimates enough to make further consideration of uncertainty fruitful. Cost estimates are not a part of this document, but FSIS is undertaking a full cost-benefit analysis as part of the regulatory rule-making process.
4			C	Overall, the conclusions are supported by the risk assessment findings, albeit with some limitations arising from the modelling approach and data availability (see comments above)	No response required.
5	28	686-710	D	The conclusions are reasonable given the scope of the work. Conclusions are presented only in the Executive Summary (Page 28, lines 686-710). A separate conclusion section in the main report is missing. Please include that.	**Chapter 8 Discussion** was added to the document to address this point.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q6. Evaluate whether the documentation of the data and modeling, and interpretation of the results is appropriate. If not, please provide an alternative outline, approach, and/or suggested language for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
	28-29	28-29		Limitations: The limitations of the risk assessment, data, modeling approach, results and interpretations are missing in the main report. The limitations has only been mention in the executive summary (pages 28-29, lines 711-728). Please provide the limitations in the main report.	Additional discussion was added to **Chapter 8 Discussion**, summarizing the limitations which are mentioned throughout the report.
6			E	I believe that in general, the conclusions are a fair characterization of what can (and cannot) be demonstrated by this approach to risk assessment.	No response required.
d. Is the documentation of the assumptions clear and complete?					
1			A	Documentation of the assumptions is scattered, it is very challenging to get an overview. So it is not clear and therefore I cannot judge whether it is complete. What you would need is a dedicated section or an overview of the questions, the answers to the questions (which are the conclusions), the uncertainty about the answers and an overview of the assumptions that lead to these uncertainties.	A Discussion Chapter has been added to the document that adheres to the structure proposed in this comment. A summary table documenting key assumptions has been added to the Introductory materials.
2	158-160		A	A table like Table 41 in Appendix A, that lists the assumptions and their implications for the conclusions of the risk assessment, would be helpful for each of the four risk management questions.	Such tables have been added to section **1.6 Introductory Tables and Figures**

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q6. Evaluate whether the documentation of the data and modeling, and interpretation of the results is appropriate. If not, please provide an alternative outline, approach, and/or suggested language for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
3			B	The given documentation explains mostly all the concepts in a comprehensible manner. Some explanations were found to be unclear and fuzzy, which are requested to be clarified in responses to the specific questions in corresponding sections. Overall, the work is comprehensive with respect to the available data and the implementation of the modeling approaches. All the statistical methods used for this QMRA are established and well-supported in the literature. The use of advanced bioinformatic and machine learning approaches has given additional insight and has added a novelty to the work	The clarification requests were addressed in the corresponding sections. No other response required.
4			B	The report is written well, and all areas are well addressed. The overall flow of the report is a little disconnected from the proposed Risk management questions, and it took an effort to connect the different pieces. FSIS should include a specific section to answer each risk management question.	A Discussion Chapter was added to summarize the finding from each risk management question.
5			C	While the assumptions have been clearly documented, the discussion on their impact appears insufficient and would benefit from	Responses are provided in the corresponding sections.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q6. Evaluate whether the documentation of the data and modeling, and interpretation of the results is appropriate. If not, please provide an alternative outline, approach, and/or suggested language for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
				further elaboration such as outlined in the answers to Q1-a,Q3-a,Q4-d, and Q4-e.	
6			D	Please refer to the comments in Questions 1 -5.	No response required.
7			E	Generally, yes, with the exception of those already noted	No response required.
e. Is the documented dose-response, exposure assessment, and risk characterization modelling transparent and reproducible?					
1			A	The dose response is well documented, I understand the approach and was able to reproduce the parts I analyzed in more detail.	No response required.
2			A	The exposure assessment is a combination of a distribution of what is found on the meat (line 1128 – 1130) and the attenuation distribution that was derived earlier. It is transparent and reproducible.	No response required.
3			A	As I see it, there is no risk characterization modelling performed in terms of obtaining a risk estimate from combining exposure assessment with dose response. In the approach used that is not needed, so I do not see it as a shortcoming. In the report, the analyses of intervention scenarios are considered risk characterization, but it seems the term is not used in chapters 5, 6, and 7. Please make sure that the definitions of	We have now included text that explicitly identifies the exposure distribution and risk characterization steps in our model explicitly. Risk characterization is simply the integration of a dose-response function across an exposure distribution to calculate the probability of illness per serving. Our finished product standard model requires the calculation of probability of illness per serving before (baseline) and after (new)

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q6. Evaluate whether the documentation of the data and modeling, and interpretation of the results is appropriate. If not, please provide an alternative outline, approach, and/or suggested language for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
				"exposure assessment" and "risk characterization", as they are applied in the risk assessment are well explained and in agreement with what has actually been done.	implementation of a concentration- or serotype-based standard.
4	18	427-431	A	The risk characterization mentioned here is not performed. The number of cases is derived from epidemiological data, not from the hazard characterization and exposure assessment. In principle there is nothing wrong with this approach, but I think it is not risk characterization as defined by Codex (and the glossary). The model is not calibrated to the number of cases, it is derived from it. Please clarify this in the text.	As explained above, the risk characterization step in our model is the integration of each dose-response function across the exposure distribution. These results need to be mixed to determine either a baseline or new overall probability of illness per serving. Illustrative examples of these probability of illness per serving estimates are included in our results table for both the concentration- and serotype-based finished product standards. It is true that the number of illnesses for the chicken products initially occurring (associated temporally with the baseline probability of illness per serving) is estimated exogenous to the risk assessment model. But this number is simply used to estimate the number of illnesses avoided by the various risk management options considered.
5			B	Yes, the text was adequate and the submitted scripts run correctly and are reproducible to generate correct plots and data files. Only for few parts the code-run was not possible, given the limited data/code sharing.	No response required.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q6. Evaluate whether the documentation of the data and modeling, and interpretation of the results is appropriate. If not, please provide an alternative outline, approach, and/or suggested language for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
6			B	An ideal documentation with corresponding R scripts and input datasets, that have been most accessible for the reviewer, could have been organized in a chapter wise manner. Chapters or Appendix ~ Scripts- Input data files.	No response required.
7			C	The report provides a clear explanation of the theory behind the modelling approaches used. However, to ensure reproducibility by other researchers, it would be beneficial to provide more comprehensive details about the mathematical model. Specifically, it is recommended to include tables that list how variables were integrated into the mathematical computations step by step. Some examples for consideration can be found in peer-reviewed articles, such as the risk assessment of Campylobacter in broiler chicken conducted by Swedish Food Agency, 10.1016/j.ijfoodmicro.2007.10.008, and the risk assessment of *E. coli* O157:H7 in beef products conducted by Public Health Agency of Canada, 10.1016/j.foodcont.2012.03.003. This would help to ensure that other researchers can follow and replicate the methodology used in the model, especially when R codes are not accessible	More comprehensive details about the mathematical model used in the Final Product Standards and Receiving Guidelines Chapters were outlined in a table in **section 1.6** that list the variables and their usage.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q6. Evaluate whether the documentation of the data and modeling, and interpretation of the results is appropriate. If not, please provide an alternative outline, approach, and/or suggested language for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
8			D	The documented dose-response, exposure assessment, and risk characterization modelling is transparent and reproducible but need changes as documented in comments to Questions 1-5	The changes requested by the reviewer were addressed in the comments to Question 1-5.
9			E	Having provided a significantly detailed series of equations, and source code, it is clearly transparent and reproducible, except in the few cases noted.	No response required.

Appendix A: Reviewer Information Sheet

Name	
Preferred Email	

Information on areas of expertise

Please provide an assessment of your expertise in the listed areas. It is not necessary to demonstrate expertise in all areas.

Expertise	Extensive	Medium	Minimal/None
Quantitative microbial risk assessment (e.g.: Bayesian modeling, Monte Carlo)	☐	☐	☐
R coding	☐	☐	☐
Dose response modeling	☐	☐	☐
Bioinformatics: Machine learning methods for genomic data (e.g.: random forest modeling)	☐	☐	☐
Knowledge of current laboratory methods for enumerating (e.g., qPCR/characterizing *Salmonella* with statistical analysis of test results (e.g., variability)	☐	☐	☐
Epidemiology and surveillance of salmonellosis	☐	☐	☐
Knowledge of chicken production and/or slaughter processes	☐	☐	☐
Knowledge of turkey production and/or slaughter processes	☐	☐	☐

Conflict of Interest Information

Please list current or in-pipeline projects and other relationships with the following entities. Activities listed below do not necessarily disqualify you from participation. RTI will evaluate your responses for any conflict of interest. All information you provide RTI will be kept strictly confidential.

	List of projects/relationship ⇩ and funding type ⇨	**Grant**	**Contract**
Industries that may be affected by related rules and regulations			
1		☐	☐
2		☐	☐
3		☐	☐
4		☐	☐
Organizations or associations representing above industries			
1		☐	☐
2		☐	☐
3		☐	☐
4		☐	☐
Organizations or associations that advocate specific policies regarding chicken, turkey and/or *Salmonella*			
1		☐	☐

2		☐	☐
3		☐	☐
4		☐	☐
Government agencies related to monitoring or controlling *Salmonella* in chicken and/or turkey meat			
1		☐	☐
2		☐	☐
3		☐	☐
4		☐	☐
Any other relevant information that you would like to disclose			
1			
2			
3			
4			

Appendix B: Summary of Expertise and Conflict of Interest

Tables B-1 and B-2 summarize the information obtained from experts regarding their expertise and conflict of interest using the form from Appendix A.

Table B-1. Summary of ALL POTENTIAL Peer Reviews' Expertise. Highest possible ranking is 3.

	Experts									
Expertise Related to the Peer Review	1*	2	3	4*	5	6*	7*	8*	9	10
Quantitative microbial risk assessment (e.g.: Bayesian modeling, Monte Carlo)	3	3	3	2	3	3	3	3	3	3
R coding	2	3	3	3	3	3	3	2	3	3
Dose response modeling	3	3	3	2	3	3	2	3	1	2
Bioinformatics: Machine learning methods for genomic data (e.g.: random forest modeling)	3	1	2	3	2	2	2	1	1	2
Knowledge of current laboratory methods for enumerating (e.g., qPCR)/characterizing *Salmonella* with statistical analysis of test results (e.g., variability)	3	3	2	2	3	2	3	2	2	2
Epidemiology and surveillance of salmonellosis	3	2	2	3	2	2	3	2	3	3
Knowledge of chicken production and/or slaughter processes	3	1	2	2	2	3	3	3	3	2
Knowledge of turkey production and/or slaughter processes**	2	1	2	2	2	2	3	2	2	1

*Selected experts for the peer review.

** Not relevant for this peer review

Table B-2. Years of Experience and Funding Support for ALL POTENTIAL Peer Reviewers

		Experts									
		1*	2	3	4*	5	6*	7*	8*	9	10
Years of experience in the field		>20	>20	15-20	10-15	15	>20	15-20	>20	<10	10
List of Projects/Relationships	**Industries that may be affected by related rules and regulations**										
	Organization/associations representing above industries						X				
	Organizations/associations that advocate specific policies regarding, chicken, turkey and /or *Salmonella*							X			
	Government agencies related to monitoring or controlling *Salmonella* in chicken and/or turkey meat**			X	X		X	X	X		X
	Any other relevant information that you would like to disclose	X									

*Selected experts for the peer review.

** In this category were included work done for governmental agencies in other countries, expert panels such as NACCMF, FAO/WHO, and EFSA.

Appendix C: Overview of the peer review materials for the Quantitative microbial Risk Assessment for *Salmonella* in Raw Chicken and Raw Chicken Products

- **Chicken_SRA_Review_1.30.23 1400_RTI**: PDF document describing the QMRA. This is the main document you need to review
- **FinProdStds_Jan25_23.R**: R code for the final product portion of the QMRA
- **Polynomial2:** CSV file with polynomial coefficients for dose-response model in FinProdStds_Jan25_23.R
- **APC_perf_standards_review.R**: R code for the process control portion of the QMRA
- **Chicken_carcass_exp_APC_review**: CSV file with data for process control portion of the QMRA
- **CHICKEN SAS:** Word document with the SAS code used for the NHANES serving size estimates
- **01_NCBI_Parse:** Rmarkdown describing initial processing of isolate assembly metadata
- **02_Sistr_Parse**: Rmarkdown illustrating the serovar prediction and QC check from SISTR
- **03_example_Prokka_Slurm.slurn**: Example code for gene annotation and identifying virulence factors in isolates
- **04_clustering_code_example.R**: R code for unsupervised random forest and analyzing cluster stability
- **NonRedundant_VFDB_PATRIC.faa**: Fasta file of all virulence factors considered in the algorithm
- **poultry_VF**: CSV file with the basic information/description of all virulence factors considered
- **RF_input_193**: CSV file with the presence/absence matrix used as input in the random forest
- **sal_prodigal_training.trn**: Prodigal training file on the reference *Salmonella* assembly
- **sistr_poultry_cat**: CSV file with the output resulting from the SISTR prediction algorithm