FSIS Response to Peer Review Comments on: *Quantitative Risk Assessment for Salmonella in Raw Turkey and Raw Turkey 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 microbial risk assessments for *Salmonella* in turkey 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 turkey. 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 0216627.002

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. Is the overall modeling approach used to address comminuted turkey appropriate?
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 comminuted turkey 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 comminuted turkey products)?

b. Are the data analyses 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 lots 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 turkey product leaves processing?

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

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

h. Are the conclusions drawn from the analysis appropriate?

5. **Evaluate whether the documentation of the data and modeling, and discussion, and interpretation of results is appropriate. If not, the reviewer must provide an alternative outline and/or approach 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 modeling 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 chicken products (RTI Project 0216627.003), 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 12 reviewers to determine their availability and interest in participating. They were all asked to provide an up-to-date curriculum vitae and to fill 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 time 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 with 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 chicken.

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 20, 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.

Timothy J. Johnson is a Professor in the Department of Veterinary and Biomedical Sciences at the University of Minnesota and Director of Research and Development at the Mid-Central Research and Outreach Center in Wilmar, MN. Dr. Johnson has a BS in microbiology and a PhD in molecular pathogenesis. Dr. Johnson has extensive experience in poultry microbiology, foodborne pathogens, and bioinformatics. He is intimately familiar with the poultry industry, especially turkey, and his research focuses on developing tools that enable poultry producers to rapidly identify emergent *Salmonella* strains that present an enhanced risk to cause human illness. Dr. Johnson published over 160 peer-reviewed papers, has more than 12,700 citations, and has given several presentations.

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.

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 by FSIS to the responses that were 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 Turkey and Turkey 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			A	First, I want to applaud the efforts of this report. An immense amount of work went into the modeling, procuring and extraction of data, and thorough review of existing literature. At its core, this type of work is greatly needed, and the overall modeling approaches used here appear to be sound and utilize widely accepted approaches. Most of the remaining comments in my review are critical, but they are primarily focused on the work by EpiX establishing risk multipliers based on genomic data. I do not want those negative comments to be perceived as an overall opinion of the work. This work is necessary and appreciated.	Thank you.
2	103	Table 18	A	The overall approach in this model in many ways ignores the fact that *Salmonella* ecology is always evolving. Table 18 provides data that, at face value, is compelling. However, this is based on many historical outbreaks. Many serovars are recently emerged as poultry-associated outbreak serovars (e.g. Infantis and Reading). Because these are recently emerged, in the context of all historical data their proportions will be relatively small. In contrast, continually problematic serovars (e.g. Enteritidis and Typhimurium) will dominate the outbreak landscape. As a	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
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				result, the assignment of those continually problematic serovars to cluster 1 (and assignment of emergent serovars to cluster 2) will paint a misleading picture that most outbreak strains fall in cluster 1. In fact, most of the RECENT outbreak strains likely fall in cluster 2. The modeling does not account for this important reality. A recency weight was applied, but in my opinion it was not sufficient to overcome this problem. An alternative approach would be to provide more clarity on recent versus historical outbreak serovar distributions, and to address this at the serovar level instead of cluster level.	
3	45	1061–1065	A	I'm still not clear on the outbreak definition (use of historical outbreaks and how they were defined) and how these data were used. I could go and read NORS, but most won't. It would be good to more clearly define the outbreaks which were included, which were not, number of illnesses associated with each, etc. A supplementary table describing this would be appreciated.	The text has been updated to include additional details on how the outbreak data was used in the risk assessment.
4	52	1210–1212	A	Lines 1210–1212, there is very limited data from which to draw the conclusion regarding distribution of clusters #1 and #2 in any lot, i.e. that they are 0.3 and 0.7. I do not believe that there is a robust enough dataset to generalize this observation.	Turkey analysis continues to be restricted by data limitations. However, this generalized distribution provides a starting point for risk assessment scenarios. Furthermore, we conducted a sensitivity analysis (**Chapter 5 Final 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) that considers alternative distributions and analyzed the impact on illness reduction. This analysis demonstrated that the dataset is sufficiently robust as used.
5	62	Figure 18	A	Figure 18, I again do not think there is enough isolates and establishments in these analyses to draw any generalizable conclusions regarding cluster distribution at rehang. There may not be other existing data that could help with this problem. Perhaps the inclusion of more recent data from federal sampling programs could help, but it might not. The dominance of Hadar could easily be a red herring because of this. It is a fact that *Salmonella* serovar distribution fluctuate widely by geography and time; without repeated observations over multiple years of data, I do not think FSIS should be generalizing anything. I am glad they acknowledge this later, but it is also potentially misleading to include the extensive charts and serocluster estimates on these data. Much more simply, this could be described in the text and the figures removed from the main text. Otherwise, I think it is very important to justify in the report why there is confidence in the generalizability of these data, and	Thank you for the suggestion. We have cleaned up the text and removed extra graphics to focus on the key components.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				underlying statistics to support this justification.	
6	Executive summary		A	The goals of the study outlined in the executive summary state that goals #1 and #2 ask "What is the public health impact (change in illnesses, hospitalizations, and deaths)…" – I don't see analyses in the study that actually address hospitalizations and death, only illness. This should be modified or better clarified.	This was clarified in **Chapter 8 Discussion**.
7			A	One overarching theme of this study is focused on decision-making based upon *Salmonella* subtypes. In this project, subtype was defined very broadly by two primary clusters of strains. CDC and other stakeholders in particular have stated on many occasions that they would like to focus *Salmonella* control (and possibly adulteration) at the serovar or even strain level. This project doesn't really address or align with those goals. In fact, each cluster (even if the high virulence versus low virulence holds true) contains serotypes of high risk and low risk. I think this is a major problem with the study that needs to be specifically addressed, and modified accordingly. There are a couple of ways in which this could be addressed. First, utilize other existing data to expand the approaches for establishing a risk multiplier	The goal of this project was to make practical use of available genomics data to answer the risk management questions. Current data lacks the resolution to establish meaningful risk multipliers on a serotype level for all serotypes. *Salmonella* epidemiology is a dynamic process and will continue to be explored/assessed to adjust risk multipliers as research expands on the links between strains and public health outcomes. Further details of the necessary decision points of this approach have been added to the **Chapter 2**.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				at the serovar level. Second, expand the genomic analyses to look for genes or gene sets which correlate with known serovars or strains that have been classified based upon patient outcome (this would require digging into public health data of outcome at the serovar or strain level). It may be that serovar drives patient outcome, but it is also possible that gene differences at the strain level drive patient outcome. Finally, it is possible that it will be difficult to identify gene sets which correlate with risk of infection or risk of severe outcome. In this case, reverting to the serovar level would be most appropriate. One final option would be to expand the current gene sets used to include the entire *Salmonella* pangenome, then perform the same clustering methods applied here, and then set resolution higher than two large clusters.	
8			B	Overall conceptual model implemented for risk assessment looks good and is in agreement with the standard method following all the elements of Codex Principals and guidelines.	No response required.
9			B	The study has considered the most plausible scenarios for risk management for public health benefits. However, the implementation of models was limited by the	An uncertainty analysis has been developed and added to the report, see **section 5.7**.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				scarce data availability for turkey carcasses and turkey parts. The final product scenario was only successfully implemented with comminuted turkey. Due to limited data estimation of variability in the exposure assessment has not been estimated and uncertainty in the estimates has also not been modeled.	
10			B	A novel dose-response model to focus on highly virulent serotypes dose-response and the impact of their management on public health gave an efficient way using the current available data for comminuted turkey. This model will also serve as an useful method for future proposed more testing and data procurement plans to implement efficient risk management policies for other product types.	No response required.
11			C	It is striking that so little data is available, especially on the concentrations. This type of data is very important to understand exposure to consumers and for risk assessments. I do not blame the risk assessment for that lack of data, but wonder whether it has implications for future sampling activities, or whether	We have added to the **Chapter 8 Discussion** that highlights the limited data available for the turkey risk assessment. We have also added a new *Research Needs* section **(section 8.1)** to the report.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				recommendations for the future should be made. Interestingly, line 639 refers to "Recommendations made in this risk assessment about future research", but I could not find them. So please include these recommendations or refer to the paragraph in the document where these recommendations are done.	
12	23	6641–646	C	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 developed and added to the report, see **section 5.7**.
13			D	Overall, given the scope of the risk assessment, data, assumptions, and	Specific comments are addressed below.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				analyses are reasonable and appropriate. However, some assumptions must be explained further, some analyses must be redone, and some missing references must be provided for some statements. Please see below for specific comments and details in Q1 a–c.	
14			E	Same as for chicken risk assessment: 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 the document such that reviewers, the public and the ultimate risk managers have a sufficiently clear understanding of the foundations and limitations. This summary table would benefit both the chicken and turkey risk assessments, and would also demonstrate both the key similarities and key differences between the two (e.g., the lack of critical data for turkey products, use of chicken-based attenuation values for turkey).	A table of assumptions has been added to the document, see **section 1.3 Table 7**. Areas where it was necessary to use chicken data, as turkey data was not available, have been highlighted.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: 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.					
1			A	The team did a very nice job of citing relevant works related to this study. The interpretation of those past studies is appropriate. I would like to see more consideration of previous efforts to identify associations between *Salmonella* strains/serovars and illnesses/hospitalizations. There is a lot of nice justification for the modeling approaches used, but very little background on what has been done already on this topic and how FSIS could utilize past data or assumptions. Some examples: https://academic.oup.com/jid/article/198/1/109/840110 and https://academic.oup.com/cid/article/38/Supplement_3/S149/354299 - the use of high virulence and low virulence categories extends beyond the risk of infection/illness. It also – more importantly – includes risk of severe outcome. In fact, in this document's definition of virulence it says "The ability of an organism to cause severe illness. In risk assessments, this is usually modeled as the probability of severe illness given infection. Virulence in bacteria is mediated by genes often called "virulence factors". Both pathogen and host factors contribute to whether disease occurs and to disease	We agree that more can be investigated in terms of virulence and risk of severe outcome. However, this relationship was not fully explored in the current analysis given the assumptions for developing the dose-response relationship. At the same time, available studies that compare the virulence of *Salmonella* serotypes are typically small or only include a limited number of serotypes. That is, the data is still mostly limited to the most virulent strains and this is generally captured using outbreak data. Previous work by EpiX Analytics analyzed similar seroclusters in regards to severe patient outcomes and clinical case presentations associated with *Salmonella* in beef as well as within serovar virulence subpopulations (Fenske, 2022). The pathogenesis is still not completely understood, but further links with patient outcome will help expand this research.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				severity." By this definition, the use of high virulence implies that this category can cause increased severe illness. However, the study is focused on infectivity, not severe illness. The assumption that there is a 1/1 relationship of infected/ill individuals used in this work exemplifies that as well. Severity of outcome is perhaps more important than infectivity, since the focus here is centered on virulence. Therefore, patient outcome and severity of illness should be a part of this approach.	
2			B	The current QMRA Risk assessment and suggested intervention utilize all relevant data sets curated from all national databases, managed by FSIS and CDC. The methodology used for data description and statistical analysis used for QMRA model development is standard and as per established procedures and guidelines. The assumptions made while using and modeling datasets are valid. There were a few minor textual and organizational issues that were observed and are enumerated as follows:	Specific comments are addressed below.
3	21	569–571, 572–574	B	Points made in the statements in these two-line locations 569–570 and 572–573 are similar and not very clear, please revise the language to make a coherent statement.	The two lines have been revised to clarify the points.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				In line 571, there seems a mistake in the line (italicized text): the presence of *Salmonella* at post-chill or *a log reduction in EB*	The correction has been made.
4	25	713	B	(insert bold italicized word) process—the public health impact on various risk management approaches.	The correction has been made.
5	35	919	B	Figure 3 can be replotted with similar faceting as in Figure 4 for better visualization of the data	**Figures 3** and **4** have been updated to a similar facet/style.
6	37	958	B	Which imputation method was used, is not clearly stated.	The text has been revised and a reference has been added to clarify: van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software, 45(3), 1 - 67. https://doi.org/10.18637/jss.v045.i03
7	37	969–970	B	LOD limits are given as 0.3 MPN/g and left censored samples are mentioned as (<0.03 MPN/g), was the censored limit kept as LOD/10 purposely? This line should be rephrased and checked for any typos in the numbers.	The typo was corrected, and the sentence rephrased.
8	37	970–971	B	Figure 6 legends (Y-axis density) do not match the intended description of frequency. Also in line 961, the counts of positive samples from the screening test are	The word “frequency” was deleted to correct the error.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				stated, so the figure should reflect the count data.	
9	45	1053–1054	B	Before breaking down into subsections, a line saying that section describes different data sources can also be helpful to make the flow of reading better	Thank you for the comment.
10	45	1055	B	Title Outbreak does not make any sense here,	The title was removed.
11	46	1112–1113	B	Please add specific references to this work. "More recently, CDC conducted a structured expert 1136 judgment (SEJ) to…"	The reference was inserted.
12	54	1279–1280	B	The text needs a proper reference "be used (Ebel refs):…"	The reference was inserted.
13	55	1321–1322	B	The statement does not match the given equation notation α=n/L. Not clear, should be rephrased. Looks like n should be the number of failing units, as by giving the definition alpha is a fraction of failing units that are diverted. If n is the total number of the tested unit, then, this notation suggests that all should fail. Does not make sense.	This was rephrased to clarify.
14	60	1430–1431	B	Add Basis /reference for this suggestion about the sampling method and diagnostic test's LOD.	References added.
15	70	1658	B	"a mixed compliance fraction in the case of non-mandatory standards (α= 0.5)…"	The figure was corrected.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				Assumingly, is it plotted as alpha in the graph, in blue color, clearly state	
16	76	1782	B	(Table from Pop Description); please add table number? and expand Pop?	A cross reference to the table in **Chapter 3 Microbial Profile** was added.
17			C	In general, the vast majority of referenced literature is from the U.S., much previous work from the authors of the risk assessment. This is well understandable (it is a U.S. 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 so it was better informed. This is a shortcoming, as important studies, unknown to the authors, may have been missed. 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.	FSIS conducted a literature review on *Salmonella* in poultry as part of the *Salmonella* Risk Profile alongside this risk assessment (available here). The Risk Profile was externally peer-reviewed and will accompany this risk assessment when a proposed rule is released. 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.
18			D	In general, given the limited data availability, this reviewer appreciates the efforts FSIS made in generating useful data over the	These points were addressed later in this 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.					
				years and used those in this risk assessment. The key studies and data have been identified. However, clarifications and justifications are needed for several instances and are identified later in this review.	
19	17 17 17	432 434–439 440–448	D	"Finally, FSIS does not currently enumerate turkey carcass samples." Turkey parts-FSIS currently does not assess *Salmonella* contamination in turkey parts. Comminuted turkey: "FSIS established a *Salmonella* performance standard for comminuted turkey in 2016 and the Agency also maintains a *Salmonella* sampling program for comminuted turkey. In current sampling….. FSIS comminuted turkey performance standard in a 52-week period." As an example, a few statements presented above have been mentioned, which clearly indicate that this risk assessment data, results, and analyses were mostly focused on comminuted turkey rather than raw turkey and turkey products (e.g., carcass or turkey parts). It is not clear, why the title of this risk assessment and document is "Quantitative Risk Assessment for *Salmonella* in Raw Turkey and Raw Turkey Products" rather than mentioning this is only	Effort was made to describe all raw turkey products and significant time is spent evaluating all available data. The methods outlined in this report could be applied to other products were data made available. As such, we respectfully disagree that the title of the document should be modified.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				for comminuted turkey. Accordingly, the title needs revision.	
20	18 37 134	451–454 958 301–302	D	"Enumeration data is limited to comminuted turkey samples. In particular, in PR;HACCP samples, a portion of detected samples were further tested for *Salmonella* levels using the "most probable number" (MPN) estimation method with a limit of detection (LOD) of 0.3 MPN/g. Missing results were addressed using an imputation procedure." "Missing results were addressed using an imputation procedure." "These missing results were addressed by randomly imputing sample results from those samples with results." Imputation has been mentioned multiple times as shown by the above statements. However, the procedure used for imputation has not been mentioned/described. Please include how imputing was performed.	The text has been revised to describe the procedure used for imputation and a referenced has been added to clarify: van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software, 45(3), 1 - 67. https://doi.org/10.18637/jss.v045.i03
21	18	464–466	D	"Although Enteritidis is most frequently associated with human salmonellosis, it is rarely observed (or detected) on turkey carcasses or in comminuted turkey products." Please provide reference or evidence to support this statement.	Text has been added to support this statement.
22	24	682–683	D	Risk Management Question # 4: What is the public health impact of implementing combinations of the risk management	Risk management question #4 has been addressed in **Chapter 8 Discussion**.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				options listed above? The analysis for risk management question number 4 was not performed. This reviewer could not find these analysis and results in the report related to the combined effects of risk management questions 1, 2, and/or 3 (i.e., risk management question number 4). Please include the analyses for this question in the report or delete this question from the report if the analyses have not been done.	
23	52–53	1220–1228	D	"A simplifying assumption was adopted due to the lack of complete data in terms of turkey products. To develop the dose-response functions, a lognormal distribution (Log10Normal(-3.037117, 1.279985)) was used that reflected the initial contamination of *Salmonella* in a mixture of the raw poultry products – chicken carcasses, chicken parts and comminuted chicken – according to their relative frequencies of consumption (see Chicken Risk Assessment), and an attenuation distribution that considers all the effects of partitioning, mixing, growth, and attenuation that typically occurs between poultry production and consumption defined as Log10Normal(-5,1.91) (E. Ebel & Williams, 2015). Together the initial contamination distribution and attenuation distribution constitute an (log10) exposure distribution."	We have expanded upon this assumption in the report. Lacking complete data to develop a refined model, the attenuation distribution for chicken has been used (see **section 1.5**). Furthermore, we have considered the impacts of this via sensitivity analyses (see **section 5.6**).

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 contamination of *Salmonella* in chicken does not reflect the same for turkey. For this turkey risk assessment, the authors have used the exact distribution of *Salmonella* contamination they used for chicken assessment. In case of no data availability, the use of surrogate data may be helpful. However, as the authors mentioned in this risk assessment, there are enumeration data available for comminuted turkey. Please see Page 18, lines 456–458: "The compiled comminuted turkey dataset consisted of 1,178 samples, of which 157 were positive on the screening test. The lognormal concentration distribution of the population derived from the data were estimated as μ=−4.857 and σ=2.333." Data clearly indicates that average contamination level in turkey (μ=−4.857) is less than chicken (μ=−3.037117). Hence, the dose-response function for turkey needs to be revised and redone by using the turkey data, which is already available to FSIS (please see above comment, page 18, lines 456–458).	
24	66	1577–1578	D	"A logistic regression model was fit to the data, but no significant relationship was uncovered." Please mention the significance level and elaborate the relationship between the fraction of samples with no detectable	Thank you for pointing to this line, it was incorrect. The relationship between no detectable AC and *Salmonella* proportion at post-chill has been expanded.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				APC at post-chill and the proportion positive for *Salmonella*.	
25	28	809	E	The turkey risk assessment uses the attenuation distribution derived for *Salmonella* in chicken products (Ebel and Williams, 2015) for the exposure assessment for comminuted turkey products. This is a major assumption that deserves some discussion as to the appropriateness of the assumption. This stretches the assumption used in the chicken risk assessment (which claimed a single attenuation distribution spanning whole carcasses, chicken parts and comminuted products), possibly too far. If this is somehow justified, or is essentially unimportant for some reason, this should be clearly explained. While it is stated that the attenuation distribution assumption's "derivation" is in the Appendices, this is not the case	We have expanded upon this discussion in the document (**section 1.5**) and clarified the development of the distribution. Lacking complete data to develop a refined model, the attenuation distribution for chicken has been used. Furthermore, we have considered the impacts of this via sensitivity analyses.
26			E	I am not aware of any additional studies that should have been used. The data seem to have been correctly analyzed and interpreted, except as otherwise noted.	No response required.
b. Have the strengths and limitations of the data been transparently explained?					
1			A	The team does an excellent job of pointing out the potential limitations of the data throughout the report.	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.					
2			A	The limitations are nicely addressed throughout the study, except in the executive summary where there are only two small paragraphs. Because the executive summary may be the only part read by many, it should include more about those limitations. I would also suggest that this section also includes the strong points of the work (i.e.: what data/results brings the highest level of confidence from the team?).	Added discussion of the high confidence results has been added to the **Executive Summary** and **Chapter 8 Discussion**.
3			B	Yes, quite well.	No response required.
4			B	Limitations of data availability for carcasses and comminuted turkey for prevalence are well documented.	
5			B	Limited availability of data has also been explained due to suspended surveillance activities during the COVID pandemic and limited accessibility of FSIS to turkey surveillance at carcasses and final product and its impact on the robustness of the model is well explained.	
6			B	The absence of uncertainty analysis has also been mentioned.	An uncertainty analysis has been added to the document (**section 5.7**).
7			B	The strength of this work over usual Risk assessment models is the use of WGS-based virulence determination of virulent serotypes. This gave non-obvious insight as more virulent serotypes are not the highest reported serotypes or most prevalent in	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.					
				numbers. The segregation of virulent/non-virulent serotypes is insightful as ultimately, it's the clinical infectiousness of the pathogen is important as it is directly related to the impact on public health. The identification of virulent serotypes via genomic data clustering can be utilized for more effective risk management for better public health outcomes.	
8			C	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. Adding this would increase the transparency.	An assumption table has been added to **Chapter 1 Introduction**. There is also a summarized description of the main data sources in **Chapter 3 Microbial Profile**.
9			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.
10	19	505–508	D	"Additionally, in the FSIS 2008–2009 young turkey carcass baseline study, less than 5% of rehang samples and less than 1% of post-chill samples were quantifiable. The lack of ample data limits FSIS' ability to assess a concentration threshold performance standard for turkey carcasses."	Thank you for the comment. Future data collection is beyond the scope of this document. This data need has been addressed in the new *Research Gaps* **section 8.1** of 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.					
				As the reliable estimation of concentration distribution of *Salmonella* on post-chill carcasses is not possible currently, how FSIS is planning to overcome this in the future. If possible, FSIS should plan on collecting the needed quantitative data for *Salmonella* in turkey carcasses, which would be helpful in the future risk assessment.	
11	21	572–576	D	"As a result of these weak relationships between indicator organisms (i.e., APC and EB) and *Salmonella* prevalence, it follows that the correlation between APC or EB and *Salmonella* serotypes or levels is also weak. Therefore, it was not possible to assess the risk management question regarding the public health impact (illnesses, hospitalizations, and deaths) of monitoring/enforcing process control from rehang to post-chill in the same manner as it was estimated for final product standards." It is appreciated that the low correlation is mentioned; however, it is not clear for what final products the correlation was estimated/used; please explain.	As outlined in the document, it was only possible to consider process control for carcasses.
12	23	626–629	D	"The overall lack of turkey data, as compared to chicken sampling data, does limit the ability of risk managers to shape FSIS policy regarding *Salmonella* illnesses from turkey products. That said, the final	The data limitations and interpretation of the results have been included in **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.					
				product standards developed for comminuted turkey were developed using more extensive data, and they can therefore be considered with greater confidence by risk managers." This reviewer appreciates mentioning about this data limitation and interpreting the results. Please emphasize this in the data analysis and results interpretation and clearly in the report, so that there is no confusion about this risk assessment focus was toward comminuted turkey, not for others such as carcass and parts, and the results and conclusion should be used accordingly with caution.	
13	23	647–649	D	"The lack of data to determine within lot and between lot variability of bacterial occurrence and levels severely limits the ability to assess the effects of diversion options, in particular for the final product standards in this risk assessment." Please explain why this is important and how FSIS is planning to handle this in the future. Research has shown that fecal sampling is an effective way to determine *Salmonella* in turkey flocks (Arnold et al. 2009. *Journal of Applied Microbiology*. DOI: 10.1111/j.1365-2672.2009.04273.x). This reviewer thinks this might be an effective approach to determine bacterial contamination in turkey flocks.	This has been outlined in the new *Research Gaps* section (**8.1**). Fecal sampling is an excellent resource for consideration of this problem. However, it cannot be used to overcome the primary data limitation in assessing final product standard: the implausibly low *Salmonella* recovery on post-chill carcasses.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
14	34	896–899	D	"However, it is likely that the low rate of *Salmonella* recovery from turkey carcasses is a function of sponge sampling and rinsate limitations than truly indicative of the true *Salmonella* prevalence in turkey and such conclusions should be applied judiciously." This is an important sampling concern. How the sponge sampling will affect the determination of the presence of *Salmonella* and subsequent risk estimation. Please elaborate. This reviewer suggests the use of alternative sampling methods for turkey carcass in the future. Here is one approach, the authors may consider: McEvoy et al. 2005. *Journal of Food Protection*. DOI: 10.4315/0362-028x-68.1.34: The authors compared a carcass rinse, modified rinse with carcass supported in a swing, whole carcass (inner and outer) swab, one- and two-site swabs and excision of skin tissue to identify best *Salmonella* recovery method in turkey carcass. The researchers suggested, "whole-carcass sampling by rinsing or swabbing is necessary for optimum *Salmonella* recovery."	Thank you for providing this additional information. *Salmonella* sampling in turkey carcasses was identified as an area for future work in the Research Gaps section (**8.1**) of this document. However, providing alternative suggestions on sampling methodology are beyond the scope of this document.
15	36 49	933–936 1153–1159	D	How the random sampling and sponge sampling affected *Salmonella* contamination determination in carcass?	In short, no carcass determinations could be made due to the sampling limitations outlined. This, however, is a significant research gap in the model and, as such,

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
					has been added to the Discussion Chapter and to the new Research Gaps section (**8.1**) of the document.
16			E	The inability to quantify the levels of *Salmonella* in turkey carcasses and turkey parts due to a significant lack of data is clearly explained and justified. The decision to choose not to quantify the impacts related to quality control processes and serovar-based lot rejection is appropriate to the lack of quantifiable evidence.	No response required.
17			E	As discussed under Overall Comments, the fact that the exposure assessment for turkey products is based on chicken products is not very transparently explained.	Additional text explaining the use of chicken data in the turkey products exposure assessment has been added to **section 5.3 Modeling Approach**.
c. Is the overall modeling approach used to address comminuted turkey appropriate?					
1			A	Given that there is limited data with which to work, the team has done a nice job with their conceptual model using pieces of data that are available. I have no real concerns with part 1 of the work addressing control points, and utilizing *Salmonella* prevalence/load plus APC and Enterobacterial counts as criteria for assessment. This all fits and makes sense. The team has considered a lot of factors into this model which is appreciated, such as growth and die-off, the impact of older versus newer data, etc. They have used	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.					
				appropriate math and previous data to establish some of these numbers.	
2	55	1300–1323	B	The model is clearly explained and accurately implemented in R. Checked and runs perfectly.	No response required.
3	52, 55	1225–1226 and 1291–1292	B	The use of lognormal distribution to represent the heterogenous serotype population is adequate. The calculation of the dose distribution as a combination of initial concentration modeled with lognormal distribution is also appropriate. While 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.	Additional text has been added to **section 1.5 Model Approach** discussing the usage and limitations of the attenuation distribution. An assumption table has been added to the new **section 1.6** describing the limitations and a sensitivity analysis (**section 5.6**) was developed that further highlights the utility of this assumption.
4	55	1298–1299	B	Further use of integral equations for calculating the conditional probability of illness in different pass/ fail and comparatives scenarios is legitimate. Modeling is appropriately described and	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.					
				referenced, and R code is provided which runs correctly.	
5			C	In general, yes, but see my points below.	No response required.
6	19	490–501	C	I guess Table 3 refers to Figure 17. It would be helpful if that is clarified. It makes me wonder why an LOD of 0.1 CFU/g is not considered, as that should reduce the number of illnesses most, according to the Figure. The authors should include this LOD or explain (in the document) why this not included.	The 0.1 cfu/g scenario has been added to Table 3 for completeness.
7	52–53	1220–1227	C	The distribution of concentrations (μ =-4.857 and σ =2.333, line 975 p. 37) in comminuted turkey should have an impact on the exposure assessment, but it is not clear to me how it has been used. Due to a lack of data, the exposure assessment used to derive the DR relation is identical to that from chicken, not the one from turkey. Please clarify	We have expanded upon this assumption in the document. Lacking complete data to develop a refined model, the attenuation distribution for chicken has been used. Furthermore, we have considered the impacts of this via sensitivity analyses.
8			D	The overall modeling approach used for comminuted turkey is reasonable and appropriate.	No response necessary.
9	22	614–618	D	"This quantitative risk assessment examines the relationship—where feasible— between the amount of *Salmonella* and/or presence of certain *Salmonella* serotypes on turkey received for slaughter and/or on turkey products (i.e., carcasses, turkey parts, and	An expanded discussion of the high confidence results has been added throughout the **Executive Summary** and to **Chapter 8 Discussion**. Effort was made to describe all raw turkey products and significant time is spent

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				comminuted turkey) and the probability of foodborne illness. It also examines the relationship between changes in microbiological indicator organisms (i.e., APC) on turkey carcasses from rehang to post-chill and changes in foodborne illnesses." It would be nice to clearly mention given the data limitation, for what product estimating the risk was feasible using quantitative risk assessment. As the analysis was mostly focused on comminuted turkey, it may be better to emphasize that and perhaps, to consider changing the title of the risk assessment.	evaluating all available data. As such, we respectfully disagree that the title of the document should be modified.
10	23	642–646	D	"Typically, quantitative risk assessments include an uncertainty analysis. The intent of this analysis was to provide a high-level comparison of the effectiveness of controlling *Salmonella* levels and/or more virulent serotypes and ensuring process control during poultry slaughter. A more in-depth analysis of these options and the uncertainty around the point-estimates will be explored in subsequent analyses." Please explain what are the possible more in-depth analysis FSIS is considering? FSIS should consider performing/testing different uncertainty analyses for the situations identified in the above statement. As an example, in Lambertini et al. 2019, uncertainty analyses were conducted to	FSIS has conducted and included in the document an uncertainty analysis (**section 5.7**).

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q1: Please evaluate the available data and the underlying assumptions used in this risk assessment.					
				assess the impact of different thresholds for *Salmonella* concentration assumptions on model outcomes. They also carried out scenario analyses to determine the impact of various processing and post-processing-related interventions on final *Salmonella* concentrations (Lambertini et al. 2019. Microbial Risk Analysis. DOI: https://doi.org/10.1016/j.mran.2019.06.002)	
11			E	The overall approach to address comminuted turkey products is largely appropriate. Use of a lognormal distribution, despite the highly censored data, is appropriate and justifiable. The use of a calibrated risk assessment, to answer the specific risk management questions related to changes to the initial contamination levels in raw products, based on lot rejection and replacement. The main threat to "appropriateness" is the use of an attenuation distribution for chicken. To the extent that the conclusions are not highly sensitive to this assumption, which is entirely possible, this should be clearly explained. The assumption of replacing rejected/diverted lots with an average lot, is appropriate, and arguably superior to alternate assumptions of the impact of diversion.	Thank you for the feedback. A sensitivity analysis has been included in the document (**section 5.6**).

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			A	There are many problems with using the proposed database of virulence genes to cluster and make assumptions about relative risk to human infection and severe illness. My biggest concern is that clustering these isolates by "virulence gene profile" - or even virulence gene profile plus core genes from one reference genome – results in a very similar tree if isolates were instead phylogenetically analyzed based upon core genome SNP or core genome MLST analysis. This is in part why serovars such as Infantis, Heibelberg, Reading, Kentucky, etc. fell under cluster #2 (lower virulence) and Enteritidis, Typhimurium, etc. fall into cluster #1.	This analysis is driven by the presence/absence of those Virulence Factors/genes (VF) that are informative for clustering. Another analysis may cluster serotypes differently. In this work, the clusters were validated by linking to epidemiological data (documented outbreaks attributed to poultry sources). Strains with higher association with outbreaks were grouped together in cluster 1. To test whether the clustering approach performed by EpiX Analytics and other methods produce similar groupings, we compared the k=4 cluster results from the risk assessment with results obtained using reference free SNPs (Timme[1]), core genome (Worley[2]), or O-antigen grouping[3], see **Table 12**. Clustering in this report distinguished groups similarly, in general, to those derived using core genome and SNP clusters. For example, serotypes in Timme group A2 were divided into Clusters 1 and 2, and Worley group A was divided into Clusters 1, 2, and 3. O-group D1 included serotypes in Cluster 1 (Dublin

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 Enteritidis) and Cluster 2 (Javiana). O-group C2-C3 included serotypes in Cluster 1 (Hadar, Muenchen, and Newport) and Cluster 3 (Kentucky), and O-group B included serotypes in Cluster 1 (I 4,[5],12:i:-, Newport and Saintpaul) and Cluster 2 (Heidelberg, Reading and Schwarzengrund). Additionally, although the k=2 scenario was selected for the risk assessment, two groups are not necessitated by the method; further differentiation is possible (k>2), but it is not necessarily desirable or stable.
				The strain classification system ignores the innate serovar-specific abilities of *Salmonella* strains to cause disease in humans, which is likely unrelated to the gene sets assessed in the clustering approach. For example, a large part of a gram negative pathogen's innate virulence potential (particularly E. coli and *Salmonella*) is due to its cell wall and outer membrane composition. The use of the VFDB plus a single reference genome's amino acid sequences will miss serovar-specific cell wall / LPS differences. If these genomic differences correlate to an innate ability of a serovar to be cause disease, the entire clustering approach is invalid.	[1] Timme, R. E., et al. (2013). "Phylogenetic diversity of the enteric pathogen *Salmonella* enterica subsp. enterica inferred from genome-wide reference-free SNP characters." Genome Biol Evol 5(11): 2109-2123; [2] Worley, J., et al. (2018). "*Salmonella* enterica Phylogeny Based on Whole-Genome Sequencing Reveals Two New Clades and Novel Patterns of Horizontally Acquired Genetic Elements." Mbio **9**(6).; [3] WHO Antigenic Formulae of the *Salmonella* Serovars

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:					
				Undoubtedly, within cluster 2 there will be a range of serovars and even strain types with different infectious doses, different invasiveness, different abilities to colonize, etc. This is well established, and the limited literature available indicates that some of the serovars listed in cluster 2 are more invasive and possibly more infective than others (including some of those in cluster 1; see Teunis publication). This is counterintuitive. I appreciate the effort to look at proportions of illnesses relative to proportions of what is recovered from the product for use in this approach, but in my opinion it is way more complex than that. Therefore my biggest issue in my comments throughout center around the lack of resolution using two broad clusters. This is extremely important because the entire dose-response model rides on the multipliers used	https://www.pasteur.fr/sites/default/files/veng_0.pdf The cluster differences were linked to the ratio of the subtype or serotype among isolates from outbreaks and food. For FSIS, understanding virulence differences between serovars is very important for more focused risk management strategies targeting serovars with a larger inordinate impact on public health. For cluster 1 (C1), this ratio was much larger and significantly greater than 1, meaning higher than expected association with outbreaks. The comment assumes that the VFDB and BV-BRC (formerly, PATRIC) databases will not include serovar-specific cell wall/LPS differences. However, genes such as rfb I, G, and H, which regulate biosynthesis of the O antigen chain, were included in the analysis. For further clarification, FSIS developed Bioinformatics Supplemental Materials (available here), which includes a "Virulence Factor" section and note that the role has not been considered in this clustering approach; nevertheless, some genes that perform these functions have been included.

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:					
				from genomic analyses. Please see my comments in Q1 about suggested ways to improve this analysis. These suggestions are not all-encompassing, merely some thoughts about how better resolution could be obtained towards a more accurate risk multiplier.	We appreciate your concern about the ability of the clustering method to resolve strains with higher and lower infectious doses, invasiveness, etc. Indeed, it would be valid to measure properties associated with individual serotypes or clonal lineages. However, the genetic basis of *Salmonella* virulence has not been fully elucidated and is likely to be complex. The current work sought to develop a framework for assigning risk multipliers relevant to public health risk management. The identified clusters contain multiple serotypes, and some may not contribute substantially to poultry related illnesses. Cluster 1 (C1) contains about a third of the poultry isolates and almost three quarters of the poultry related illnesses, resulting in a risk multiplier that is significantly larger than the other clusters. Further, when resolved to k=4 clusters, thereby constructing clusters that exclusively contain serotypes Infantis and Kentucky individually, the C1 risk multiplier remained about the same (2.1), while the Infantis and Kentucky clusters had lower risk multipliers (0.31 and 0.094, respectively).

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:					
					One of the underlying assumptions of this work was that isolates and data collected in Teunis are representative of isolates causing illness in the U.S. Furthermore, in the absence of similar studies focused solely in the U.S., we have assumed that Tenuis provides the best estimate available for a dose response model relevant to C1 serovars, as C1 serovars are predominantly those used in the Teunis model. Finally, improving resolution any further, such as by increasing the number of clusters to k > 4 made the clusters unstable. A more thorough analysis of subpopulations within each cluster or serotype (such as the one conducted in Fenske (2022)) could yield additional information but would be more problematic for modeling risk management options. Therefore, the analysis may not be sufficiently sensitive to detect highly infectious subtypes rarely detected in poultry products, which supports future runs of the algorithm routinely to detect emerging strains of public health importance.
2			A	The genomic clustering occurs prior to use of epidemiological data, so it assumes that serovars similar to one another based on genetic	A number of issues are raised in comment Q2. A2.1.

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:					
				clustering will behave the same in humans. Then, clusters are lumped together for the risk multiplier using additional observational data from surveillance. I don't follow the rationale for this approach. Furthermore, this approach will ignore small-scale differences which are known to have a big impact on strain-level virulence. The *Salmonella* Reading and Infantis outbreak strains are perfect examples. Infantis acquired a large plasmid which is established as giving those strains enhance infectivity and virulence. However, other serovars lack this plasmid. The clustering approach at this resolution will ignore this and cluster Infantis with other lineages in cluster 2, even though plasmid-containing Infantis strains are clearly more virulent than other strains in that cluster. The critical plasmid genes would not be relevant using this approach. In Reading, small scale mutations (gene deletions and single gene acquisitions) occurred which gave the recent outbreak strains a competitive advantage in the gut and during the invasion processes. Again, these mutations would be missed with this approach. In both cases, contemporary strains of Infantis and Reading clearly should be classified as high virulence and they are not. Clearly the clustering approach at higher resolution (k=4) did identify pESI genes as differentiating between strains of	First, the rationale for this approach. In contrast to the reviewer's point, the clustering method relied on genes lost or gained in the isolate collection, not phylogenetic similarity as measured by SNPs, core genes or O-antigen genes (**Table 12**). Clustering was also agnostic to the virulence factor's biological function. Second, this clustering method was insensitive to point mutations and insertions/deletions. These mutations can modify gene function resulting in public health risk, as illustrated by the emergence of *Salmonella* Reading. Indeed, *Salmonella* Reading was assigned to the lower virulence cluster 2. However, 2% of the strains (N = 26, see **Table 32)** were assigned to the higher risk cluster 1, which may indicate the emergence of a higher risk Reading subtype distinguished by gene gain/loss. Third, the risk of illness associated with plasmids bearing strains of *Salmonella* Infantis is unclear. When distinguished into k=4 clusters, both cluster 2 and 3, containing Infantis with and without the pESI plasmid, had low-risk multipliers. Further analysis was included in the *Chicken Risk Assessment* as not only are

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:					
				plasmid-containing and plasmid-lacking Infantis. [This brought another question about the methods because the methods do not describe use of a complete pangenome for clustering – I assume that the inclusion of pESI identification using clustering is based only on Yersiniabactin which happened to be in the VFDB?] However, ultimately the lower resolution approach was pursued. But, it demonstrates that there are indeed cluster 2 isolates that very likely have high virulence potential and are misclassified as low virulence potential. On a side note, the authors should also acknowledge that pESI contains more than just Yersiniabactin. It has numerous proven or predicted VFs which are relevant to this question. There are several studies demonstrating the relatively high virulence of this clade (PMID 24320043/DOI: 10.1111/1462-2920.12351, PMID 34197273/ DOI: 10.1080/22221751.2021.1951124). I think this relates to my comment above that only Yesiniabactin was in the database ultimately used for clustering.	consumption rates higher in chicken compared to turkey, but the proportion of Infantis in broilers roughly tripled between 2016 – 2019 whereas comminuted turkey observed a rise in Infantis detections later in 2020-2021. The delay in rising Infantis detections in comminuted turkey restricts the comparative analysis with illness data which generally lags behind detections in raw commodity samples. Nevertheless, the increase in proportion of Infantis in chicken is not proportional to changes in the FoodNet Infantis case rate, as would be expected if consumption of chicken were a major contributor to illnesses (*Chapter 2 of Chicken Risk Assessment*). The proliferation of pESI-containing strains may be due to genes that convey selective advantage to growth in poultry or poultry environments.
3	91	140–141	A	Lines 140–141, "To remediate this, we sought to identify genomic markers which correspond to virulence potential" This statement implies that all genes in this database have been shown to correspond to virulence potential. They have not.	Virulence genes in *Salmonella* are heavily influenced by gene acquisition facilitated by horizontal gene transfer and gene loss through pseudo-gene formation. Therefore, gene acquisition and loss were

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:					
				Furthermore, they are in no way weighted to reflect that some genes play a much larger role in virulence than others. Because of this, the use of clustering and summative data regarding the carriage of these putative VFs is flawed from the beginning. If the focus is on true virulence factors, then the ones chosen should have been experimentally validated (either phenotypically or genomically) as VFs. See https://www.sciencedirect.com/science/article/pii/S0023643821018545 and https://www.sciencedirect.com/science/article/abs/pii/S0963996921007171 as examples of genomic approaches previously used to identify VFs correlated with *Salmonella* disease.	useful for clustering serovars. The unsupervised random forest algorithm is agnostic to the biological meaning of the VFs, and thus, strains were clustered based on similarity. The approach was validated using epidemiological information (relative proportion of each strain among poultry outbreaks and in poultry product samples). The approach described by the reviewer would also need to be validated.
4			A	In addition to seeing the justification for the choice of two seroclusters versus k=3 or k=4, there should be more ample justification for the choice of the genomic methods employed over other options. Many other options for utilizing the genome-based data exist. It was acknowledged in several instances that certain methods were employed to meet the deadlines set by FSIS. This implies that other genomic options were certainly available, but could not be effectively employed due to time constraints. In fact, the methods employed are what I would consider to be a very low-resolution genomic approach. Other phylogenetic methods, methods using the	The goal was to develop a computationally efficient genomic approach that could use the largest number possible of isolates to group *Salmonella* serovars by their risk to human health, and to link these differences to a dose response model. This analysis was integrated into a larger model that was used to estimate the impact of FSIS risk management options. Because poultry associated serovars may change over time, it was important that the method be repeatable so that FSIS could reanalyze data over time. The pangenome contains ubiquitous and lineage-specific features.

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:					
				entire pan genome, methods using SNP-based mutational data, methods examining functional versa non-functional genes, and others, could have been tested. The corresponding biological data from literature on limited serovar infectivity and illness for more serovars could have been employed to get more accurate estimates of serovar-level capacity. I think this warrants an explanation on why this specific approach was chosen over others that may be more complex and time consuming, but ultimately more accurate and informative? The assignment of risk clusters in this project will likely have long-lasting precedent. I have highlighted in my comments many issues I have with this approach, and why I believe it is flawed. Given the potential impact of setting this precedent, I think more thought needs to be given to the underlying reasons why this approach was taken in the first place. Two options exist to remedy this situation: 1) expanding the genomic efforts to better explore correlations between genotype and patient outcome, as highlighted above; or 2) utilizing biological data available from additional serovars to expand the model beyond Typhimurium and Enteritidis biological data.	The unsupervised random forest (URF) clustering method focused on lineage-specific features (i.e., VFs present in at least 10 assemblies and no more than 95% of assemblies) to maximize clustering. These include horizontally- and vertically-transmitted genes. A phylogenetic tree based on SNPs or core genes may not have captured the epidemiological differences of this method. EpiX Analytics did limit the strains used for URF to those with assemblies in NCBI to avoid having to assemble all genome sequences. They assumed no systematic bias between strains preassembled in NCBI. The method is bioinformatically parsimonious but was not chosen to meet FSIS deadlines, as this clustering methodology was already independently developed by EpiX Analytics (Fenske, 2022).
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	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:					
				perspective of public health as well as the product industry.	
6	92	198–201	B	Were all the virulence factors considered, or some filtration criteria was used?? When compiling the virulence factor database. Please clarify that in the text.	Initially, all virulence factors (VFs) from the VFDB and BV-BRC (formerly, PATRIC) databases, including genes from *Salmonella* and other Enterobacteriaceae were considered. The data were made non-redundant by removing duplicates with 90% or greater similarity. Lastly, the data were filtered to include only those VFs present in no more than 95% and in at least 10 assemblies. Based on these criteria, there were only 193 VFs variably present in the 36,647 enterica assemblies representing human clinical cases in the US and poultry and beef associated isolates. To further elucidate the clustering process, FSIS developed *Bioinformatics Supplemental Materials* (available here).
7	92	201–203	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.	Open reading frames (ORFs) from *Salmonella* Typhimurium reference strain LT2 was used to help identify and annotate isolates against the virulence

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:					
8	92	202–203	B	Not a proper rationale or reference is provided to support the choice of the reference genome of Typhimurium serotypes, for this task.	factors (VFs) forming 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 stain is also commonly used in *Salmonella* genomics.
9	92	206–207	B	Not very clear, 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 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 the VF factors or (2) processing the annotation on the isolate assemblies to determine presence/absence of each VF to identify

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:					
					the gene profiles. Text in **Appendix A** has been amended to clarify this process.
10	94	256	B	"Using NORS and FSIS data, we estimated RRi"; no code/data to review this step.	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 Federal Register at 8460). For this reason, a portion of the work that was conducted under the Cooperative Agreement (with an external private sector collaborator) was not made available to the peer reviewers, per the OMB exemption. Nonetheless, all the methods were fully documented in **Appendix A** of this FSIS risk assessment report.
11	94	260–263	B	No data set for these variables' names, is mentioned in the text in these lines. Please provide them.	These refer to FSIS sampling projects which is available on FSIS's web site, see Laboratory Sampling Data https://www.fsis.usda.gov/science-

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:					
					data/data-sets-visualizations/laboratory-sampling-data
12	94	264	B	Not sure where to find the code of the implementation process, could not find it in the provided R code. So I was not able to check it.	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 Federal Register at 8460). For this reason, a portion of the work that was conducted under the Cooperative Agreement (with an external private sector collaborator) was not made available to the peer reviewers, per the OMB exemption. Nonetheless, all the methods were fully documented in **Appendix A** of this FSIS risk assessment report.
13	98	416–421	B	Model assumption and description is appropriate. However, no R code for these steps were provided and so this reviewer was not able to check its reproducibility.	
14	103–104	478–516	B	There are certain gaps in sharing the scripts for estimations as described in the text, like: Implementation of multiplier calculation after the random forest reported and tabularized in these lines (478–516). However, later added sample code provided proof of concept, that adequately connected textual concept and implementation. Nimble implementation was not shared, the textual concept is clear and adequate.	
15	104–105	518–524	B	Sensitivity estimates of the risk multipliers, the text explains it adequately in concept, R script implementation was not provided earlier, example code was shared later, shows the proof of concept appropriately.	

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:					
16	106–108	550–579	B	No R code for the DR model fitting with variability and uncertainty calculations/plotting as in Figure 33–35, example code shared later and is good to show the performed steps.	
17	117		B	*S1: Outbreak attribution*: No code for outbreak attribution model implementation/figures was provided, as explained in the S1 section. Please include it	
18	88–99		C	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 is required.
19	95	256–258	C	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 256–258, 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). An explanation of the added value of the bioinformatics clustering is missing, 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 Factor profiles, 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 https://www.medrxiv.org/content/10.1101/2022.12.13.22283417v1which exhibited

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:					
					differences by cluster (see **Figure 4** and **Supplemental Table S4**). Although the work could have begun with serotypes, FSIS sought to avoid ignoring the 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 focused on a single subtype, which was not appropriate for this risk assessment.
20	Appendix A		C	It is quite hard to read the report in Appendix A, it is difficult to differentiate between headings and subheadings and the terminology is not always consistent. For example, equation 1 on p. 94 describes the probability that strain s belongs to cluster Ci as $Pr(s \in C_i)$ 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 process in the new *Bioinformatics Supplemental Materials (available here).*
21			C	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	This reviewer is highlighting the conclusion (**Chapter 5 Final Product Standards**) that imposing serotype- or serocluster-based standards on turkey carcasses doesn't

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:					
				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). It shows; however, that you cannot do that (lines 536–543), which makes me wonder what this bioinformatic clustering actually adds to the risk assessment. I doubt whether more detailed analyses are needed here. It is very hard to perform the clustering for the scenario analyses, based on limited data. This seems to have much more impact on the uncertainty than details of the bioinformatics analysis.	work because of limited positive post-chill data, and subsequently, it makes an accurate representation of the underlying distribution of seroclusters within a turkey flock infeasible. Additional data is required for comminuted turkey from two sample locations per lot to potentially appreciate further use of the different clusters in terms of risk management.
22			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 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).	Response provided below.
23			E	(No comment from this reviewer)	No response is necessary.
a. Was the *Salmonella* genomics data appropriately curated and processed?					

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:					
1			A	No. See my comments below on concerns about the database methods. In short, many of the curations in the spreadsheet provided are not accurate annotations. They are lifted from NCBI and suffer from extensive genome rot, where function gets assigned based on very little similarity.	Virulence factors from the Enterobacteriaceae family were considered as these are more peripheral markers that may correspond to pathogenesis while also providing the ability to find differences between serovars that the core genome would not uncover. Other methods as mentioned by the reviewer may be more robust for defining and explaining *Salmonella* virulence, however, the underlying goal of this approach was to determine groups based on infectivity/virulent traits that could be analyzed using the largest possible number of isolates and validated post-hoc using epidemiological data.
2			A	See my comments below, but I don't understand why numerous well-known E. coli virulence factors (not typically observed in *Salmonella*) were included in this database? What about SPIs? While the VFDB includes some known virulence factors, we have not fully functionally characterized every gene within every *Salmonella* genomic island. Many of these genes (not included in the VFDB) may be important towards virulence, even though they have not yet been characterized. A better approach would have been to consider the pan-accessory genome, including every SPI and its genomic context. This undoubtedly would have provided further separation of serovars into clusters and would have been more accurate. The use of VFDB plus a reference genome is a "quick and easy" way to develop clustering data for thousands of genomes. Other approaches mentioned above (e.g.: phylogenetic methods, methods using the entire pan genome, methods using SNP-based mutational data, and methods	

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:					
				examining functional versa non-functional genes) would have resulted in better data and should instead be used.	
3			B	Genomic data was properly collected from all available reliable sources, including the latest data from all FSIS (HACCP, NARMS), and FDA(NARMS) databases, Clinical cases were also included from CDC PulseNet and NORS databases.	No response is necessary.
4			B	Filtering and selection criteria to curate the complete, clean, and relevant datasets for the objective are clearly defined and well implemented in R.	No response is necessary.
5			C	I am not able to judge that	No response is necessary.
6	89 91 91–92	100 163 170–175	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 turkey. 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	The EpiX Analytics team assumed that the clustering results would not depend on the species where the isolates originated from (see **Table 54**). 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

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:					
				sources – 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.	stable cluster designations (k=2, 3, 4) with more isolate stability for less common 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 respectively disagrees with the need to conduct a new clustering analysis.
6			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	There are concerns on the generation of the custom VF database used and the validity of the hits recorded. I downloaded the 193 protein sequences deposited as those used for clustering analyses. At first glance something appeared incorrect, as many of the gene annotations for the 193 protein sequences were annotated as genes typically found in E. coli, not *Salmonella*. For example, Afa, CS17, F17, Fae	This analysis was driven by presence/absence of putative virulence factors (VFs) informative for clustering without reference to their biological meaning or function. The VF database was expanded to include other members of Enterobacteriaceae (i.e., *Salmonella*, *Escherichia*, *Shigella*, and *Yersinia*). Open reading frames (ORFs) with at least 90%

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:					
				(K88 or K99), Pap, Tsh, and others. I then took the representative sequences for some of these proteins (column A of sheet) and searched the corresponding representative amino acid sequences found in the non-redundant_VFDB_PATRIC.faa file. A few examples: 1. AAS89777 csbA CS17 fimbrial subunit A search of the corresponding AA sequence for this against NR *Salmonella* proteins in NCBI found that the closest match to *Salmonella* had only 40% sequence identity. There was a clear match to E. coli sequences (100%). This protein was found in a large % of study isolates so I would expect database matches for *Salmonella*. CS17 is an ETEC virulence factor, not *Salmonella* VF. 2. APECO1_O1CoBM73 tsh Tsh In this case there are actual hits to Tsh for *Salmonella* in the NCBI database. This is a well known avian E. coli gene, and it resides on a plasmid. This plasmid has been shown to be present in some *Salmonella* strains, but a role for Tsh in *Salmonella* virulence has never been established. This exemplifies the problem of looking at rare genes that are on plasmids and have no established role in *Salmonella* virulence. 3. NP_755467 papA P pilus major subunit PapA	protein identity to the non-redundant VF database were identified in *Salmonella* assemblies, therefore matches with other Enterobacteriaceae are not unexpected. The 3 particular VF examples discussed in this comment (csbA, tsh, papA) are indeed *E. coli* VF and rarely present in the isolate assemblies considered in the unsupervised random forest (URF). In particular, csbA was annotated in 16 isolates, tsh in 1,028, and papA in 51 out of 36,647 isolate assemblies. Each of these VF surpassed the minimum threshold requirement to be included in the analysis. See FSIS *Bioinformatics Supplemental Materials* (available here) for further detail. As observed in a previous study conducted by EpiX Analytics (Fenske, 2022), many of the most influential VFs were originally annotated from *Escherichia coli* and *Salmonella* genomes. In terms of the 3 *E. coli* VFs identified by the reviewer, only 2 (csbA and papA) were included in that previous analysis, however, these were much less informative to the cluster

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:					
				Closest match in *Salmonella* was type I fimbrial protein at 56%. No pap proteins were identified in *Salmonella*. In this case there is clearly no matching *Salmonella* containing the pap operon, which is found in ExPEC and related to UTI. I am unsure how this could have been identified as a core VF in a subset of *Salmonella*?	derivation (ranking near the bottom of the 183 VFs considered).
2			A	There are limited virulence gene databases available for *Salmonella*. The one chosen for these analyses, at the time of its application, was likely one of the only choices available without extensive manual curation. However, groups at FDA have openly acknowledged that this VF database is insufficient. In fact, they have recently released their own *Salmonella* VF database, which is much improved both in terms of curation and genomic context (https://virulence.preprod.fda.gov/). Why was this not communicated to the group at EpiX, when it is well known even to academics that it has been utilized for several years, even for publication purposes (e.g. PMID 35960531/ DOI: 10.1089/fpd.2022.0005)? Even if this database was not available to EpiX, I have the same concerns as noted above regarding the choice of genes to search. I do not feel that the database used for this analysis is the best one currently available, and that the FDA database should be considered. However, I still feel that an approach	Thank you for the suggestion. FSIS is exploring how to utilize the VirulenceDB tool developed by U.S. FDA (https://virulence.preprod.fda.gov/). Furthermore, VFDB and BV-BRC (formerly, PATRIC) will continue to evolve as new information is discovered and validated. Hence, this genomics-driven clustering and analysis could continually require refinement to incorporate and update new information into the model.

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:					
				using the entire pangenome of *Salmonella* is a more robust approach than a selected database.	
3			A	Overall, the use of the PATRIC plus VFDB was convenient but not ideal. I mentioned the efforts already underway by FDA to create an improved *Salmonella* VF database, that should have been used in collaboration with FDA. Additionally, there is no consideration for experimental evidence or aspect of pathogenesis these proteins are involved in. Many are actually fitness factors and not true virulence factors	
4	92	203	B	Overall steps and methodology implementation for annotation for virulent factors are good. However, no proper rationale or reference is provided to support the choice of the reference genome of Typhimurium serotypes. Also, the language to implement the steps should be rephrased to make it clearer.	Open reading frames (ORFs) from *Salmonella* Typhimurium reference strain LT2 was used to help identify and annotate virulence factors (VFs) 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.

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:					
					FSIS has developed a *Bioinformatics Supplemental Materials* (available here) to clarify steps of the process.
5	92	206–207	B	Not very clear, was the parsing step performed after PROKKA annotation or before?	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 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 the 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.
6			C	I am not able to judge that.	No response is required.
7	92	197–201	C	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 is 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:					
8			D	The databases and methods used to determine the virulence factors to form the genetic matrix are appropriate.	No response is required.
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 required.
c. Is the clustering algorithm accurately described, utilized, and appropriate for its intended use?					
1	93	229	A	Line 229, "The unsupervised random forest algorithm is agnostic to the biological meaning of the virulence factor genes and will cluster observations solely based on similarity." This is the precise problem with this approach. There are many unanswered questions about the genomic context of *Salmonella* virulence. Simply clustering based on "known" VFs, then assuming that those clustering together will behave similarly, is flawed. Now, all serovars and isolates clustering together will be assigned the same risk multiplier. We know that there is a very uneven distribution of genomes and isolates by serovar. If even one dominant serovar which should be considered as "high virulence" is included in that cluster, the corresponding epidemiological data (individuals sick / individuals exposed) will be highly skewed towards that serovar. Others will be lumped in that may not be "high virulence" but will be done so because of similarity in limited	The genomic context of *Salmonella* virulence is beyond the scope of the present work. Rather, a subset of virulence factors variably present in a collection of over 36 thousand *Salmonella* assemblies were used as features to distinguish clusters. A risk multiplier was estimated using the relative presence of these clusters in poultry outbreaks and food samples. The risk of illness with a strain from cluster 1 (C1) is 2.1 times higher than the risk before knowing that the strain belonged to C1. Similarly, the risk of illness with a 2 strain was 2.63 times lower than before knowing the strain belonged to C2. The risk multiplier is an average over multiple serotypes and subtypes. However, the average is weighted based on presence in poultry outbreaks and food. Serotypes commonly found in poultry (e.g.,

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:					
				gene profiles. The same vice versa effect would occur in cluster 2 (high risk strains will be lumped into the cluster skewed because of a dominant low virulence serovar).	Typhimurium, Enteritidis, Kentucky) contribute more to the estimate than serotypes uncommon in poultry (e.g., Newport, Dublin, Cerro). Therefore, the estimate is skewed towards strains to which a poultry consumer is likely to be exposed.
2			B	All steps in the clustering algorithm are clearly defined and R code implementation also works fine. Overall concept and implantation is appropriate and useful.	No response is necessary.
3			C	The term “clustering algorithm” only occurs in Table 24. I guess you refer to the machine learning algorithm?	Yes, it refers to the machine learning algorithm; namely, unsupervised random forest for classification.
4			C	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.
5	88	78–87	C	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 30) 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.	Clustering is accomplished using virulence factors (VFs), 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 32**). Most serotypes were assigned to a single cluster, but some (e.g., Infantis) were

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:					
					assigned to two clusters. In the case of Infantis when k=4 clusters, 88% 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 a *Bioinformatics Supplemental Materials* (available here) to provide additional clarity and understanding of the methodology. EpiX Analytics has also updated the caption of **Figure 30** to clarify the process.
6			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	Responses to the reviewer's questions are provided 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:					
				development of virulence matrix. Please see below for the details and specifics:	
7	92	185–186	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 93, lines 213–219 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 93, lines 240-245: This ultimately brought the number of isolates allocated to clusters to 40,038; ~4,000 newly added 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:	It is possible that by not including these "rare" isolate assemblies, influential or important virulence factors could be missed or overlooked. 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, 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 commonly passed through horizontal gene transmission. The analysis performed by EpiX Analytics overcame this by including selected *E. coli, Shigella* and *Yersinia* VFs in addition to *Salmonella* VFs (complete list of 193 VF are described in FSIS' *Bioinformatics Supplemental Materials* (available here)). Future iterations should investigate modifying the lower threshold requirement of 50 isolates per serotype and other potentially informative genomic and VF data.

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:					
				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.	
8	92	176–181	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-, 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.	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 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.

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:					
9			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)] 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).	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 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

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:					
					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.
10	52 99 100–101	1215 – 1219 439–447 Table 17	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." "Serovar assignment for *k*=2, 3, and 4 clusters are provided in Table 17. The serovars composing Cluster 1 remained consistent at the three levels of *k* (Figure 31). 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-	EpiX Analytics has added clarification in the text and now more clearly reflects the information in **Table 34**. 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 (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 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:					
				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 99, lines 439–447) and the Table 17. 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 17). 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 moves to cluster 4. Also, please clearly mention as the cluster goes from 1 to 4, the virulence decreases.	
11	95	291–295	D	"These total consumption rates per product led to a weight of 11% for chicken carcasses, 6% for ground chicken, and the remaining (83%) for chicken parts. For turkey, the weights were 75% for carcasses and 25% for ground for product (weight between programs). We applied a final weight of 5/1 for chicken vs. turkey. All weights were provided by FSIS." Why the consumption rates of chicken are being used for turkey? This does not seem appropriate.	A section on the risk multipliers was added to **Chapter 2** to discuss the different weighting scenarios that were assessed and how this affected the associated risk to each cluster. The overall lack of turkey data, as compared to chicken sampling data, does limit the ability to assess risk management options regarding *Salmonella* illnesses from turkey products. Therefore, this simplifying assumption was adopted to

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:					
				Please provide rationale and more information behind this weightage from chicken to turkey and the correlation between consumption rates of chicken and turkey. Also, please explain how this affected the results.	reduce the complexity induced in the modeling due to the lack of complete data in terms of turkey products. That said, the final product standards developed for comminuted turkey were developed using more extensive data, and they can therefore be considered with greater confidence by risk managers. Turkey consumption weights are 75:25 for carcasses versus ground, which are not the same as for chicken . Since chicken is consumed more frequently than turkey, a final weight of 5:1 chicken vs turkey was applied. The risk multiplier results for a 2-cluster model were robust to modeling options including not weighting different products or using chicken or turkey only (**Table 14**).
12			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			A	Beyond all of the elegant math and modeling employed here (which is commendable), the results often do not make biological sense. Again, I point to specific examples within clusters #1 and #2. Lines 861–865 indicate that based on this analysis, 10,000 cells from cluster #1 strains are predicted to result in 57% infection rate. In contrast, 10e10 cells result in only 40% infection rate in cluster #2 isolates. Based on knowledge of the impact of certain serovars within cluster #2 (specifically contemporary Infantis and Reading, Heidelberg as examples), it is highly improbable that such a dose with those strains would only result in a 40% infection rate. This again points to the flaws of the underlying genomic methods towards establishing these numbers, and all other estimations using the modeling approaches. In so many ways, I am discouraged about this overall approach. I applaud the extensive efforts at modeling and dose-response modeling, but there are huge biological gaps that should have been better addressed during the past 15 years. This is not a direct comment towards this	Although those values might seem unrealistic to the reviewer, they simply reinforce the empirical evidence observed in human surveillance data and how that relates to the presence of *Salmonella* across poultry products. This was discussed previously in the response to comment Q2. A2.1. Furthermore, the reviewer is referring to the probability of illness integrating the variability of all strains from cluster 2. As shown in **Figure 32** there is great uncertainty in the probability of illness across all doses, even at 10^10 cells. **Table 38** indicates an estimate of 0.40 with 95% confidence interval (0.22-0.61) at 10^10 cells. Nonetheless, a sensitivity analysis (**section 5.6**) has been added to consider alternative dose response functions and their impact on the reduction in illnesses. Data limitations will always exist and FSIS continues to explore sampling programs and resources to inform and enhance decision making and fill research or data gaps.

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:					
				study but a more general comment to the federal agencies. If they had foresight to gain better biological data on specific serovars and strains, there would be more accurate assessments in this study on virulence which were derived from real data. There are many USDA-ARS scientists (Drs. Shawn Bearson, Michael Rothrock, Michael Kogut, Allen Byrd, etc.) who could have generated serovar-specific biological data using in vitro assays or animal challenge models, which could have helped to inform the modeling in this project. Also, if a more robust post-harvest sampling program was in place, with an emphasis on more representative and more thorough sampling at the stage of slaughter (from ceca) and at point-of-sale (more retail meat samples), I would not be questioning the generalizability of data used in this study. There is a severe lack of foresight here which inclines me to be more critical of the assumptions used in the present study. In other words, this was completely avoidable. Instead, this left the modelers in this study grasping for straws in the data in many respects. It is clear from the writing style that they are uncomfortable with many of the datasets which were used to inform the models, and that is understandable.	

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:					
2	83	2035	A	Line 2035, it is hard to believe that the risk multiplier for the Infantis cluster is 0.31 when splitting into additional clusters (k=4). Given the prominence of Infantis, I highly doubt this reflects reality. This also illustrates the flaw of the approach, again. The authors also mention Kentucky, which they acknowledge includes a highly virulent Group 2 which is currently not prevalent in the US. However, this approach will still cluster Group 2 Kentucky into cluster #2. So by default it will still be classified as low virulence. Where is the underlying biological proof that this is truth? I understand the concepts behind using epidemiological and surveillance data to infer risk of infection, but this overall lumped approach is flawed. See my detailed comments on this under Q2.	The risk multiplier calculation considers the proportion of outbreaks and proportion in poultry products for each cluster. In particular, the risk multiplier is based on proportions of poultry-attributed outbreaks in cluster 1 versus cluster 2 relative to their occurrence in poultry with consideration of a time-series component, underreporting and strength/severity, to name a few. There are limitations with this approach due to using outbreaks to classify higher versus lower virulence in broad clusters where risk is therefore inherently driven by the more virulent strains within each cluster. In the k=4 cluster scenario, the Infantis cluster (Cluster 3) is estimated to have a risk multiplier of 0.31 (95% CI 0.0095-0.89). The comparison between FoodNet and Infantis in poultry was discussed previously in the response to comment Q2. A2.1. Moreover, since the risk multiplier calculation considers poultry-attributed outbreak data from NORS specifically, we have also added some description to **Chapter 2** on the extent of Infantis outbreaks included. Currently, the recent increase in proportion of Infantis in poultry is not proportional to changes in the FoodNet case rate or NORS poultry-

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:					
					attributed outbreak related to Infantis, as would be expected if consumption of poultry were a major contributor to illnesses. The reviewer also predicts that the applied clustering method would classify new Kentucky isolates (presumably the European Kentucky serovars with high antimicrobial resistance levels, as opposed to high probability of infection/virulence) into cluster 2. Clearly, such a prediction can only be made with actual isolates and would be solely based on genomic data. We note that this is not a one-size-fits-all approach and it will need to be reevaluated due to the evolving epidemiological dynamics of *Salmonella* to improve granularity in the future. Furthermore, as noted in the *Choice of Two Seroclusters* section, the more virulent Kentucky Group 2 (Soltys, 2021) has been recently isolated from chicken samples in the United States (Cameron P Thompson, 2018), so these findings should be revisited periodically to determine if *Salmonella* Kentucky maintains its low-virulence status.

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:					
3			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.
4			B	Proper methods to account for all details of the modeling like the Bayesian approach, bootstrapping, and sensitivity and specificity analysis have been used in the modeling process and are appropriate.	No response required.
5	98	407–411	C	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.
6	106	546–561	C	It is not clear to me whether you use Teunis et al. data (caption fig 33 and line 546–561) or the outbreak data from cluster 1 (line 546) 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	The dose response model was derived from Teunis (2022) data on cluster 1 strains: Enteritidis and Typhimurium. Teunis (2022) provided new and corrected data from Teunis (2010).

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:					
				based on one data set to another data set? Please explain the approach better. 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".	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., asterisks vs blue circles).
7	108	575	C	I like the approach to use this fitted equation. Very pragmatic!	No response required.
8			C	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 can decide whether (s)he thinks it is worthwhile to focus on it.	The dose-response model 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 function, such as the median estimate in the uncertainty dimension.
9			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 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.	
10			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 function below the first inflection point. A 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,

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:					
					relative to *Salmonella*, and the dose-response would be right shifted. 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			A	I have read these papers and they are widely cited and accepted. Therefore, I believe that use of this model is indeed appropriate. The approaches make sense and the underlying data for the approach in the Teunis papers is extensive and well described.	No response required.
2			A	If I am reading the methods correctly, the focus of the DR model was on cluster 1. It is unclear to me (could use some more description) if this was based purely on the data for Enteritidis/Typhimurium or additional serovar data for all other serovars of cluster 1. It also appears that serovars in cluster 2 were not considered. However, there was biological data presented on other serovars in Teunis 2022 (using previous human challenge	The dose-response model was derived from Teunis data on cluster 1 strains: Enteritidis and Typhimurium. The dose-response model for cluster 2 was estimated using the risk multiplier relationship. In Teunis (2022), there are very few data points for the derivation of individual serotype-specific dose-response models outside of Enteritidis and Typhimurium. As stated by the author, "The other serotypes

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:					
				experiments and outbreaks) which actually contradicts the cluster assignments in this report. In that study, several serovars identified as cluster 2 in this study (Heidelberg, Infantis, etc.) actually have lower infectious and illness doses at 1% than Typhimurium. This again suggests that the cluster assignment is not supported in real life. I'm curious if this was considered, and if not, why it was not?	found in outbreaks (apart from the above Enteritidis and Typhimurium) were rare, leaving too little information to obtain precise estimates" (Teunis, 2022). This is evident from Teunis (2022) Figure C3d showing a 95% range that goes from almost 0% to 100% probability of illness for most dose ranges.
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	111		B	All assumptions made during the modeling and use of parameters are very logically enumerated in Table 24.	No response required.
6	97–99	366–426	C	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.
7			C	Still, some consideration should be given to the fact that the DR model is based on	This is a valid point. If outbreak strains are more infective, one would expect 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:					
				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.	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 lower virulence cluster 2 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.
8			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.	Thank you for sharing another dose-response approach using genomic data. The approach implemented by EpiX Analytics aims to derive reasonable dose-response models comprising a wider range of serotypes in poultry; however,

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:					
				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.	improving the resolution is an important step for future risk analyses.
9			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 10^6 cfu/g, for example) is strictly not required, it is ultimately not applicable	No response required.
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.					

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:					
1			A	The risk multiplier is applied at the "cluster" level, when clearly it should be considered at least at the serovar level, and possibly strain level. The resolution at the cluster level is not high enough to provide a blanket multiplier across all serovars and strains within those clusters. Therefore, the over-riding concept of "high" and "low" virulence and the associated multipliers given to these clusters will in many cases be invalid. All subsequent dose-response models rely on this critical information and are thus misleading. See my comments in Q2 about methods to obtain a more accurate risk multiplier. These are not criticizing the use of the multiplier in the modeling scheme, only the genomic methods used to obtain said risk multiplier.	We understand the need for improvements in the granularity, and that will continue to be an important goal for FSIS. As noted previously, a sensitivity analysis (**section 5.6**) has been added the document to consider alternative dose response functions and their impact on the reduction in illnesses.
2			A	The risk multiplier is based directly on proportions of isolates in outbreaks in cluster 1 versus cluster 2 relative to their occurrence in poultry. If you actually break this down by more contemporary versus more historical outbreaks, I am confident that the multiplier would change dramatically and probably even flip. I would recommend first determining if the risk multiplier changes if you were to break this apart over time. For example, if you consider 3–5 year intervals distinct from	FSIS developed *Bioinformatics Supplemental Materials* (available here) as well as added several sections to **Chapter 2** to provide additional transparency and detail on the risk multiplier derivation, underlying outbreak data, and sensitivity analysis discussion. EpiX Analytics applied recency weights that account for the differences across time and place more weight on recent outbreaks compared to historical data (i.e.

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:					
				one another, does the risk multiplier change by cluster? If it does, this points to a problem with the temporal adjustment factors. I don't believe that the adjustments for recent versus historical prevalence adequately deals with this problem. I also fail to understand why this multiplier is generalized to poultry and not specific to turkey. It wasn't clearly justified why this was not a concern. Again, running these analyses separately (chicken vs. turkey) would confirm or deny if they should be lumped together. In both of these cases, if the breakdowns of time and source yield essentially the same risk multipliers, then the concerns presented here are attenuated.	exponential decay function). This method is also used in IFSAC's outbreak attribution and provides the best estimate available today. EpiX Analytics considered 3 different temporal scenarios (no recency weighting, and recency weights starting at 1 year and 5 years (baseline)) in **Table 4**. The multipliers did not change dramatically and only when not considering any recency weights did the risk multiplier for cluster 2 increase (as more historical outbreaks were relatively split between clusters (**Chapter 2**). The second question posed by this reviewer as to why the multipliers were generalized across turkey and chicken, and not product specific, is presented in EpiX Analytics' report in Appendix A, specifically **table 35** and the accompanying text and discussed in **Chapter 2**. In particular, the risk multiplier calculation using only turkey is constrained due to the limitations on the available data. Out of the 216 unique serotype-outbreak combinations, only 44 were definitively attributed to turkey with a majority occurring prior to 2017 (i.e., historical). At the same time, within-product weights to estimate the proportion in poultry

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:					
					considers 75% turkey carcasses to 25% comminuted turkey. As discussed throughout the *Turkey Risk Assessment*, turkey carcass data is limited (roughly <10 detections annually), and thus, serotype proportions fluctuate dramatically. Hence, additional data is required to appropriately refine the weights to consider turkey alone under this approach.
3	112	Table 24	A	Assumption #6 in Table 24 is problematic. Again, lumping these serovars into two broad clusters ignore the established fact that different serovars grow better under different conditions. Ideally, a re-analysis of the genomic data should be performed using different methodologies (see my suggestions in Q2 above) to discern more accurate genomic clusters. I do not believe that simply increasing the number of clusters (k) using the same analysis will solve the problem (see my comment on Kentucky and Infantis above). In particular, some serovars survive better than others on raw meat. Secondly, some serovars grow better in enrichment procedures than others. This undoubtedly has an impact of the MPNs by serovar on the final product, and the ability to accurately detect such strains/serovars and their proportions on product. (See PMID 19435216/ DOI:	The genomic clusters were developed using aggregated data to derive overarching dose-response relationships by cluster. Improving the resolution of the clusters may allow future consideration of a more refined attenuation component (inactivation/growth) to improve upon this simplifying assumption for clusters of serovars. As noted previously, a sensitivity analysis has been added to **Chapter 5 Final Product Standard**.

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:					
				10.4315/0362-028x-72.4.707 and 22217013/DOI: 10.1089/fpd.2011.1016, for example). There may not be an easy solution for this but it needs to be addressed. Why not focus on the distributions of serovars on raw retail product, as opposed to the distribution of serovars at the earlier stages of harvest? This would eliminate the need for prediction of outcome based on hang and pre/post-chill samples.	
4			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.	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 Federal Register at 8460). For this reason, a portion of the work that was conducted under the Cooperative Agreement (with an external private sector collaborator) was not made available to the peer reviewers, per the OMB
5	94	256, 260–262, 264	B	The methodology used is conceptually correct, however, no data or code was provided to review the attribution steps. Steps take to implement the random forest clustering, distance matrix, and multiplier are appropriate.	
6	95–96	302–322	B	No code for attribution steps provided: No code for poultry steps	
7	96	330–352	B	No code for attribution steps in clusters	
8	96–97	353–364	B	“Comparison with FoodNet Data”, no code/data provided to review	

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:					
9	97, 98	377–385, 416–421	B	The conceptual framework and assumptions made for the implementation of the Bayesian framework with modified Teunis model in these lines are appropriate. However, the R implementation was missing for these steps and the direct output of Bayesian posteriors and uncertain/variable derivation of posteriors was supplied to plug in the DR model that runs correct.	exemption. Nonetheless, all the methods were fully documented in **Appendix A** of this FSIS risk assessment report.
10			C	I would not know any other data sources or methods that should have been used as alternatives.	No response required.
11			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.	Response is in previous comments.
12			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. Note: the “rare event” nature of the contamination is not obviously more applicable to the case for turkey products. While comminuted turkey products have a	No response required. Comments addressed in Q1.

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 prevalence (and a lower implied prevalence in 325g samples from the lognormal distribution fitted by maximum likelihood) and the median is lower (-4.8 log cfu/g), the fitted lognormal distribution has a much higher standard deviation, such that the arithmetic average concentration per gram for comminuted turkey is higher than for comminuted chicken. The arithmetic average concentrations, in addition to the geometric averages provided, should be made apparent to the reader for initial concentrations and for distributions of doses.	
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	The basic premise of the underlying model math appears to be widely accepted approaches which are sufficient. My problem is with the inputs, specifically the derivation of RR1 and RR2 for subsequent use. The underlying genomic data contributing to the ultimate probability of illness given mixed populations is problematic, a point I have already exhausted.	No response required.
2			A	I also have concerns about the underlying data used to estimate proportions of cluster	Thank you for the suggestion. Additional details about the datasets have been

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:					
				1 versus cluster 2 entering the food chain. Again, the math is fine but the inputs are based on very limited data. It would be helpful throughout the report to note the number of samples included in each analysis (i.e. Figs 8–9). Why were retail meat samplings from FDA not a part of this project? It seems from the FDA NARMS site that there are lots of samples and isolates to work from, and these are closer to what actually enters the consumer's home than any other sample type. It seems more logical to use that as baseline as opposed to post-chill plus several additional assumptions.	incorporated into the document to improve transparency on the limited datasets. FDA NARMS isolates were included in the clustering and dose-response models derivation.
3			B	The conceptual and R code implementation of the two-curve dose-response model is appropriate and reproducible.	No response required.
4			B	Nice work.	No response required.
5	136	367	C	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-	We have enhanced the **dose-response model section (5.3)** in the document to improve the clarity on how the dose-response is used.

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:					
				curve" does not occur in the document either.)	
6	46–47		C	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 between the lines. It should be clarified and discussed.	A section (**4.4**) describing the "risk per serving" as described by the dose response models was added to the document making it clear that the derivation of the dose-response models was for the purposes of evaluation of control measures. The baseline probability of illness estimates serve as the empirical, rather derived hazard characterization, and the section (**4.3**) title has been modified to clarify that point.
7	18–20		C	The dose-response model is used to derive the final product standards for comminuted turkey. Serotype-based final product standards were not feasible. If I understand it well, the two-curve DR relation are used to compare scenarios, but not to implement different interventions for different serotypes. I believe the model is applied appropriately, but I am not convinced that the result would have been very difference if a simpler one curve DR relation would have been used. The added value of the	A sensitivity analysis (**section 5.6**) has been added to the document outlining the effect of alternate dose-response models on the illnesses avoided estimates.

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:					
				two-curved DR relation has not been clarified.	
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.	The specific comments are addressed below.
9	92	201	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	94 97	Footnote 2 373–374	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 will 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	94 96	263–279 323–328	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 serovars 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 35** in **Appendix A**. Text describing the baseline scenario has also been added immediately prior to the multiplier tables.
12	95	300	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 document 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

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:					
					were summed to calculate the proportions by cluster and determine the 95% confidence intervals described in **Tables 18-19.**
13	98	393	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 97, lines 367–368. 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 models 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.

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:					
14	99	425–426	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 interpret. But the explanation and rationale provided here for this choice is unclear and incomplete. Please provide the explanation.	A polynomial was fit to the DR functions with 95% credible intervals. This generalization assists with portability and efficiency of implementation without loss of detail/information.
15	100–101	Table 17	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	102	463–469	D	The authors mentioned that two serotypes (Berta and Saintpaul) switched serotypes during bootstrapping. Table 17, pages 100–101 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	112	Table 24	D	Assumption 7: *Salmonella* inactivation and growth are not product-specific. This is not	The clustering and dose-response models were developed on the aggregated

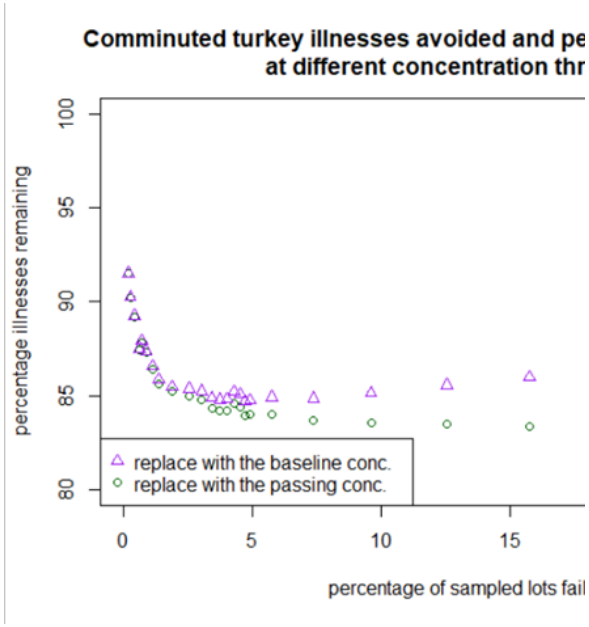
Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q3. Please evaluate the two-curve dose-response model used to estimate the probability of illness for a given exposure dose of *Salmonella*, giving specific consideration to the following:					
				appropriate, as *Salmonella* inactivation 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-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.	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 potentially result in contradictory dose-response models for each cluster by product.
18	112	Table 24	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.	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**.

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:					
				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).	
19			E	While it is not possible to conclude that the dose-response models are not "appropriately used," the level of complexity 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.	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
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 comminuted turkey 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			A	This is probably the most solid and most actionable part of the report. At the level of *Salmonella* as a whole, I think this report very nicely describes the existing data, its analyses, and the potential impacts of increased mitigation procedures at the plant level. I honestly don't know what else they could have done using the limited data available. If this results in recommendations to improve total bacteria or Enterobacteria counts through the processing chain, that is something that could have impact.	Thank you. No response required.
2			B	Risk management scenarios proposed to evaluate the public health impact are reasonable. However, the feasibility to model for implementing an efficient risk management approach is limited by the scarcity of data availability. The modeling feasibility for different turkey product types was different due to less surveillance frequency of some product types like parts and also needed different assumptions as per the data availability.	Thank you. 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 comminuted turkey 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			B	However, the exposure assessment exercise performed with all available data for different turkey products was useful to give an overview of hazard (*Salmonella*) prevalence and its public health impact. This was helpful to provide a basis for making choices in dose-response modelling and risk characterization.	Thank you. No response required.
4	21 50	533 1138	C	The approach is selected because FSIS believes in it. Please explain the origin of that believe, describe clearly why it is chosen and provide appropriate references.	Further explanation of the selection of this approach has been added to the start of the **Chapter 5 Final Product Standards**.
5			C	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. Their Figure 4 is giving the relation between the fraction illnesses avoided and the fraction non-compliant lots for several thresholds. and is comparable with Figure 17.	Thank you for your suggestion. We have incorporated additional tables for risk managers regarding the number of lots failing/diverted for different approaches. The Nauta paper is referenced in the risk assessment. 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 comminuted turkey 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:					
				I used the R model provided to make a new graph similar to Nauta et al 2012 Figure 4, 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 17% of illnesses is prevented here, whereas there it 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.	

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 comminuted turkey 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	The scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on comminuted turkey products are reasonable. However, there are some concerns regarding the use of attenuation multiplier (please see below Q4 (e)).	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 comminuted turkey 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			E	(no comment provided by 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 comminuted turkey products)?					
1			A	The scenario analysis technique is very well described. In fact the entire document is thorough with this regard. I do believe it is appropriate for the estimations of possible changes in *Salmonella* levels. Given the variation in strains and serotypes, I question if the data is robust enough to provide accurate inputs on distribution of serotypes. This becomes especially problematic when further lumping serotypes into two clusters. See my comments in Q2 regarding clustering, and in Q3 regarding estimation of inputs.	A sensitivity analysis has been added to the document (**section 5.6**) to assess the distribution of serotypes on the reduction in illnesses.
2			B	Scenario analysis for risk characterization suggested the implementable suggestion for concentration threshold for replacing with average or passing lot of products, with an estimate of illness reduction of approximately 16%, which is efficient.	No response required. Thank you.
3			C	Yes, the scenario analysis technique is well described and appropriately applied	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 comminuted turkey 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	55	1292	C	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**. The standard definition was used.
5	55	1291–1293	C	This is just the exposure distribution. Why write it differently here? That is confusing. Please explain that, when calculating the dose at consumption, you actually assess the exposure.	The sentence was rephrased to make explicit reference to the exposure assessment.
6			D	The scenario analysis was appropriately described and utilized within the current scope of work.	No response required.
7	20 49	530–532 1138–1141	D	"A major assumption of this modeling approach is that consumer demand for raw turkey products will continually be met by the industry, and so every lot removed (as a result of a new standard) will ultimately be replaced by another average lot." While this is a reasonable and valid assumption, please provide evidence to support this.	An explanation of this assumption has been added to document at beginning of **Chapter 5 Final Product Standards**.
8	21	552–557	D	"Given the lack of robust data, it was not possible to estimate the public health impact of performance standards at receiving that address either *Salmonella* levels or serotype. That said, attempts were made to develop a hypothetical model (Appendix B) of serocluster distributions	Two seroclusters were derived from aggregated isolate assemblies which included both chicken and turkey isolates. The attempt presented here using a hypothetical model was developed on limited turkey carcass sampling data at

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 comminuted turkey 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:					
				that would allow for the development of performance standards at receiving. This hypothetical model does demonstrate that focusing on reducing the Cluster 1 distribution at receiving could have an improved effect on reducing illnesses." This reviewer understands that there are limited data for turkey. However, FSIS must explain how without relevant data for turkey, how this hypothetical model will represent serocluster distributions reasonably relevant to turkey. Seroclustering was done using predominantly chicken data and must be revised in the future with the availability of turkey data.	rehang from a historical FSIS microbiological baseline study. The concept illustrates an approach that can be evaluated in the future if robust turkey data becomes available and a 2 serocluster structure was defined.
9	21	585–588	D	"Scenarios were run assuming underperforming establishments adjust their practices toward meeting a level of control according to the indicator organism metrics listed above from rehang to post-chill. That is, by setting a log reduction or presence fraction target/guideline, the overall prevalence that results from that change can be assessed." Although this is a reasonable assumption, please provide more information on the log reduction and presence fraction target and guideline.	We have updated the text to improve the clarity regarding the specific process control scenarios.

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 comminuted turkey 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 are the specific values used for these scenarios.	
				On Page 21, lines 572–576 it is mentioned about the weak relationships between indicator organisms (i.e., APC and EB) and *Salmonella* prevalence and hence, it was not possible to assess the risk management question regarding the public health impact of monitoring/enforcing process control from rehang to post-chill in the same manner as it was estimated for final product standards. Scenario analysis is not clear for process control regarding from rehang to post-chill. Specifically, two process control standards that were investigated using the available data from 2008–2009 must be further explained.	We have updated the text to describe the process control scenarios, as well as how that model differs from the final product standards approach.
10	22	610–612	D	"Nevertheless, indicator organisms were readily measured and quantified compared to *Salmonella* levels at both sampling locations. Future analyses would require more current information to validate appropriate targets for APC and EB." Do both locations imply rehang and post-chill? It is not clear regarding the correlation between indicators and *Salmonella* in these locations. This reviewer suggests performing a correlation analysis when	Yes, "both sampling locations" identify the two poultry carcass sampling point locations: rehang and post-chill.

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 comminuted turkey 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:					
				data are available in the future so that appropriate targets for APC and EB are validated.	
11	26	739–741	D	"Scenarios (that is, options) for receiving guidelines, process control monitoring, and final product standards serve as inputs to the overall model, and the reduction in per serving risk, as well as an estimate for annual illnesses avoided are the outputs for each scenario, whenever appropriate." Please elaborate "whenever appropriate" or provide examples for which this was done.	We have added the line to clarify what is meant by "whenever appropriate."
12	39	995–1001	D	"The distribution of serotypes fluctuates widely as there were few positives in turkey carcass across time (Figure 8), whereas comminuted turkey has relatively similar serotype proportions annually (Figure 9). Other top *Salmonella* serotypes from the CDC FoodNet summary include serovars I 4,[5],12:i:-, Infantis, and Typhimurium, which are on an upward trend in the proportion of detections in turkey over the last few years. Although serovar Enteritidis is most frequently associated with human salmonellosis, it is rarely observed (or detected) on turkey carcasses or in comminuted turkey products." The cluster	Additional description/discussion has been included into the *Serotype* section under **Chapter 3 Microbial Profile** about the Risk Multiplier. The temporal dynamic (i.e., serotype fluctuation) is incorporated in the risk multiplier estimation. Further details on the risk calculation are provided in FSIS' *Bioinformatics Supplemental Materials* (available here) to improve transparency of the process.

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 comminuted turkey 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:					
				analysis does not appear to take this serotype fluctuation into consideration. Please provide more information or clarification.	
13			E	The overall approach seems appropriate to its intended use in answering the risk management questions.	No response required.
14			E	The inability to answer other risk management questions is also adequately described.	No response required.
b. Are the data analyses and R source code accurate for the aims of the study?					
1			A	The R source code looks complete and is adequate. I have no concerns about the availability of source data.	No response required.
2			A	A glance at the R code indicates that it is accurate and fits the methods described in the publication.	No response required.
3			B	The proposed model implemented correctly and coded accurately, runs perfectly.	No response required.
4			C	I could not identify anything inappropriate here, the R program is well structured and seems to work well.	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 comminuted turkey 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			C	From the R model I see that the exposure assessment does use the lognormal distribution of concentrations (μ =-4.857 and σ =2.333) in comminuted turkey and combines it with the attenuation distribution. The dose-response relation; however, is obtained using another distribution. This confuses me and should be explained better. I assume the reason is that the number of cases that is specifically attributable to turkey is not known, so a simplifying assumption has to be made, but this would have to be explicitly mentioned.	The assumption has been more clearly stated.
				To derive the DR model "The distribution (Log10Normal(-3.037117, 1.279985)) was used that reflected the initial contamination of *Salmonella* in a mixture of the raw poultry products – chicken carcasses, chicken parts and comminuted chicken – according to their relative frequencies of consumption (see Chicken Risk Assessment)". Where does the turkey come in? What implications does this have for the risk estimates (in absolute numbers for the same DR model)? Please clarify and discuss.	A sensitivity analysis (**section 5.6**) has been conducted to address the role this dose-response relationship—based primarily on the more robust chicken data—has on the analysis and 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 comminuted turkey 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	The R code is clear, and relatively well documented. 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	""“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

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 comminuted turkey 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:					
				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 appropriate, given the conclusions rely on this estimate.	might fail and some lots just above the LOD might pass). Furthermore, as explained in *Chicken Risk Assessment* the "*Accuracy of quantitative PCR methods*" **Appendix B** section, the current method for accurately quantifying samples is a much more relevant concern than the Poisson variability associated with sampling. As pointed out in our 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			A	If this question is related to page 28, then I do think these methods are appropriate and described well.	No response required.
2			B	Not a very clear definition of production lots has been described in the document. Please make sure it is clearly described. However, the sampling frequency of FSIS is clearly mentioned and methods of sampling and testing are properly described.	Production lot sizes are described in **Chapter 1, section 1.1, Purpose and Scope**. The definition has also been added to the Assumption Table for further clarity.

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 comminuted turkey 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			C	I am not sure I understand this subquestion (c), and struggle with what it refers to.	No response required.
4			C	For comminuted turkey meat, the dataset holds 156 (page 134, line 297 or 157 (page 17, line 457 and page 37, line 961) positive results among 1178 samples. The distribution fitted through these data (lognormal(-4.857, 2.333) must be very uncertain. Technically, the approach seems OK (although it is not explained sufficiently well), but this uncertainty is not addressed. The uncertainty should be expressed and discussed.	The parameters of the distribution are not subject to a substantial degree of uncertainty. The variance-covariance matrix generated during the maximization of the likelihood function is given by 0.0547 -0.0343 -0.0343 0.0271, so the ellipsoid of concentration describing the joint variability in the mu and sigma parameters is relatively compact. The effect of uncertainty in the parameters of the contamination distribution is further reduced, relative to other components of the risk assessment, by noting that illness reductions are calculated using I_avoid=I_turkey (1-P_new (ill)/P_baseline (ill)) because the effects of uncertainty in the dose distribution appear in both the numerator and denominator of ratio. Thus, the effects of uncertainty in the concentration and attenuation distributions and the dose-response models mostly cancel out. This leaves the highly

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 comminuted turkey 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:					
					uncertain estimate of the number of illnesses from turkey as the dominant source of uncertainty.
5	37–38	975–979	C	The approach uses a threshold value for a continuous distribution of 1/325, standing for 1 cfu in a 325 g sample. Given the discrete nature of bacteria, this is not fully correct. A more suitable distribution is actually poisson-lognormal, not lognormal (Gonzales-Barron et al. 2010. *Int. J. Food Microbiol*. 136 (3), 268–277). A simple Monte Carlo simulation with a poisson-lognormal with intensity randomly sampled from a lognormal(-4.857, 2.333) distribution gives 19% probability of 1 or more cfu, i.e. 3% more than the 16% mentioned.	**Section 3.2 *Salmonella* Concentrations** has been updated to include this suggestion.
6	38	Figure 7	C	I would stress that this distribution is hypothetical. Although not well explained, I assume that it is assumed that there is a lognormal concentration distribution, and this is fitted through the 157 data points from the 1178. (page 134, line 303 in Appendix C says it is done with a weighted maximum likelihood routine, but there is no reference to that, and it still does not tell me what has been done.) The distribution that comes out contains a majority of	Clarification was added to the text as and these references are a useful additional resource. Williams, M. S., Ebel, E. D., & Cao, Y. (2013). Fitting distributions to microbial contamination data collected with an unequal probability sampling design. *J Appl Microbial*, *114*, 152-160. https://doi.org/10.1111/jam.12019 Williams, M. S., & Ebel, E. D. (2012). Methods for fitting a parametric

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 comminuted turkey 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:					
				concentrations that implies that the turkey is not contaminated, so the distribution is fitted on the basis of the tail only (only containing values larger than the fitted mean). Have zero-inflated models been fitted as alternative? Is their performance compared with the performance of the risk assessment based on this hypothetical lognormal distribution? It will anyway be useful to explain the peculiarities of the approach better.	probability distribution to most probable number data. *Int J Food Microbiol*, *157*(2), 251-258. https://doi.org/10.1016/j.ijfoodmicro.2012.05.014
7			D	The methods used to describe the contamination of the lot from samples are appropriate.	No response required.
8	50–51	1184–1193	D	"*Salmonella* serotype detection is generally limited to the most abundant serovars due to the current sampling techniques employed. It is assumed that if a particular serotype is in high abundance within a flock/lot, then the results of *Salmonella* positive samples at two points in the slaughter process (i.e., rehang and post-chill) should regularly agree. If a variety of serovars is present within a flock/lot, as observed in chicken carcasses (C. P. Thompson et al., 2018), then the rehang sample can become a poor predictor of the	Additional references and emphasis on the limited turkey data has been included.

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 comminuted turkey 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:					
				serotype identified at post-chill. Subsequently, the assumed distribution mixture of serotypes in that flock/lot. In the young turkey baseline study, in particular, there was a single rehang sample (1 out of 144 positives) that was identified to contain multiple serotypes. Surprisingly, the paired post-chill sample (i.e., from the same flock) was negative for *Salmonella*." Although the assumption is reasonable, the authors must emphasize that the turkey data is limited.	
9			E	The treatment of the sampling data appears to be appropriate. In addition, the determination that is not possible to quantify contamination levels for turkey carcasses and turkey parts at this time is appropriate. One possible exception is the question of whether the Poisson nature of sampling data (e.g., that sampling data is the result of a Poisson-Lognormal process) is adequately captured in the determination of the likelihood of a positive test result. In some cases, the PLN nature of the process seems to be explicit, while in other cases, it is not clear.	See the above response to Q4b comment E8.

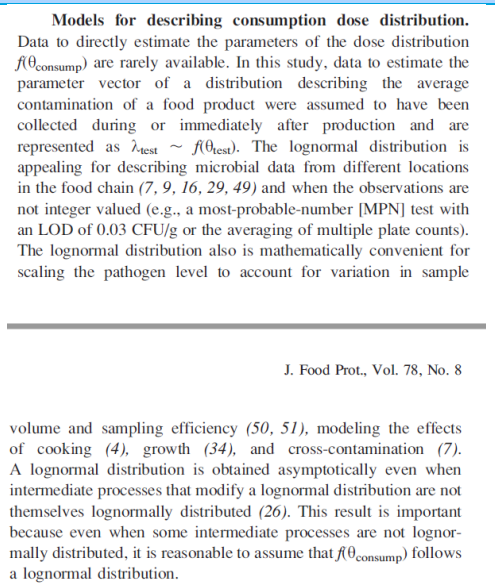
Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q4. Please identify limitations, weaknesses, or inadequacies of the scenario analyses conducted to evaluate the public health impact of changes in *Salmonella* levels and/or presence of certain serotypes on comminuted turkey products. Please 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:					
d. Is the assumption that multiple serotypes are present within lots appropriate and how else can the mixture of serotypes (i.e., "serotype scheme") be described?					
1			A	I worry that this assumption is problematic. There is so little evidence of the nature/context of multiple serotypes within a lot. I do not doubt that this is truth, but in some cases this was described as only a few examples showing serovar-level similarities or differences within lot samples. Other researchers have explored this in much more detail (i.e. Nikki Shariat at UGA). This body of work may actually be more useful than the FSIS data for informing models. Please see DOI: 10.1016/j.fm.2022.104149, DOI: 10.1128/aem.00204-22, DOI: 10.4315/0362-028X.JFP-19-166, and DOI: 10.1128/AEM.01859-18 for data that might be useful towards establishing a better assumption in the model.	Thank you for the provided additional references. We have expanded the discussion on the assumption that multiple serotypes are present within a lot by adding a table of assumptions to **Chapter 1**.
2			B	The serotype scheme is based on the WGS virulent factor-based clustering, so has a lot of support from the genomic data and is appropriate.	No response required.
3			B	Also properly referenced and common distributions have been used to model the	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 comminuted turkey 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:					
				serotype mixture and support the assumptions	
4			C	The evidence seems to be very poor. There is very little data to support either the hypothesis that usually only one serotype (or cluster) is present and that more serotypes (or clusters) can be present. I find it hard to imagine how, with so little data, you can derive and implement a final product standard based on serotype (if you find one cluster, the other may very well be present). It makes the analyses done a bit of an academic exercise, in practice it is highly uncertain whether the assumption is correct, and the implications of that uncertainty are not assessed or analyzed.	We have expanded the discussion on the assumption that multiple serotypes are present within a lot. We have also incorporated a Sensitivity Analysis (**section 5.6**) on the serocluster distribution with the final product standards model.
5	52	1210–1212	C	The resulting average proportions of cluster 1 and 2 in any lot are 0.3 and 0.7. It is not clear how you obtained these numbers, or how certain you are about them. They seem to be of crucial importance though and should be carefully explained.	As noted above, we have added a sensitivity analysis regarding this parameterization.
6			D	The assumption that multiple serotypes are present within flocks is appropriate but needs further references to substantiate this, please see below.	Additional references have been 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 comminuted turkey 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	64	1502–1508	D	"Data collected during this period identified incoming turkey flocks containing some amount of contamination, although less than 10% of rehang samples were test-positive for in a sample, and ultimately, a flock. A recent poultry study observed that contaminated chicken carcasses typically contain multiple *Salmonella* serotypes (C. P. Thompson et al., 2018). This individual sample hypothesizes a potentially similar situation, although more data is certainly required to validate." In addition to the reference cited, are there any other research reported or citations available to corroborate this. Given the limited data for turkey, it would be better to test for multiple serotypes in the future to validate and revise the approach, if possible.	Thank you for the suggestion. Efforts to improve the resolution should be a high priority for future risk assessments and has been added as a research gap to the **Chapter 8 Discussion** (**section 8.1**).
8			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	Additional references have been added to support this assumption.

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 comminuted turkey 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:					
				dominant), so multiple serotypes would seem to be the safer assumption.	
e. Were any considerations missing from the development of the attenuation multiplier to adequately describe *Salmonella* growth and die-off after raw turkey product leaves processing?					
1			A	I understand the concept behind the attenuation factor, but honestly I don't fully understand how it was developed and used in this study. This could be my limitations in understanding the underlying math. It would be helpful to explain this and other key concepts described in readable language. I think of this not only for people like myself but for the rest of the scientific community that has not been trained extensively in this area.	Additional text describing the attenuation factor has been added to improve clarity.
2	52	1221–1222	B	The assumption for dose for use in the DR model considered the mixture of all raw poultry products and had been modeled with lognormal distribution is adequate The variability (which can be explained by the embedded Poisson process) and uncertainty (with gamma or beta distribution) around doses can be added here before using the doses in the DR model. This will complicate the Dose-response model; however, it can be helpful	We have added explanation and support for this assumption to the report. The chicken attenuation multiplier was adopted due to the severe lack of complete data on turkey products. Further, given the available data, only comminuted turkey products could be adequately modeled. Comminuted turkey and comminuted chicken are typically handled, prepared, and consumed in a similar manner.

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 comminuted turkey 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:					
				to address the described limited availability of the data.	
3	52–53	1225 – 1227	B	The attenuation multiplier was used using the literature mentioned (E. Ebel & Williams, 2015), lognormal distribution parameters with overall consideration of all effects of partitioning, mixing, growth, and attenuation, essentially modeling all processed in farm-to-fork steps using one distribution, is a shortcut way of modeling all process in lumpsum manner. Was due to limited data for these intermediate steps or based on some approximation, this average multiplier was used? Added rationale and explanation around this assumption will be helpful to rationalize its use in the model.	
4			C	This attenuation distribution is obtained from a referenced paper, without any reference to assumptions and shortcomings of the method. Yet, it is an accepted peer reviewed paper, so we have to assume it is suitable.	We have added more description to section 1.5 of the attenuation distribution. Our referenced paper provides support for use of the lognormal distribution (Ebel Williams, 2015):
5			C	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	

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 comminuted turkey 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 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.1539-6924.2010.01481.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. 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 doing some sensitivity analyses with a different model (such as the model presented by Nauta et al 2012)	 **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. We have now also conducted sensitivity analysis on the attenuation distribution and included this variable in our examination of uncertainty. Thank you for the suggestion. It is definitely possible that the attenuation distribution is somewhat different for turkey than chicken.

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 comminuted turkey 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:					
				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 what we actually observe in epidemiological data. (see https://doi.org/10.1111/risa.12153 and https://doi.org/10.1111/risa.12538). In the approach used here all the "error" is 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	However, without complete data on consumer handling of turkey products, it is difficult to develop a more refined model. In addition, we could only model comminuted turkey. The preparation and handling of comminuted poultry should be relatively similar.

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 comminuted turkey 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:					
				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, 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 chicken and turkey 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 relative risks from differences between products/models are usually not that large.	
6			D	There are many steps from raw turkey 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	Explanation of the full scope of steps that the attenuation distribution describes has been added to **section 1.5**.

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 comminuted turkey 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:					
				can cause increase in pathogen load. This has been missing in the modeling approach.	
7	28 52 138	809–812 1224–1228 407–409	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	As it relates to absolute risk, we agree the attenuation distribution is important and we have included new language to address the assumptions used in the turkey risk assessment. Nevertheless, the ultimate output of the finished product standards assessment is the proportional reduction in illnesses. Therefore, the change in risk (before and after risk management option implementation) 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 comminuted turkey as for chicken products is based on 1) a common target internal cooking temperature recommendation for all poultry products (the log10 reduction average should be similar for any product), 2) the serving size being similar across products, and 3) alternatives not being readily available or refined due to the lack of complete data for turkey products.

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 comminuted turkey 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:					
				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 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	
8			E	See above, in response to Q1a. The assumption that an attenuation multiplier applicable to chicken (all products) is specifically relevant to all turkey all products. This assumption should be thoroughly discussed. It may be that there is simply no other choice available to FSIS with the current state of knowledge, but the failure to confront this assumption (e.g., that the exposure assessments for chicken and turkey are identical, other than for initial contamination levels) may be more	We have incorporated this into the Assumptions Table (**section 1.3**) and expanded upon the sensitivity analysis (**section 5.6**).

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 comminuted turkey 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:					
				problematic than the fact of making the assumption. ideally if it can be shown that the key conclusions in response to the risk management questions are relatively robust to the exact nature of the attenuation distribution (as may very well be the case in this risk assessment, e.g., under the rare event assumptions elsewhere described).	
f. Does the Monte Carlo simulation approach adequately model the scenarios?					
1			A	I am not an expert in this area, but my knowledge of Monte Carlo approaches would suggest that it was appropriately applied.	No response required.
2			B	Yes, sufficiently large samples of threshold concentrations were considered and simulated for estimating the conditional probability of illness for replacing passing or failing lots respective to threshold log reductions.	No response required.
3			B	The code is efficient, runs accurately, and reproduces the results.	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 comminuted turkey 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.	No response required.
5			D	Monte Carlo simulation technique adequately modeled the scenarios. Please see below the comment for clarification and justification.	
6	55	1307	D	"Using Monte Carlo simulation, sample from a truncated form of initial concentrations, *x*, where its minimum is defined as negative infinity and its maximum is $\log_{10}(T)$. Multiply the vector of initial concentrations less than the concentration criteria threshold by a vector of samples from the attenuation distribution to simulate exposure doses from passing units." In this case truncation is valid as the minimum is defined as negative infinity. Please provide the minimum value at which truncation was done. Also, maximum is set for $\log_{10}(T)$. Does the maximum value has the potential to reach positive infinity or any truncation was done for the maximum value? Please provide this information. In addition, please provide more information on "passing unit" here, as it would be helpful to understand the context.	Truncation went from -Infinity to a range of log10(concentration thresholds). The absolute maximum concentration threshold simulated was 100 cfu/g.

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 comminuted turkey 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	57	1353–1354	D	"Point estimates and pass/fail results from 10 million Monte Carlo iterations are provided in Table 13." The number of iterations is sufficient.	No response required.
8			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	Simplified approaches were considered in the **section 5.4,** subsection *Comparison to model approximation 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 comminuted turkey 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:					
				and the use of lognormal distributions that span 12 orders of magnitude (+/- 3 standard deviations with sigma = 2 $\log_{10}$ units).	
g. What approach could be taken to assess uncertainty in these conclusions?					
1			A	I am not qualified to comment on this question.	No response required.
2			B	Posteriors samples from the Bayesian approach can be used to estimate the uncertainty. For this purpose, mc2d R Package can be used for the simulation to model the randomness (uncertainty) most commonly with gamma or beta distribution. Also, the variability in the absence of fewer data could have been modeled with Poisson distribution.	An uncertainty analysis has been developed and included in **Chapter 5 Final Product Standards**.
3	103	488	B	In the appendix describing the DR model, the uncertainty was simulated for 5,001 sets for 5,000 variable sets of transformed parameters for dose-response models and lower, and upper Confidence intervals have been created for multipliers for corresponding clusters 1 and 2. However for Dose-Response modeling for Final products only mean values are considered (R Code script FinalProductStds_CommTurkey.R,	A sensitivity analysis was conducted and included in **Chapter 5 Final Product Standards** assessing the effect of the dose-response model on the final product standard estimates and incorporated into the uncertainty estimation.

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 comminuted turkey 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:					
				line115–116). So, coefficient data is available from the Bayesian approach posteriors to estimate the uncertainty in the dose-response models.	
4			C	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 **Chapter 5 Final Product Standards** to address the overall uncertainty in the conclusions and sensitivity analysis was also 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 the update to improve clarity.
5	53	12747	C	Please clarify "substantial" uncertainty and its impact on the conclusions.	An uncertainty analysis was added 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 comminuted turkey 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	There are methods the authors may refer to assess uncertainty. For example, the knowledge uncertainty and stochastic variability for multiple simulations could be better 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 examples: 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	Thank you for the advice. FSIS carefully evaluated these expert suggestions and integrated them into the sensitivity and uncertainty analyses that have been added was added to **Chapter 5 Final Product Standards**.
7			E	Modeling uncertainty with respect to the conclusions from an entirely bottom-up approach (propagating uncertainty in each input toward an overall characterization of	Simplification was explored (**section 5.4**, subsection *Comparison to model approximation approach*) and the insights developed were incorporated into an

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 comminuted turkey 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:					
				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 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.	uncertainty analysis (to **Chapter 5 Final Product Standards**).
h. Are the conclusions drawn from the analysis appropriate?					
1			A	Again, based on my knowledge, everything appears to be adequately done here. I just cannot provide a lot of recommendations for improvements because my knowledge is limited in this area.	No response required.
2			B	Given the limited data for and feasibility of collection and risk management without too much burdening the industries, the suggested conclusion for changing the threshold reduction for final parts is reasonable to reduce ~16% of illness. As mentioned, variability around exposures and uncertainty around doses for dose-	An uncertainty analysis was 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 comminuted turkey 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:					
				response should be taken into account for modeling and more extensive estimates.	
3			C	It lacks well formulated conclusions, and therefore it is hard to evaluate whether they are appropriate.	**Chapter 8 Discussion** has been added to address this short coming.
4	20	521–529	C	Having said that, a conclusion seems to be that "a comminuted turkey performance standard that diverts test-positive lots based on a concentration threshold of 1CFU/15g is the most effective risk management option, with 2,700 illnesses avoided annually, which equates to slightly over 15% of the approximately 18,000 comminuted turkey illnesses estimated to occur annually. A comminuted turkey performance standard that diverts test-positive lots based on a concentration threshold (at the current LOD, or screening level) of 0.033 CFU/g is a similarly effective risk management option, with 2,500 illnesses avoided annually, which equates to 14% of the approximately 18,000 comminuted turkey illness that occur annually." The analysis is sound, but given that the data and assumptions behind it are not anchored in strong evidence I am not convinced it is really appropriate to use this	We agree. An uncertainty analysis was added to **Chapter 5 Final Product Standards** to best inform decision makers.

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 comminuted turkey 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:					
				as a basis for decision making. I would say it maybe is the best you can get, but, in the risk assessment, you would need to stress this is all uncertain. To support the decision makers, you should somehow give a judgment of the uncertainty and indicate how large it is.	
5	20	536–543	C	Another one that it is “infeasible to estimate the public health impact of performance standards that focus on serotype for all turkey products”. This is appropriate.	No response required.
6			D	The conclusions drawn from the analysis are reasonable and appropriate.	No response required.
7			E	There are limited conclusions related to comminuted turkey, and these are based on sound analysis and adequately explained.	No response required.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Evaluate whether the documentation of the data and modeling, and discussion, and interpretation of results is appropriate. If not, the reviewer must provide an alternative outline and/or approach for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
				General Comments	
1			A	The data provided is extensive. This is great. There are some places where data is described and provided, but how it was precisely used was not fully explained. For example, see my comments on use of the VFDB and refinement of those genes, but missing pieces about what exactly drove the analyses. It would be helpful to have additional descriptions/explanations of the files provided and what they contain. Overall, though, the team did a great job of providing the underlying data.	These comments were addressed in other responses. Thank you.
2			B	The report reads fine with some section organization issues that can be easily addressed with some reorganization, especially for model approach section and result section of Process control, as suggested following in the response 5.a. and 5.b.	No response required.
3			B	Overall great work with limited data for turkey.	No response required.
4	13	314	C	It seems that the fourth risk management question is not answered, I didn't find any reference to it.	This has been addressed in **Chapter 8 Discussion** that was added.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Evaluate whether the documentation of the data and modeling, and discussion, and interpretation of results is appropriate. If not, the reviewer must provide an alternative outline and/or approach for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
5	38–39	—	C	Figure 8 seems to contain more information than Table 8. In the table, the 2008–2009 baseline contains 1 finding of Muenchen, Agona and others, but in the bar chart their frequencies differ. I do not understand this, please explain.	Error in the graphic was corrected.
6			C	A main concern about the risk assessment report is that it lacks 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 are too vague, what does it mean that product standards “can therefore be considered with greater confidence by risk managers”? This is ambiguous. I note that there is more attention for scientific challenges like 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 by far the most important message to the risk managers that required this risk assessment. Clearly, if new methodologies	The conclusions have been summarized in a new **Chapter 8 Discussion** that is structured according to the risk management questions. Language has been added upfront plainly stating the limitations that result in some of the ambiguity. The quoted fragment has been restated as "are thus more informative to risk managers." While the **Executive Summary** does provide a complete overview of all findings and estimates in the risk assessment, a **Discussion Chapter** was added to address this imbalance. This concern boils down to the fraction of C1 among *Salmonella*. We’ve added a sensitivity analysis (**section 5.6**) exploring this effect. As outlined in previous comments: it does matters with respect to the calculated probability of illness, but it doesn’t change conclusions which

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Evaluate whether the documentation of the data and modeling, and discussion, and interpretation of results is appropriate. If not, the reviewer must provide an alternative outline and/or approach for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
				are introduced, they should be critically evaluated, but it would be helpful to also explicitly address 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. This seems a sound approach to me. But I get the impression that “the assumption that the average proportion of Cluster 1 versus Cluster 2 in any lot is approximately 0.3 and 0.7 (line 1210)” 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.	estimate the effectiveness of the risk management options.
7			D	The report is not well written, as there are many redundancies, in some cases lack justifications (as mentioned in comments to questions 1–5), 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.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Evaluate whether the documentation of the data and modeling, and discussion, and interpretation of results is appropriate. If not, the reviewer must provide an alternative outline and/or approach for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
8			E	To the extent that there is redundancy between the Chicken RA and the Turkey RA, the common components of the RAs should be identified as such explicitly, and the differences noted explicitly.	Use of model components that were developed with chicken data, due to turkey data limitations, have been identified in the document more explicitly.
		a. Is the report clearly written and complete?			
1			A	The report is enormous, and I appreciate the overall attention to detail. The report was broken down into executive summary, multiple main parts, and appendices. It was sometimes challenging to jump between these sections to identify information. However, I am not sure if there is a better way to structure this report.	No response required.
2			A	The writing is easy to read and understand. The team attempts to explain themselves throughout the report, which is appreciated. I think a disconnect occurred when referencing the work by EpiX. I did not gather much from the main report, then got confused in the EpiX appendix and had to go back to the main report, and so on. I feel like there can be better cohesiveness in these sections.	Additional text has been added to **Chapter 2** to clarify the EpiX Analytics Appendix and its utility in the full risk assessment. FSIS developed the *Bioinformatics Supplemental Materials (available here)* that will accompany the primary document, to provide further explanation and clarifications for items cited by reviewers.
3			B	The report is written well, and all areas are well addressed. The overall flow of the report is a little disconnected from the	A new **Chapter 8 Discussion** was added that summarizes the conclusions from each section

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Evaluate whether the documentation of the data and modeling, and discussion, and interpretation of results is appropriate. If not, the reviewer must provide an alternative outline and/or approach for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
				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.	of the report in the same order as the risk management questions.
4	27	787	B	The model approach section could be restructured to enhance the flow of the reader, by describing the three research management questions in the coherence of the scenarios. The sequence of paragraphs seems arbitrary which makes it difficult to comprehend, however, all pieces of information are provided, yet they are disconnected and can be better organized. If they can be organized in the same order as suggested in the response to the following question, it will make more sense.	Several improvements were made in the structure and organization of the document in response to reviewer suggestions.
5	70–76	1652–1745	B	The result section of Process control is not written coherently. Text and figure captions are all mixed and all over the place, making it difficult to comprehend very clearly. If it can be restructured for more clarity about risk management scenarios for both indicator organisms, it will be good. The overall takeaway message from these results is also not very clearly stated, so that needs improvement.	Figure captions have been corrected. The text was edited.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Evaluate whether the documentation of the data and modeling, and discussion, and interpretation of results is appropriate. If not, the reviewer must provide an alternative outline and/or approach for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
6			C	As a non-native English speaker, I do now and then struggle with the terminology. As the foreseen readership of this report will be in the United States, that is probably OK, but please be aware of it.	No response required.
7			C	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 have been clarified throughout the document.
8	16	409–413	C	Where is the answer to risk management question #4?	This has been addressed in a Discussion **Chapter 8** that was added to the document.
9	29	831	C	Figure 2 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, 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 suggest that you estimate the annual number of illnesses from the exposure + dose-response	The figure has been revised to address these considerations.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Evaluate whether the documentation of the data and modeling, and discussion, and interpretation of results is appropriate. If not, the reviewer must provide an alternative outline and/or approach for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
				relation, but you don't; the annual number of illnesses is estimated from the epidemiological data. Please revise this figure.	
10	30	832	C	Part I does not include an exposure assessment as defined in the glossary. The word exposure occurs only twice in this part.	We have now included text that identifies the exposure distribution and risk characterization steps in our model explicitly (see **Chapter 5**).
11	46	1107–1118	C	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, 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.	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. 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%).

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Evaluate whether the documentation of the data and modeling, and discussion, and interpretation of results is appropriate. If not, the reviewer must provide an alternative outline and/or approach for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
					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.
12			D	Please check for these errors and correct accordingly.	The following typos were corrected:
13	3	22	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.	Deleted.
14	10	—	D	Hazard Identification-The identification (of) biological agents capable…missing "of".	Corrected.
15	11	—	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" .	Deleted.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Evaluate whether the documentation of the data and modeling, and discussion, and interpretation of results is appropriate. If not, the reviewer must provide an alternative outline and/or approach for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
16	11	—	D	Pathogen Reduction; Hazard analysis and Critical Control Point (PR;HACCP): Please change to “Analysis” instead of “analysis”.	Corrected.
17	36	936	D	Please delete the period before “to”.	Deleted.
18	54	1279	D	Please complete “(Ebel Refs):”.	The missing reference was added.
19	81	1956–1960	D	The same reference Thompson et al. in the reference list has been written back to back. Please delete one.	The reference list was corrected.
20	129	183	D	Table 28: The average daily turkey consumption in grams of turkey commodity on a population basis. What is GRM[4]? What is superscript 4?	The superscript has been corrected to an “a.”
21	129	196	D	Table 30: Average daily consumption in grams turkey commodity by consumer domain. Same as above. What is GRM[4]? What is superscript 4?	
22	134	307	D	Please add “comma” before “respectively.”	The comma was corrected.
23			E	The document is clearly written and complete, with the exception of uneven treatment of certain key assumptions already noted.	No response required.
	b. Does the report follow a logical structure and layout?				
1			A	Yes, the layout is generally fine. See my comments above.	No response required.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Evaluate whether the documentation of the data and modeling, and discussion, and interpretation of results is appropriate. If not, the reviewer must provide an alternative outline and/or approach for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
2	27	771–777	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 (Ch5) Q2. About Final product ~(Ch4) Q3. About the process control ~ (Ch6)	Thank you for your suggestion. The 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 (Ch5) à Ch4 Final product (Ch4)à Ch5 Process control (Ch6)àCh6	
4			C	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. I would like to see a separate section with “Conclusions” or “Answers to the questions” that is not a discussion of what you did (as p. 22–23) but gives the answers to the questions. It should be explained why you develop a new DR model and where and how you use	A Discussion **Chapter 8** was added that summarizes the conclusions from each section of the report in the same order as the risk management questions. A sensitivity analysis was added that considers the impact of the dose-response relationship on the model results. Motivation for the development has been addressed.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Evaluate whether the documentation of the data and modeling, and discussion, and interpretation of results is appropriate. If not, the reviewer must provide an alternative outline and/or approach for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
				it, and then evaluate whether it was worth the 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	
5	46–47 56	1117–1127 1332–1339	D	There are many redundancies in this report. Please check this throughout the report and revise accordingly. For example: "It is estimated there are 42,669 turkey-associated *Salmonella* illnesses per year based on latest IFSAC attribution rate (0.059). This value is calculated as the product of total 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 then divided by the FoodNet population coverage (15 percent). The total cases are subsequently allocated by commodity using NHANES consumption statistics. In particular, that 0.42 of all turkey-associated *Salmonella* illnesses result from exposure to comminuted (ground) turkey products, which is approximately 17,921."	Given the length of the document, certain key points are intentionally repeated for clarity.
6			E	There is no fundamental issue with the structure and layout of the report. Many readers will be interested in both the	The two risk assessments are structured the same way to aid parties interested in the

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Evaluate whether the documentation of the data and modeling, and discussion, and interpretation of results is appropriate. If not, the reviewer must provide an alternative outline and/or approach for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
				chicken and turkey RAs and a structure that makes it easier to identify the key points of overlap, as well as the key differences would be welcome.	comparison. The key repetitions have been mentioned throughout this report.
	c. Are the conclusions supported by the risk assessment?				
1			A	I believe that the conclusions are supported by the assessment. I think FSIS was cautious in their recommendations based upon the results of this work, which is good. But, it will also leave some readers wondering "so what?" The main conclusions / recommendations from this are 1) implement new Enterobacterial count standards and 2) consider diverting lots which fail. This results in modest predicted reductions in illness. How does this work further our overall efforts to improve *Salmonella* food safety? How does it get to the "strain-level" mitigations that are currently being considered at the regulatory level? I am not convinced that it does these things.	Thank you for the comment. The stated goal of this quantitative risk assessment was to answer the risk management questions to the extent possible given the underlying data. This resulted in modest reductions in illnesses from final product standards on comminuted turkey and marginal reductions in *Salmonella* on turkey carcasses via process control, which requires more data to support. The analysis also uncovered different data and research gaps that would assist overall efforts in improving *Salmonella* food safety in turkey products. Strain-level mitigations have multiple components that need to be assessed when modeling. For example, the efficacy of preharvest interventions, such as vaccinations or other mitigation measures on specific strains, is an evolving field of study. Additionally, without more robust data regarding the serotype mixture in comminuted turkey lots (or turkey flocks), one can only make assessments on the final product samples. We have included additional analysis

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Evaluate whether the documentation of the data and modeling, and discussion, and interpretation of results is appropriate. If not, the reviewer must provide an alternative outline and/or approach for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
					investigating key strains of interest for risk managers based on the data collected (see Section 5.5). Strain level final standards are a subset of concentration-based standards. The virulence capacity is still being studied.
2			B	As for overall conclusions for this QMRA, only a robust risk management scenario was suggested for comminuted turkey using concentration-based final product standards for *Salmonella*. For other two products due to limited data availability robust modeling was not possible. The suggested concentration threshold is shown to reduce ~16% comminuted turkey-associated illness per year.	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 (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	An uncertainty analysis was developed to address this concern (**section 5.7**). 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.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Evaluate whether the documentation of the data and modeling, and discussion, and interpretation of results is appropriate. If not, the reviewer must provide an alternative outline and/or approach for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
				Aug;27(4):831–44. doi: 10.1111/j.1539-6924.2007.00926.x. PMID: 17958495.	
4			C	Yes. But see my general comment above about the lack of a section that provides the answers to the risk management questions in a clear way.	A Discussion Chapter was added that summarizes the conclusions from each section of the report in the same order as the risk management questions.
5	22–23 23	613–640 642–649	D	The conclusions are reasonable given the scope of the work. Conclusions are presented only in the Executive Summary (Pages 22–23, lines 613–640). A separate conclusion section in the main report is missing. Please include that. Limitations: The limitations of the risk assessment, data, modeling approach, results and interpretations are missing in the main report. The limitations have only been mentioned in the executive summary (page 23, lines 642–649). Please provide the limitations in the main report.	A separate Discussion Chapter has been included, which outlines the stated limitations
6			E	The conclusions are supported by the risk assessment. Equally, the lack of conclusions (inability to answer certain RM questions) is supported by the lack of evidence.	No response required.
d. Is the documentation of the assumptions clear and complete?					

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Evaluate whether the documentation of the data and modeling, and discussion, and interpretation of results is appropriate. If not, the reviewer must provide an alternative outline and/or approach for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
1			A	Assumptions are clearly identified and are complete. I don't agree with all of them, but they are documented.	No response required.
2			B	All assumptions are mentioned and documented in the report. However, the different assumptions about data such as the use of historical data and missing data are made throughout the report and were an effort to find while reading. A consolidated table in the introduction chapter describing all assumptions about data and modeling scenarios and parameter estimations could help readers more to be aware of all assumptions.	Both tables were added to the document.
3			C	Documentation of the assumptions is scattered; it is very challenging to get an overview.	A table of assumptions has been added to the document.
4			C	A table like Table 24 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.	A table of assumptions has been added to the document.
5			D	Please refer to the comments in Questions 1–4.	No response required.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Evaluate whether the documentation of the data and modeling, and discussion, and interpretation of results is appropriate. If not, the reviewer must provide an alternative outline and/or approach for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
6			E	As discussed in earlier comments, the assumptions are generally transparently stated, clear and complete. However, the degree to which certain assumptions are discussed is "uneven" (e.g., the use of a global chicken attenuation multiplier for comminuted turkey). The authors may be aware of how certain assumptions are viable and do not threaten the credibility of conclusions, but this is not made apparent to the reader.	This inequity has been addressed throughout the report.
e. Is the documented dose-response, exposure assessment, and risk characterization modelling transparent and reproducible?					
1			A	Yes.	No response required.
2			B	Exposure assessment was limited for carcasses and comminuted turkey due to limited data availability and the effect of the pandemic on data collection.	No response required.
3	37	958	B	The use of historic data for carcass is justified for baseline estimations. Also, the imputation methods used to fill the gaps of comminuted turkey MPN data make sense, however, which method has been used should be mentioned clearly.	The test 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. https://doi.org/10.18637/jss.v045.i03

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Evaluate whether the documentation of the data and modeling, and discussion, and interpretation of results is appropriate. If not, the reviewer must provide an alternative outline and/or approach for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
4			B	All submitted scripts run correctly and are reproducible to generate correct plots and data files.	No response required.
5			B	The dose-response model for Final Product standards runs perfectly and is in sync with the conceptual model as explained in section 5.3. Limitations of lack of data are given for not considering the variability and the model gives the point estimates. However, no modeling has been done to consider the uncertainty in the estimates, while the serotype's cluster multipliers for confidence Intervals (CI) are available and can be included in DR model to calculate the CI . In script 'FinalProductStds_CommTurkey.R' line 115 and 116 only '*est="Mean"*' has been used to calculate the variable *'p.ill.dose.C1'* and this value was incorporated in the lines 176 ,182 and in following lines to calculate the 'avoided illnesses' while simulating the different thresholds of p.ill of passing lots in the function in lines 144–212. Within same function upper and lower CI bounds value provided in the imported dataframe 'Poly' from serotype cluster multiplier, can be used	A sensitivity analysis (**section 5.6**) has been added to the report addressing this comment.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Evaluate whether the documentation of the data and modeling, and discussion, and interpretation of results is appropriate. If not, the reviewer must provide an alternative outline and/or approach for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
				to calculate the CI bounds of 'avoided illnesses'. It is mentioned in the text that uncertainty will be considered for the future work, will that work be done with some new assumptions or already available CI bound data will included and estimates will be revised. Please clarify.	
6	61	1454–1455	B	Please add appropriate referencing or supporting findings or rationale for the suggestion of improving the testing measures. It's a direct statement made, without mentioning the basis, whether it's inferred form the lab work or from a reference, please add clear background for this statement.	The highlighted scenarios are not a suggestion for improving testing measures. Limitations of the model were explored on a range of concentration scenarios from 1cfu/2,600g to 100cfu/g, many of which are not practical.
7	37	973–975	C	Please explain how you obtain the indicated parameter values for the lognormal distribution of concentrations. (I find a description in Appendix C, I would say that this explanation should not be hidden there). In Figure 7, add the thresholds for 1 cfu per turkey and for the prevalence estimate.	The derivation of analytical methods is summarized in **Appendix C** to best tailor the structure of the document to risk managers' needs. The figure has been corrected.
8	37–38		C	The exposure assessment for the comminuted turkey is a combination of a distribution of what is found on the meat (line 973–982) and the attenuation	No response required.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Evaluate whether the documentation of the data and modeling, and discussion, and interpretation of results is appropriate. If not, the reviewer must provide an alternative outline and/or approach for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
				distribution that was derived earlier. This is transparent and reproducible.	
9	37	950–952	C	An exposure assessment of carcasses is not possible and not performed. This is transparent and reproducible yet striking. For the derivation of the dose-response relation the *Salmonella* concentration distribution in chicken meat is used, anchored in the data on the number of cases for turkey. This gives me the impression that you assume something here. This is not clear for me.	Data limitations that required the use of chicken data have been outlined in the new Assumption Table (**section 1.6**).
10	45–47		C	The chapter on exposure assessment does not describe an exposure assessment, it is more related to something I would call risk characterization. See the definitions in the Glossary. The actual exposure assessment is described in lines 1220–1227, in a section on hazard characterization. The authors should either refer to exposure assessment only in a chapter with the title “exposure assessment” or change the title of the chapter.	We have now included text that identifies the exposure distribution and risk characterization steps in our model explicitly (**Chapter 5**).
11			C	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.

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Evaluate whether the documentation of the data and modeling, and discussion, and interpretation of results is appropriate. If not, the reviewer must provide an alternative outline and/or approach for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
12	Section 4.3 and other places		C	It is not clear how the baseline probability of illness, given in Table 13, is derived. The section on baseline probability of illness (4.3) does not give the baseline probability of illness, just number of illnesses. You cannot expect a reader to do these crucial calculations her/himself.	This oversight has been corrected. The empirical probability of illness estimates are the ratio of total turkey illnesses to total turkey servings (Hsi, 2015).
13			C	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 “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.	We have now included text that 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. Clearly, our finished product standard model requires the calculation of probability of illness per serving before (baseline) and after (new) implementation of a concentration-based standard.
14	15	359–363	C	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	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. We’ve

Comment #	Page #	Line(s) #	Reviewer ID	Comment	FSIS Response
Q5. Evaluate whether the documentation of the data and modeling, and discussion, and interpretation of results is appropriate. If not, the reviewer must provide an alternative outline and/or approach for adequately and clearly documenting this risk assessment. Specific consideration should be given to the following:					
				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.	reported illustrative examples of these probability of illness per serving estimates in our results table for the concentration-based finished product standards (**section 5.4**). It is true that the number of illnesses for the turkey 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.
15			D	The documented dose-response, exposure assessment, and risk characterization modelling is transparent and reproducible but need changes as documented in comments to questions 1–4.	No response required.
16			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		☐	☐

List of projects/relationship ⇩ and funding type ⇨		Grant	Contract
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	1
R coding	2	3	3	3	3	3	2	3	3	3
Dose-response modeling	3	3	3	2	3	3	3	1	2	2
Bioinformatics: Machine learning methods for genomic data (e.g., random forest modeling)	3	1	2	3	2	2	1	1	2	3
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	2	2	2	3
Epidemiology and surveillance of salmonellosis	3	2	2	3	2	2	2	3	3	3
Knowledge of chicken production and/or slaughter processes**	3	1	2	2	2	3	3	3	2	3
Knowledge of turkey production and/or slaughter processes	2	1	2	2	2	2	2	2	1	3

*Selected experts for the peer review.
** Not relevant to 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	>20	<10	10	>20
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*										
	Government agencies related to monitoring or controlling *Salmonella* in chicken and/or turkey meat**			X	X		X	X		X	
	Any other relevant information that you would like to disclose	X									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 Microbiological Risk Assessment for *Salmonella* in Raw Turkey and Raw Turkey Products

1. **Turkey_SRA_Review_1.30.23 5pm_RTI**: PDF document describing the QMRA; this is the main document you need to review
2. **FinProdStds_CommTurkey.R**: R code for the final product portion of the QMRA
3. **Polynomial2:** CSV file with polynomial coefficients for dose-response model in FinProdStds_CommTurkey.R
4. **ProcessControl_turkey.R**: R code for the process control portion of the QMRA
5. **Turkey_carcass_process_control**: CSV file with data for process control portion of the QMRA
6. **TURKEY_NHANES_Consumption:** Word document with the SAS code used for the NHANES serving size estimates
7. **01_NCBI_Parse:** R Markdown describing the initial processing of isolate assembly metadata
8. **02_Sistr_Parse**: R Markdown illustrating the serovar prediction and QC check from SISTR
9. **03_example_Prokka_Slurm.slurn**: Example code for gene annotation and identifying virulence factors in isolates
10. **04_clustering_code_example.R**: R code for unsupervised random forest and analyzing cluster stability
11. **NonRedundant_VFDB_PATRIC.faa**: Fasta file of all virulence factors considered in the algorithm
12. **poultry_VF**: CSV file with the basic information/description of all virulence factors considered
13. **RF_input_193**: CSV file with the presence/absence matrix used as input in the random forest
14. **sal_prodigal_training.trn**: Prodigal training file on the reference *Salmonella* assembly
15. **sistr_poultry_cat**: CSV file with the output resulting from the SISTR prediction algorithm