Quantitative Risk Assessment for *Salmonella* in Raw Chicken and Raw Chicken Products

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

USDA Food Safety and Inspection Service

Contributors to the Risk Assessment

Anant Venu

Natalie Baker

Tameru Berhanu

Iva Bilanovic

Eric Ebel

Sunil Kumar

Davi LaBarre

Drew Posny

Gurinder Saini

Michael Williams

Joanna Zablotsky Kufel

Acknowledgements

The authors would like to acknowledge and thank the following for their assistance, contributions, and review in the preparation of this document:

EpiX Analytics partners, Francisco Zagmutt, Régis Pouillot, Jane Pouzou, Dan Taylor, Gavin Fenske, and Solenne Costard, for their work to advance the use of whole genome sequence data for *Salmonella* in quantitative microbial risk assessment for food safety. Specifically, we thank this team for their collaboration under FSIS Cooperative Agreement (FSIS-02152022). Their expertise in applying advanced bioinformatics to make a genomic determination of which *Salmonella* are more virulent was instrumental in enhancing the hazard identification component of this risk assessment. Their use of this genomic information to develop an innovative *Salmonella* dose-response model that takes into consideration differences in virulence among *Salmonella* serotypes greatly enhanced the hazard identification and hazard characterization components of this risk assessment.

University of Maryland Joint Institute for Food Safety and Applied Nutrition (UMD-JIFSAN) collaborators: Clare Narrod (UMD-JIFSAN), Wendelyn Jones (Institute for the Advancement of Food and Nutrition Sciences; IAFNS), Caitlin Karolenko (IAFNS), Joe Scimeca, and Andrea Stumpf (Structured Partnerships), and Eric Owusu (UMD-JIFSAN). Specifically, we thank them for establishing strategic partnerships and facilitating discussions on industry data sharing between the FSIS and key industry representatives. This work was conducted under FSIS Cooperative Agreement FSIS-02152022 in an effort to further strengthen the exposure assessment component of this risk assessment.

BioMérieux scientists for their contributions to enhance our understanding and interpretation of the *Salmonella* level data used in this risk assessment. Specifically, we thank Stan Bailey, Senior Science Advisor, for engaging in open discussions with FSIS' risk assessment team to answer questions concerning the bioMérieux Gene-Up Quant *Salmonella* assay for level of *Salmonella* in poultry samples. Adam Joelsson and his research team in Philadelphia and Vik Dutta and John Mill from the Scientific Affairs team in St. Louis were instrumental in providing our team with insight on how the Quant assay algorithm was developed to correctly interpret data on the amount of *Salmonella* in chicken for use in this quantitative risk assessment. This exchange of information serves as a model for private industry and government collaboration on science and information sharing to improve public health.

Interagency Risk Assessment Consortium members, including Steven Foley (Food and Drug Administration (FDA), National Center for Toxicological Research); Marc Allard (FDA, Center for Food Safety and Applied Nutrition); Michael Ollinger and Sandra Hoffmann (United States Department of Agriculture (USDA), Economic Research Service) and Dayna Harhay, Jonathan Frye, and Thomas Oscar (USDA, Agricultural Research Service) who generously offered their time to review drafts of the risk assessments.

USDA-ERS economists, specifically Michael Ollinger and Sandy Hoffman, for sharing studies related to behavior economics to further inform the development of this risk assessment.

We thank the USDA's Office of Risk Assessment and Cost Benefit Analysis, specifically Mark Powell and Linda Abbott, for providing practical and useful input during consultations early in the development of

this risk assessment.

FSIS thanks Margo Schwab, Senior Science Policy Analyst with the Office of Management and Budget, Office of Information and Regulatory Affairs, for her guidance to support information sharing for transparency and stakeholder engagement, while ensuring equitable access to high quality and credible findings from FSIS' risk assessment.

The Centers for Disease Control and Prevention, specifically Patricia Griffin, Daniel Payne, Erika Rose, Hazel Shah, and Cary Devine, for facilitating access to and assisting in the interpretation of several public health surveillance datasets, including FoodNet and NORS.

USDA Agricultural Research Service (ARS) scientists whose research on *Salmonella* in poultry helps to fill FSIS research priority needs. We appreciate the studies they're conducting under an FSIS/ARS Interagency Agreement in support of the FSIS *Salmonella* Initiative. ARS scientists have expanded their *Salmonella* research to include poultry and are conducting projects investigating *Salmonella* serotype distribution and *Salmonella* detection methods development. Ongoing studies on *Salmonella* enumeration will ensure that fit-for-purpose methods are available.

The Office of Policy and Program Development (OPPD), Policy Analysis Staff, specifically Andrew Pugliese, Angelica Marrero, and Getachew Nigatu, for working closely with the risk assessment team to ensure alignment between the risk assessment and benefit-cost scenario analyses evaluating final product standards.

FSIS OPPD Risk Management and Innovations Staff, specifically Stephanie Hretz, Peter Evans, and Melvin Carter, for providing data from and sharing insights on the *Salmonella* Initiative Program (SIP) and developing and providing clarifications on the risk management questions.

FSIS Office of Field Operations, specifically Amber Pasko, Gabrial Eddings, and Parrish Endy for their input related to poultry processing, as well as Frontline Supervisor Chelsea Buckley and Supervisory Public Health Veterinarian Andrew Gordon who enabled the risk assessment team to gain firsthand insights to poultry slaughter and processing.

FSIS' Office of Public Health Science (OPHS) Laboratory Regulatory Operations, including Laboratory Quality Assurance, Response, and Coordination Staff and the laboratories for their work on the exploratory sampling, interpreting laboratory data, and method validation work for this risk assessment. We specifically thank William Shaw, Todd Lauzé, Hannah Anderson, Tye Boynton, Michael Day, and Jeanetta Tankson.

FSIS, OPHS, Microbiological and Chemical Hazards Staff and the rest of the OPHS team for leadership on the conduct of a risk profile *Salmonella* in poultry to provide a review of the scientific literature and foodborne illness data to enhance the hazard identification component of this risk assessment. We specifically thank Evelyn Mbandi, John Jarosh, Lesley Good, Wu San Chen, Neal Golden, Bryce Merrill, and Courtney Johnson.

Peter Evans, FSIS OPPD, for his support in crafting responses to the external peer review of the risk assessments while on detail with OPHS and sharing his bioinformatics knowledge and insights.

Wayne Schlosser (formerly with FSIS, OPHS, Risk Assessment and Analytics Staff, retired) who spearheaded the launch of this risk assessment and provided guidance and insights valuable to its development.

Janell Kause, FSIS OPHS, for her guidance and support from the beginning and throughout the conduct of this risk assessment, including FSIS' collaboration with the University of Maryland.

FSIS OPHS leadership, including Isabel Walls, Denise Eblen, Kis Robertson-Hale, and Emilio Esteban (formerly OPHS Chief Scientist) for their guidance for the conduct of this risk assessment and review to enhance the preparation of this report.

Table of Contents

List of Figures

List of Tables

Glossary[1]

Foodborne Illness Source Attribution	Identification of which foods are the most important sources of selected major foodborne illnesses.
Colony Forming Units (cfu)	Colony forming units (cfu) is an estimation of the number of viable microbial cells in a sample. They are typically expressed as a rate per unit of volume or mass such as cfu per gram (cfu/g) or cfu per milliliter (cfu/mL).
Dose-Response Assessment	The determination of the relationship between the magnitude of exposure (dose) to a microbiological organism and the severity and/or frequency of associated adverse health effects (response).
Exposure Assessment	The qualitative and/or quantitative evaluation of the likely intake of a microbial hazard via specific foods. It provides an estimate of the likelihood and level of the hazard in a specified portion of that food. The exposure assessment may also identify the frequency and amount of food consumed in a given period for a given (sub)population and may combine the information to estimate the population exposure to a microbiological hazard. The exposure assessment details the various steps of the farm-to-fork pathway so that the effect of pertinent steps/processes, or changes to them can be assessed.
Indicator Organism	Indicator organisms, such as Enterobacteriaceae (EB) or aerobic counts (AC), have been used as gauges of process control and to measure the microbial reduction from carcasses at slaughter to post-chill.
Infectivity	The ability of an organism to cause infection. In risk assessments, this is incorporated as the probability of human infection following oral exposure to any amount of *Salmonella*. This probability can vary depending on pathogen factors such as the serovar or subtype, and host susceptibility.
Hazard Characterization	The qualitative and/or quantitative evaluation of the nature of the adverse health effects associated with the hazard. For the purpose of microbial risk assessment, it provides a description of the adverse effects that may result from ingestion of a hazard, whether that is a microorganism or its toxin. This should include a dose–response relationship where possible. Those health effects include, for example, diarrheal illnesses, hospitalizations, and deaths. In the context of QMRA

[1] The definitions compiled in this glossary are largely consistent with those adopted in the 2023 NACMCF response to FSIS questions on *Salmonella* in poultry products (National Advisory Committee on Microbiological Criteria for Foods (NACMCF). (2023). *Response to questions posed by the Food Safety and Inspection Service: Enhancing Salmonella control in poultry products*. https://www.fsis.usda.gov/sites/default/files/media_file/documents/NACMCF%20Salmonella-Poultry17Mar2023.pdf).

	are usually considered to be acute, rather than chronic, health effects. This component may include identification of different adverse effects, including sequalae and their likelihood, for different subpopulations, such as neonates or immunocompromised people.
Hazard Identification	The identification of biological agents capable of causing adverse health effects and which may be present in a particular food or group of foods. It is a qualitative process intended to identify microbial hazards (infectious agents or toxins produced by microorganisms) of concern in food.
Limit of detection (LOD)	LOD is the lowest level of microbial cells that can be reliably detected using a standard test.
Limit of quantification/ quantitation (LOQ)	LOQ is the lowest level of microbial cells that can be quantified based on predefined goals 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.
Pathogenicity	The ability of an organism to cause disease. In risk assessments, this is usually modeled as the probability of clinical disease given infection. This probability can vary depending on pathogen factors such as the serovar or subtype, and host susceptibility.
Pathogen Reduction; Hazard Analysis and Critical Control Point (PR/HACCP)	The PR/HACCP rule, fully implemented in 1996, was designed to reduce the occurrence and numbers of pathogenic microorganisms, harmful bacterial, on meat and poultry products, reduce the incidence of foodborne illness associated with the consumption of meat and poultry products, and provide a new framework for modernization of the current system of meat and poultry inspection.
Public Health Information System (PHIS)	The Public Health Information System (PHIS), a dynamic, comprehensive data analytic system, was launched as part of FSIS' effort to collect, consolidate, and analyze data in order to improve public health.
Post-chill	Post-chill refers to the point in the process where the young chicken carcasses exit the chiller after all slaughter interventions have taken place but before entering coolers or proceeding to further processing.
Positive Predictive Value (PPV)	For a microbiological assay, the positive predictive value (PPV) considers the sensitivity and specificity of the assay as it relates to the level distribution (**section 3.2** . The relationships that define PPV for the application to qPCR enumeration technology here is

	$PPV = \frac{the\ fraction\ of\ true\ positive\ samples\ above\ the\ LOQ}{the\ fraction\ of\ \ positive\ samples\ declared\ above\ the\ LOQ} = \frac{true\ positives}{true\ postives + false\ positives}.$
Prevalence	The frequency of a disease in a population at a particular time point. Often expressed as a proportion or percentage.
Process control	Process control is a defined procedure or set of procedures designed by an establishment to provide control of those operating conditions that are necessary for the production of safe, wholesome food. The procedures typically include some means of observing or measuring system performance, analyzing the results generated in order to define a set of control criteria, and taking action when necessary to ensure that the system continues to perform within the control criteria.
Quantitative microbial risk assessment (QMRA)	Quantitative microbial risk assessment is a mathematical modeling approach used to estimate the risk of infection and/or illness when a population is exposed to microorganisms from a variety of sources, in this case in foods. QMRA estimates can be used to predict the potential reduction (increase) in foodborne illnesses resulting from the implementation of strategies to mitigate foodborne pathogens in foods.
Receiving	The point in the slaughter process where poultry arrive at the establishment in transport cages, are unloaded, and are hung on shackles before stunning and bleeding.
Rehang	Rehang refers to the location in the process after the hock cutter and prior to evisceration
Risk Assessment	A decision-support tool to provide risk managers with a rational and objective picture of what is known about a health risk and its causes at a particular point in time (FAO/WHO 2021). A scientifically based process consisting of the following steps: (i) hazard identification, (ii) hazard characterization, (iii) exposure assessment, and (iv) risk characterization.
Risk Characterization	The process of determining the qualitative and/or quantitative estimation, including attendant uncertainties, of the probability of occurrence and severity of known or potential adverse health effects in a given population based on hazard identification, hazard characterization and exposure assessment.
Serocluster	Genomics classified group of *Salmonella* serotypes, grouped based on *virulence* similarities (see **Chapter 2** or **Appendix A** for more detail)

Serotype	See serovar definition.
Serovar	The term serovar is used to distinguish groups within a *Salmonella* species that share distinctive surface structures, namely the O surface antigen and the H antigen that is part of the flagella. Consequently, serovars represent phenotypical differences between individual bacteria belonging to a *Salmonella* species, and do not necessarily represent evolutionary differences as elucidated in the *Salmonella* genome. Note that in this report, the term serovar and serotype are used interchangeably.
Subtype	*Salmonella* subtype is a term used to distinguish differences within a serovar (serotype), such as defined using Whole Genome Sequencing (WGS), Pulsed-field gel electrophoresis (PFGE) or Multi-Locus Sequence Typing (MLST). Subtyping provides a more detailed characterization of heterogeneity between *Salmonella* bacteria than serovar groupings as it is based on genetic differences.
Test Sensitivity	The probability that a test performed on a contaminated sample will yield a positive result. For qualitative (positive/negative) test results, this probability is affected by the limit of detection of the test, whereas for quantitative test results the probability is affected by the limit of quantification.
Virulence	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 severity.

Executive Summary

Introduction

The United States Department of Agriculture's (USDA) Food Safety and Inspection Service (FSIS) protects the public's health by ensuring that meat, poultry, and egg products are safe, wholesome, and properly labeled. The Agency is committed to reducing foodborne infections associated with FSIS-regulated products, in particular *Salmonella* illnesses attributable to poultry.

Salmonella on poultry remains a significant food safety concern in the United States (U.S). The U.S. Centers for Disease Control and Prevention (CDC) estimates *Salmonella* infection is responsible for over 1 million illnesses, 26,500 hospitalizations, and 420 deaths in the U.S. every year (Scallan, 2011). Of these illnesses, an estimated 66 percent are from food (Beshearse, 2021), with over 17 percent from eating chicken and almost 6 percent from eating turkey (IFSAC, 2022). This imposes an estimated $3.7 billion in economic burden in a typical year. Almost 90 percent of this burden, $3.3 billion, is due to deaths; 8 percent, $294 million, is due to hospitalization; and the remaining 2 percent is due to non-hospitalized cases (Hoffmann, 2021).

Reducing foodborne illness from *Salmonella* in poultry products remains a public health priority. On October 17, 2022, FSIS proposed a regulatory framework to control *Salmonella* in poultry products and more effectively reduce foodborne *Salmonella* infections linked to these products (FSIS, 2022c). Central to this effort is this quantitative microbial risk assessment (QMRA) that provides information addressing risk management questions on the predicted public health impact of controlling the prevalence, levels, and/or specific serotypes of *Salmonella* on chicken presented for slaughter, on chicken carcasses throughout the slaughter process, and/or in final chicken products. This QMRA was refined in response to external peer review (available here) provided in accordance with Office of Management and Budget (OMB) information quality guidelines[2] and in response to interagency review through the Interagency Risk Assessment Consortium.

Risk Management Questions

This risk assessment addresses the following risk management questions:

1. What is the public health impact (change in illnesses, hospitalizations, and deaths) achieved by eliminating at receiving a proportion of chicken contaminated with specific levels of *Salmonella* and/or specific *Salmonella* subtypes?

2. What is the public health impact (change in illnesses, hospitalizations, and deaths) achieved by eliminating final product contaminated with specific levels of *Salmonella* and/or specific *Salmonella* subtypes?

3. What is the public health impact of monitoring/enforcing process control from rehang to post-chill? Monitoring could include analytes such as *Enterobacteriaceae*, Aerobic Count, or other indicator organisms, analysis could include presence/absence or levels and the monitoring could

[2] https://www.archives.gov/files/federal-register/executive-orders/pdf/12866.pdf
https://www.whitehouse.gov/briefing-room/presidential-actions/2023/04/06/executive-order-on-modernizing-regulatory-review/

also include variability of actual result versus expected result, log reduction, absolute sample result, or other individual establishment specific criteria.

4. What is the public health impact of implementing combinations of the risk management options listed above?

Structure and Scope

This is a quantitative probabilistic food safety risk assessment. It examines the relationship between the amount of *Salmonella* (hereafter referred to as the 'level' or 'concentration' of *Salmonella*) and the presence of certain *Salmonella* serotypes on chicken received for slaughter and/or on chicken products (i.e., chicken carcasses, chicken parts, and comminuted chicken) and the probability of foodborne illness. It also examines the relationship between changes in microbiological indicator organisms (i.e., aerobic count) on chicken carcasses from rehang to post-chill and changes in foodborne illnesses.

This risk assessment contains the traditional four components identified in the Codex *Principles and Guidelines for the Conduct of Microbiological Risk Assessment CAC/GL-30* (FAO/WHO, 1999) and in the Food and Agricultural Organization of the United Nations and World Health Organization *Microbial risk assessment – Guidance for food* (FAO/WHO, 2021):

- hazard identification,
- exposure assessment,
- hazard characterization, and
- risk characterization.

Where possible, these four components are referenced in relation to specific sections of the risk assessment. However, these components were developed for a traditional mechanistic risk modeling approach; and this risk assessment does not directly utilize that approach, given the broad scope of the risk management questions. Therefore, this document is organized according to the risk management questions provided above.

The hazard identification identifies the *Salmonella* associated with foodborne illness from consuming chicken. This risk assessment leveraged FSIS' *Risk Profile for Pathogenic Salmonella Subtypes in Poultry* (available here) to identify *Salmonella* serotypes in chicken linked to foodborne illness. This independently peer-reviewed risk profile provided a comprehensive review of the scientific literature and foodborne illness data to identify certain *Salmonella* serotypes in poultry linked to foodborne illness. FSIS expanded on this work through a Cooperative Agreement (FSIS-02152022) with the University of Maryland's Joint Institute for Food Safety and Applied Nutrition (UMD-JIFSAN), in partnership with EpiX Analytics to differentiate *Salmonella* serotypes by virulence using advanced bioinformatics (i.e., machine learning) to evaluate genomic data.

When considering risk management options (also identified in this risk assessment as 'scenarios'), the exposure assessment provides a characterization of the amount of *Salmonella* consumers are exposed to from each chicken product serving. This exposure assessment characterizes current *Salmonella* contamination levels (colony forming units per gram (cfu/g)) on carcasses at the rehang and post-chill

slaughter steps and in final chicken products (i.e., chicken parts and comminuted chicken). Special consideration is given to the proportion of higher virulence *Salmonella* serotypes in each product. National Health and Nutrition Examination Survey (NHANES) U.S. dietary data on the product serving size is used to empirically estimate the amount of *Salmonella* consumers are exposed to in a serving of chicken. The hazard characterization utilizes a peer-reviewed beta-Poisson *Salmonella* dose-response model (Teunis, 2022; Teunis, 2010; Teunis, 2008) modified to take into consideration differences in the virulence of *Salmonella* serotypes based on genomic data (**Appendix A**). This dose-response model estimates the probability of illness given a consumer exposure to a specified amount (dose) of *Salmonella* in a serving of chicken.

Finally, the risk characterization component integrates outputs of the exposure assessment and hazard characterization components to provide risk estimates of the probability of salmonellosis per serving from consumption of each type of chicken product (i.e., meat from carcasses, parts (legs, wings, breast meat), and comminuted chicken) consumed in the U.S. The model is applied to current epidemiological foodborne salmonellosis cases attributed to each type of chicken to assess the number of illnesses prevented by various risk management options (IFSAC, 2022).

The risk assessment model for chicken serves as a decision-support tool used to evaluate the public health impact of risk management options for control of *Salmonella* in chicken products. The risk assessment model parameters were adjusted to evaluate final product standard and receiving guideline scenarios.

Separate consideration was given to Risk Management Question #3 regarding process control monitoring. Process control is addressed in terms of indicator organisms and, as such, cannot be directly tied to *Salmonella* levels and serotypes. The main model outlined above is thus not appropriate for evaluation of the process control risk management options—and the public health impact of those options—in this risk assessment. Nonetheless, the same slaughter and processing paradigm can be analyzed and the connection between *Salmonella* prevalence and levels can be made (see **Chapter 7 Process Control**), making this analysis a key part of the *Salmonella* control framework under consideration in this risk assessment.

Conceptual Model

Figure 1 illustrates the conceptual model for the general approach used in this risk assessment to evaluate the public health benefits of the risk management options for controlling *Salmonella* in chicken and chicken products. The three-component model: (1) slaughter and processing, (2) growth and die off, and (3) public health outcomes, is sufficiently flexible to describe the U.S. chicken industry and targeted enough to answer the specific risk management questions.

Scenarios (that is, risk management options) for receiving guidelines, process control monitoring, and final product standards serve as inputs to the overall model and estimates for annual illnesses prevented are the outputs for each final product standard and receiving guideline scenario. The effects of the myriad of pathways contaminated product may flow through, from the end of processing through commerce and preparation, is summarized using an attenuation distribution. This attenuation distribution captures the variability associated with mixing, partitioning, growth, cooking and serving size processes between production and consumption (i.e., component (2) growth and die off).

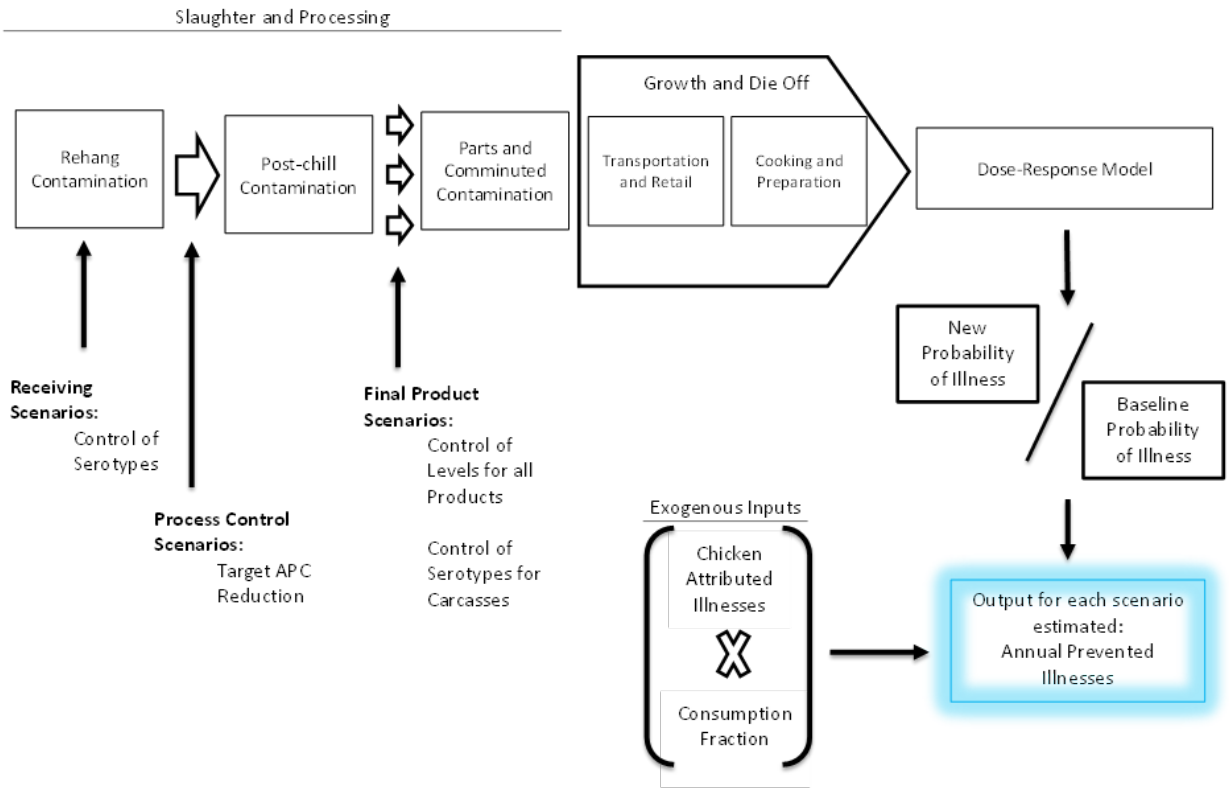


Figure 1: The conceptual risk assessment model.

Data

This risk assessment utilized a combination of FSIS pathogen testing data, consumption data, and human illness data. The FSIS data used in the risk assessment is presented by chicken product type—chicken carcasses, chicken parts, and comminuted chicken. Data from 2016 through 2021 were used for genomic analyses. Unless otherwise stated the data used for subsequent analyses of chicken carcasses are the FSIS Pathogen Reduction; Hazard Analysis and Critical Control Point (FSIS PR/HACCP) carcass *Salmonella* verification samples paired, by flock, with FSIS *Salmonella* exploratory program rehang samples.

Chicken Carcasses

- Establishment-level FSIS chicken carcass rehang samples from the 2022 exploratory program results from April 2022 through October 2022
- Establishment-level PR/HACCP chicken carcass post-chill samples from the *Salmonella* verification program results from January 2016 through October 2022
- Establishment-level *Salmonella* data from FSIS young chicken microbiological baseline study from July 2007 through September 2008

Chicken Parts

- Establishment-level FSIS PR/HACCP chicken parts samples from the *Salmonella* verification

program results from January 2016 through October 2022

- FSIS establishment-level *Salmonella* data from chicken parts microbiological baseline study from January-August 2012

Comminuted Chicken

- Establishment-level FSIS PR/HACCP comminuted chicken samples from the *Salmonella* verification program results from January 2016 through October 2022

Human illness estimates used in this risk assessment come from the CDC, including sporadic foodborne illness data from the Foodborne Diseases Active Surveillance Network (FoodNet) and foodborne illness outbreak data from the National Outbreak Reporting System (NORS)(CDC-NORS, 2021). The Interagency Food Safety Analytics Collaboration (IFSAC) foodborne illness source attribution estimates were also used. Data on the consumption of chicken in the U.S. were obtained from the National Health and Nutrition Examination Survey (NHANES).

Key Findings:

Key findings from this risk assessment are presented below. First, an overview of the descriptive data analyses conducted are presented, followed by estimates of the public health impacts of various scenarios for final product standards (Risk Management Question #2), receiving (Risk Management Question #1), and process control (Risk Management Question #3).

Data Description

Salmonella Prevalence

The amount of *Salmonella* on chicken carcasses regulated by FSIS has decreased over time. Comparison of the FSIS 2022 Exploratory Sampling post-chill data to the previous FSIS chicken carcass microbiological baseline (FSIS, 2009a) shows a 59% reduction in volume-weighted *Salmonella* prevalence (from 0.075 in 2009 down to 0.031 in 2022). In that same period, *Salmonella* prevalence at rehang has largely remained the same. **Table 1** provides current estimates for *Salmonella* prevalence statistics at rehang and post-chill, adjusted by establishment production volume.

Table 1: *Salmonella* prevalence at on chicken carcasses rehang and post-chill using FSIS 2022 Exploratory Sampling.

Sample Location	*Salmonella* Prevalence	Standard Deviation	95% confidence interval
Rehang	0.655	0.0074	(0.641, 0.670)
Post-chill	0.031	0.0030	(0.026, 0.037)

The decrease in *Salmonella* contamination at post-chill has been driven by high-volume production

establishments (slaughtering at least 10 million carcasses per year), that is, *Salmonella* carcass post-chill contamination is now predominantly in low-volume production establishments (those establishments slaughtering less than 10 million carcasses per year).
Figure 2 illustrates the inverse relationship between establishment production volume and *Salmonella* occurrence at post-chill (i.e., few large volume establishments have any *Salmonella* positives).

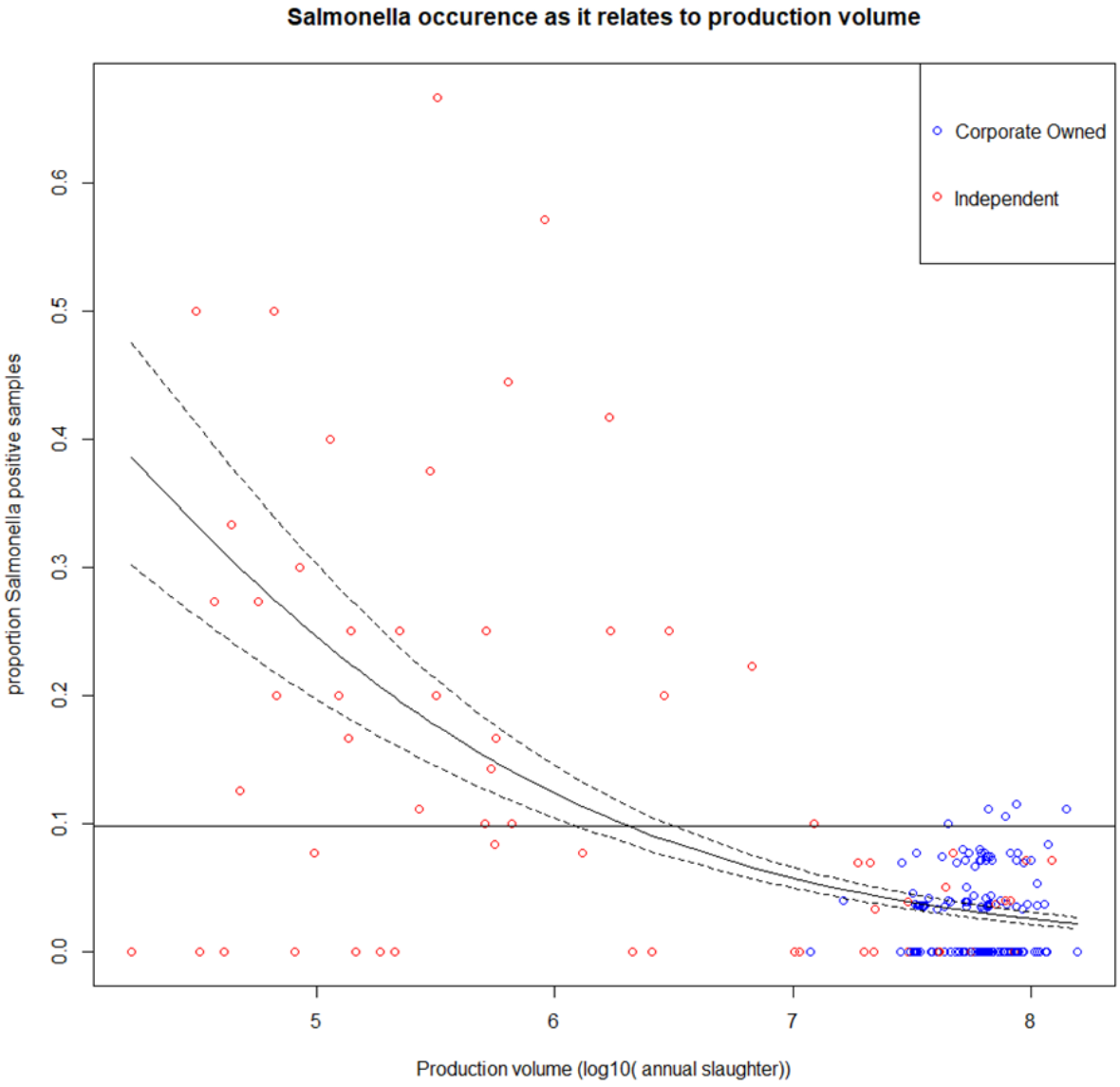


Figure 2: Relationship between establishment production volume, ownership, and presence of *Salmonella* on chicken carcasses.

Enumeration

In this section and throughout the document, findings are presented as colony forming unit per milliliter (cfu/mL) for chicken carcasses and parts, and colony forming units per gram (cfu/g) for comminuted product because FSIS sampling for carcasses and parts is conducted using a rinsate (fluid), while the comminuted chicken sampling is conducted using a direct sample of product (solid).

Data analyses indicate that current and past chicken product samples have low estimated *Salmonella* levels at post-chill (<1cfu/mL) (see **Table 2**), with 99.73% of carcasses, 99.83% of parts, and 97.12% of comminuted chicken products produced in the U.S. having a *Salmonella* level below 1cfu/mL (or cfu/g for comminuted product). It is rare for consumers to be exposed to a serving from chicken product that has at least 10 cfu of *Salmonella* per gram.

Table 2: The estimated amount of *Salmonella* positive chicken product per *Salmonella* threshold level. Further details are provided in section **3.2** . The limit of detection (LOD) of the quantitative polymerase chain reaction (qPCR) enumeration technology used by FSIS at present is 10 cfu/g or /mL (FSIS, 2022b).

	Chicken Carcasses	**Chicken Parts**	**Comminuted Chicken**
Tests *Salmonella* Negative	96.92%	93.31%	72.90%
Tests *Salmonella* Positive	3.08%	6.69%	27.10%
Tests *Salmonella* Positive and ≥1 cfu/mL or /g	9%	2%	11%
Tests *Salmonella* Positive and ≥10 cfu/mL or /g	1%	0.07%	3%
Tests *Salmonella* Positive and ≥100 cfu/mL or /g	0.10%	<0.01%	1%

Serotype

Bioinformatics were used to cluster *Salmonella* serotypes according to virulence-associated gene markers (see **Appendix A**). Available epidemiological data supported dividing chicken-associated serotypes into two clusters: a "higher virulence" cluster 1 and a "lower virulence" cluster 2 (**Table 3**). The probability of illness from consuming chicken products containing more virulent *Salmonella* serotypes ("higher virulence" serotypes) exposures is 5.66 times greater than the probability of illness from consuming chicken products containing less virulent ("lower virulence") *Salmonella* serotypes. In FSIS sampling, the average annual percentage of higher virulence serotypes among FSIS' PR/HACCP verification *Salmonella*-positive samples is approximately 26% for carcasses, 32% for comminuted product, and 35% for parts.

Table 3: Summary of the five most frequent *Salmonella* serotypes in chicken clustered based on differences in virulence. Note, all serotypes in each cluster are considered to be equally virulent for the purpose of this analysis.

Higher virulence cluster 1 serotypes	Lower virulence cluster 2 serotypes
Enteritidis	Kentucky
Typhimurium	Infantis
I 4,[5], 12:i:-	Schwarzengrund
Hadar	Heidelberg
Litchfield	Thompson

The top three higher virulence serotypes which appear most frequently in FSIS chicken samples (Enteritidis, Typhimurium, and I 4, [5],12:i:-,) are referred to as "serotypes of public health significance" in this risk assessment. The portion of FSIS PR/HACCP *Salmonella* positive samples that are sequenced as a serotype of public health significance is 24% for chicken carcasses, 33% for chicken parts, and 29% for comminuted chicken product.

Risk per Serving

Two virulence-adjusted dose-response models were developed to answer the risk management questions, one for each of the serotype clusters listed above. These dose-response models provide a description of risk of illness per serving for *Salmonella* from poultry products, when combined with an attenuation distribution. This attenuation distribution describes the variety of activities that occur between FSIS sampling a final product lot and a consumer ingesting a serving of chicken from that lot. These activities include product mixing, transportation, and cooking and can result in both *Salmonella* growth and die off.

A summary of probability of illness per serving for the main scenarios under consideration is provided in **Table 4**. The average *Salmonella* level for failing lots that test at or above each threshold level are provided, along with the average dose consumed (which is to say, the level after attenuation), and the likelihood that consumers are exposed to such servings.

The model-derived baseline probability of illness from *Salmonella* for chicken carcasses is 2 per million servings, for chicken parts is 3 per million servings, and 25 per million servings for comminuted chicken. Comparison of the threshold probability of illness to the baseline quantifies how much higher than average the risk per serving is for each scenario. For example, a serving that tests at or above 10 cfu/mL or 10 cfu/g and has a serotype of public health significance is likelier to cause illness than an average serving, with a probability of illness that is 2,000-fold higher than the average across all servings for carcass lots, 1,100-fold higher than the average serving for parts, and 590-fold higher than the average serving for comminuted chicken products.

Table 4*:* Risk of illness per serving of poultry product based on the initial threshold level of *Salmonella* in FSIS-sampled products.

Measurement	Product type	Initial threshold level (cfu/mL or g) 0.003	0.033	1	10	100
Average level for failing lots (cfu/mL or g)	Carcasses	0.30	1	10	62	433
	Parts	0.05	0.19	3	21	179
	Comminuted	17	37	163	582	2,572
Average dose consumed for average failing lot (cfu/serving)	Carcasses	0.08	0.26	3	15	108
	Parts	0.01	0.05	0.67	5	45
	Comminuted	3	6	26	92	408
Probability of illness per million servings*, higher virulence	Carcasses	230	539	2,243	6,124	15,980
	Parts	60	164	995	3,398	10,504
	Comminuted	2,353	3,632	8,000	14,859	28,487
Probability of illness per million servings*, lower virulence	Carcasses	39	95	417	1,193	3,292
	Parts	10	28	179	643	2,109
	Comminuted	438	690	1,581	3,047	6,113
Likelihood of consumer exposure to raw product at or above initial threshold level	Carcasses	11%	3%	0.27%	0.03%	0.00%
	Parts	31%	7%	0.17%	0.00%	0.00%
	Comminuted	27%	13%	3%	1%	0.17%

* Given average initial concentration multiplied by attenuation distribution

Indicator Organisms

Young chicken carcass slaughter establishments under FSIS jurisdiction have improved their process control, as measured by aerobic count (AC) reductions from rehang to post-chill, since 2008. **Figure 3** summarizes the 2022 FSIS exploratory paired carcass sampling AC levels for all samples at rehang and post-chill, where the orange vertical lines in **Figure 3** depict the average AC levels at both sampling locations. These values are 4.40 at rehang and 1.39 at post-chill, for an average log reduction of 3.01. Comparing these values to those of the 2007-2008 FSIS chicken carcass microbiological baseline study, where the average AC log reduction of 2.04 (from 4.50 at rehang to 2.46 at post-chill) was lower suggests an overall improvement in industry process control.

As depicted in Figure 3, while there have been essentially no changes in the incoming AC levels on chicken carcasses, the additional processing interventions implemented in the last 15 years are achieving about an additional 1-log reduction (i.e., on average, only 1 aerobic bacterium out of every 1000 is surviving between rehang and post-chill, as compared to 1 out of 100 aerobic bacterium in the 2007-2008 period). In a previous study (Williams, 2015) log reductions in AC, generic E. *coli* (GEC) and *Salmonella* were similar in magnitude (2.04, 2.3 and 2.08, respectively), so it is reasonable to expect reductions of roughly 3 logs in *Salmonella.*

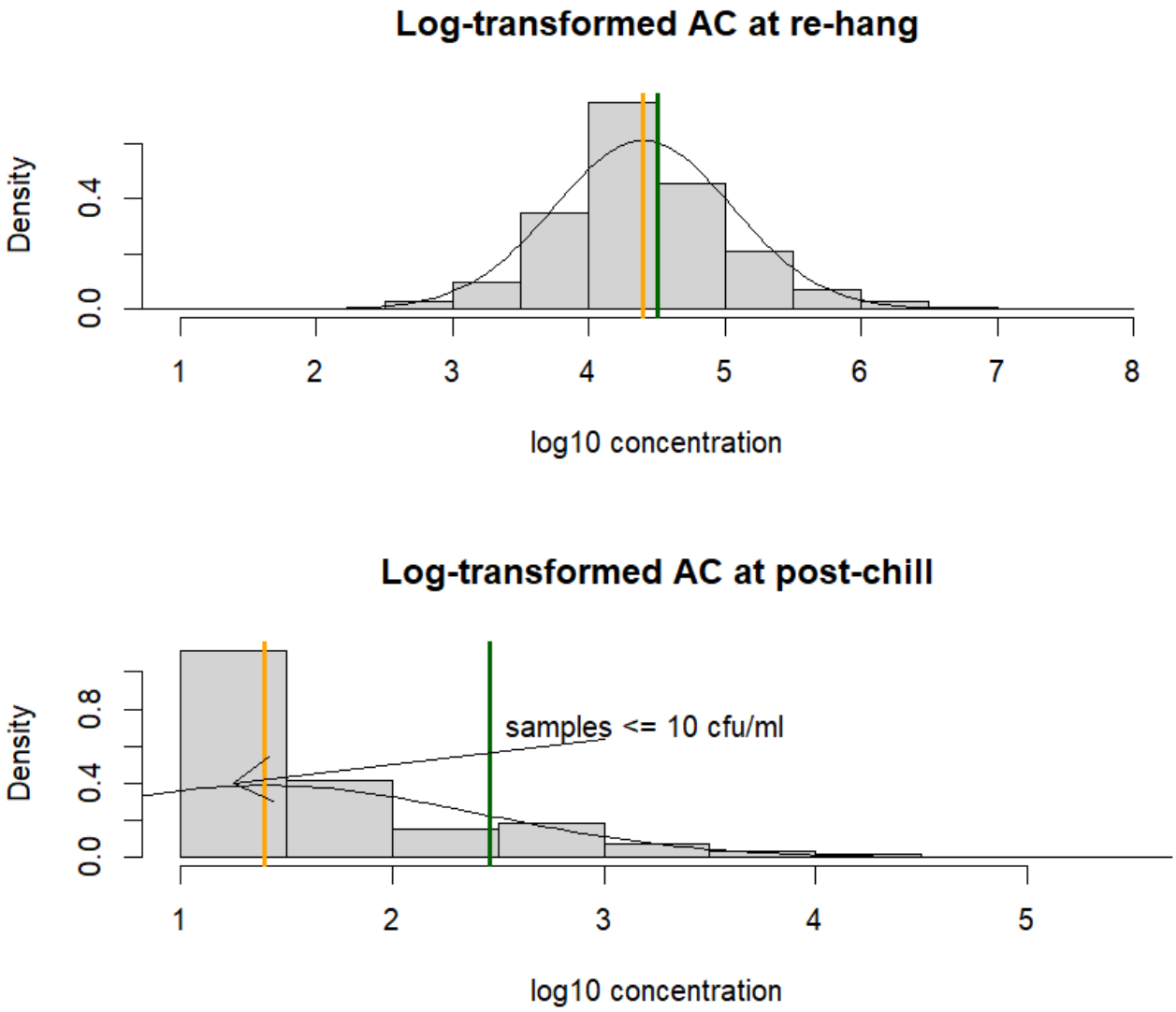


Figure 3: Distribution of AC level at both rehang and post-chill. The orange vertical lines represent the current mean level, while the green lines represent the mean level from the previous FSIS microbiological baseline study.

Results of Scenario Analysis

Final Product Standards

As there was no empirical data at the time this risk assessment was conducted on how the chicken industry will react to a final product standard, the indirect benefits of such a standard could not be assessed. The model therefore assesses the direct benefits of FSIS' implementation of such a standard. Great consideration is thus given to realistically modeling the standard, notably the continuation of status quo FSIS sampling and the determination "pass" and "fail" status of lots on the basis of enumeration and serotyping laboratory technologies with high accuracy.

An overview of the total illnesses directly prevented for the final product standard scenarios, by chicken

product, is presented in **Table 5.** Of note, it is highly likely that the regulated industry will take additional actions in response to a new final product standard, which may lead to substantial additional public health benefits.

Table 5: Summary of illnesses prevented for final product standard scenarios.

		Chicken Carcasses	Chicken Parts	Comminuted Chicken
		Illnesses Prevented (%)		
Level	**100 cfu/mL or /g**	450 (0.36%)	20 (0.02%)	600 (8%)
	10 cfu/mL or /g	1,000 (0.95%)	200 (0.2%)	1,000 (14%)
	1 cfu/mL or /g	2,450 (2%)	1,400 (1.4%)	1,400 (19%)
	Screening* Level	4,700 (3.8%)	7,850 (7.5%)	1,250 (17%)
Serotype	**Higher Virulence Serotype Diversion**	1,800 (1.7%)	NA	NA

*The FSIS *Salmonella* screen tests are sensitive down to the level of 0.03 cfu/mL for carcasses and parts, 0.003 cfu/g for comminuted product

Level-Based Final Product Standards

A chicken carcass performance standard that diverts test-positive lots based on a threshold level of 0.033 cfu/mL (i.e., 1 cfu of *Salmonella* per 300 mL poultry rinsate) is the most effective risk management option to reduce foodborne *Salmonella* from chicken carcasses, with 4,700 illnesses prevented annually, which equates to 3.8 percent of the approximately 125,000 overall chicken illnesses that occur each year. The public health impact (in terms of illnesses prevented) of the chicken carcass final product standards encompasses the illnesses estimates for all secondary chicken products, as the majority of those secondary products are fabricated from carcasses.

A chicken parts performance standard that diverts test-positive lots based on a threshold of 0.033 cfu/mL is the most effective risk management option to reduce foodborne *Salmonella* from chicken parts, with 7,850 illnesses prevented annually, which equates to 7.5 percent of the approximately 104,000 chicken parts illnesses that occur each year. The second most effective risk management option to reduce foodborne *Salmonella* from chicken parts is a level threshold of 1 cfu/mL, which prevents about 1,400 illnesses annually (a 1.4 percent reduction).

A comminuted chicken performance standard that diverts test-positive lots based on a threshold of 1cfu/g is the most effective risk management option to reduce foodborne *Salmonella* from comminuted chicken, with 1,400 illnesses prevented annually, which equates to approximately 7,500 comminuted chicken illnesses that occur annually (a 19 percent reduction).

A comminuted chicken performance standard that diverts test-positive lots based on a threshold of 0.0307 cfu/g (i.e., 1 cfu of *Salmonella* per 325 g of comminuted product) is the second most effective risk management option, with about 1,250 illnesses prevented annually, which equates to approximately 7,500 comminuted chicken illnesses annually (a 17 percent reduction), due in part to a high average risk per lot comminuted chicken.

Serotype-Based Final Product Standards

Serotype-based final product standards were developed solely for chicken carcasses due to data limitations (only 1 sampling point at post-chill) in assessing the mixture of serotypes in chicken parts and comminuted chicken product lots. In short, without at least two-points of data for a product lot, changes to the mixtures of *Salmonella* serotypes within a lot could not be estimated.

A chicken carcass performance standard that diverts lots that test positive for the higher virulence *Salmonella* serotypes (referred to as "Cluster 1") should prevent 1,800 attributable foodborne salmonellosis each year. A final product standard that targets a selection of *Salmonella* serotypes is a subset of the most effective carcass risk management options above: a level threshold of 0.0333 cfu/mL. That is, a lot must first test *Salmonella*-positive before the serotype can be identified and no serotype is without some probability of illness. Consequently, a serotype-based risk management option will not be as effective as a level-based risk management approach alone.

Achieving the public health benefits predicted for these risk management options for final product standards assumes that the *Salmonella* contamination decreases that have been achieved by the chicken industry since 2015 are maintained.

Assessing the public health impact of implementing combinations of the risk management options (Risk Management Question #4) could only be partially assessed in this risk assessment. Scenarios were conducted that assess the impact of performance standards that focus on either reducing the level of *Salmonella* on product or the serotype of *Salmonella* identified. Efforts to assess the public health impact (e.g., how many illnesses would be prevented) of a scenario where level and serotype are combined would have fewer public health benefits than each scenario (level or serotype) separately. Analytical challenges and limited data on within lot *Salmonella* variability precluded efforts to address the "or" scenarios where, for example, a lot fails if a *Salmonella* above a certain level is detected OR the sample contains a higher virulence serotype. Future research is needed to develop analytical tools to address this issue.

Uncertainty Analysis for Major Risk Management Scenarios

As is good practice, the sensitivity of the risk assessment model to inputs was analyzed. The major model inputs, including the mean of the attenuation multiplier, estimates for the *Salmonella* serotype mixture in product lots, the initial contamination distributions, and the choice of dose-response model, were systematically analyzed (see section **5.3**) and results were used to develop a robust uncertainty analysis. This uncertainty analysis was conducted for the major threshold scenarios under consideration by the risk managers. The illnesses prevented estimates with the 95 percent credible intervals are summarized in **Table 6** below.

These results suggest substantial overlap in the 95 percent credible intervals across progressively higher threshold levels, which indicates that differences in the most likely effectiveness between different threshold levels may not be meaningful. The overlap is more significant for comminuted chicken.

Table 6: Estimated annual illnesses prevented by final product concentration standards are shown for concentration thresholds of interest. Values are rounded to nearest 100 illnesses.

	Annual illnesses prevented, most likely (95% credible interval)		
Threshold level	**Chicken carcasses**	**Chicken parts**	**Comminuted chicken**
0.03 cfu/mL or /g	4600 (2000, 7100)	7900 (3300, 12700)	1500 (800, 2200)
1 cfu/mL or /g	2400 (700, 5000)	1400 (400, 3600)	1400 (600, 2100)
10 cfu/mL or /g	1000 (200, 3100)	200 (40, 700)	1000 (400, 1900)
100 cfu/mL or /g	200 (0, 1500)	20 (0, 100)	600 (200, 1500)

Receiving

As the receiving step only occurs in slaughter establishments, scenarios for this risk management question were only assessed for chicken carcasses. Rehang sampling data was used as a proxy for sampling at receiving.

Further, due to limited chicken carcass paired *Salmonella* level data at the post-chill step (only 14 of the 216 *Salmonella* screen positive post-chill samples could be enumerated for *Salmonella* in the 2022 FSIS Exploratory Sampling data), among other technical issues, this analysis does not estimate the public health impact of eliminating a portion of chicken carcasses at rehang that are contaminated with a specific level of *Salmonella*. Rather, this analysis focuses on the public health impact of eliminating specific *Salmonella* serotypes at rehang, with a focus on the same higher virulence Cluster 1 and lower virulence Cluster 2 serotypes discussed in the Final Product Standards section above.

Given the available data, two interpretations of this risk management question were considered:

1. Rehang sampling as a *verification* of *Salmonella* control strategies that were undertaken upstream (i.e., vaccination, defeathering, etc.) which result in fewer *Salmonella* positives at rehang, and

2. Rehang sampling as a *location* of potential FSIS action (i.e., diversion of product) which requires the consideration of sampling frequency and other logistics.

An overview of the total *Salmonella* illnesses prevented for both interpretations of the risk management option for chicken carcasses is presented in **Table 7**.

Table 7: Total illnesses prevented for receiving scenarios per year. A "C1 sample" is one that tests positive for a higher virulence cluster 1 (i.e., C1) serotype.

Interpretation	Risk Management Scenario	Assumption	Sub-scenario	Annual Illness Reduction (%)
Verification	Interventions taken to reduce higher virulence *Salmonella* serotypes	Reduction of higher virulence *Salmonella* serotype frequency	Reduce by 50%	22%
			Eliminate	44%
		Reduction of proportion of higher virulence *Salmonella* serotype	Reduce by 50%	7%
			Eliminate	22%
Location	Potential sampling program	Diverted flocks replaced with flocks without higher virulence *Salmonella* serotypes	**Divert Flock if:**	
			1 C1 of 1 sample	23%
			≥ 1 C1 out of 4 samples	40%
			≥ 2 C1 out of 4 samples	29%
			≥ 3 C1 out of 4 samples	17%
			4 C1 of 4 samples	7%

The results presented here reflect actions that FSIS *could* take at rehang. That said, the illness reduction estimates provided here are not realistic given current FSIS sampling capacity. This risk assessment does not presume that FSIS would realistically be able to remove all lots that test positive at rehang, given the high *Salmonella* positive rate at rehang. Therefore, these results can be used to inform potential future FSIS actions at receiving.

For the first interpretation of this risk management question—rehang sampling as a *verification* of *Salmonella* control strategies that were undertaken before—up to 55,000 (a 44% illness reduction) annual salmonellosis cases could potentially be prevented if higher virulence serotypes could be completely removed from flocks, and up to 27,000 (a 22% illness reduction) annual cases could be prevented if the proportion of higher virulence serotypes in flocks could be decreased.

For the second interpretation of this risk management question—rehang sampling as a *location* of potential FSIS action—up to 50,000 (a 40% illness reduction) annual salmonellosis cases could potentially be prevented.

Scenarios were run that consider the logistics of sampling at rehang for the second interpretation. Two sampling programs were considered: 1 sample collected per flock and 4 samples collected per flock, with 0 to 3 allowable test positive higher virulence serotype samples. With this approach, if FSIS were to increase sampling at rehang, approximately 5,000 more illnesses could potentially be prevented if flocks with higher virulence serotypes could be removed.

Process Control

Process control scenarios were assessed for the chicken carcass slaughter industry. Process control analyses

were conducted using rehang and post-chill samples. Analyses of current (2022 FSIS exploratory paired carcass sampling) and past (2007 FSIS carcass microbiological baseline) data indicates that the chicken industry is consistently achieving a large reduction in AC. Specifically, over the past 15 years, the chicken industry has achieved an overall 1-2 log10 reduction in AC. While this limits the ability to additionally reduce the overall burden of *Salmonella* illnesses from chicken by relying on changes in process control, it indicates that such improvements could be possible for segments of the industry that are not yet achieving this reduction.

Historically, indicator organisms are not strongly correlated with the presence of *Salmonella* at post-chill. Analysis of current data indicated a weak correlation between post-chill *Salmonella* prevalence and AC on two fronts:

1. AC-reduction between rehang and post-chill, and

2. the fraction of post-chill samples where no AC is observed.

The latter of these two correlations is new and did not exist when *Salmonella* rates were higher (Williams, 2015); this new correlation it will be referred to as AC-elimination.

As a result of these weak relationships between AC and *Salmonella* prevalence, it follows that the correlation between AC 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.

Given these two pieces of information, this analysis instead focused on assessing the potential of two process control performance standards to achieve the Healthy People 2030 (HP2030) illness reduction targets for *Salmonella* (HHS, 2022); or a 25% reduction in salmonellosis illnesses. These approaches are:

1. an AC-reduction standard that sets a minimum value for the difference in average log10 AC levels between rehang and post-chill,

2. an AC-elimination standard sets a minimum fraction of post-chill samples where no AC are observed with the current assay (i.e., samples with <10 cfu/mL).

As a result of the overall reductions in *Salmonella* contamination by the chicken carcass industry, the first standard described above would be effective at achieving HP2030 illness reduction targets if there was a 2.5log to 3log industry-wide AC reduction in AC between rehang and post-chill. Further, to achieve the HP2030 illness reduction targets, the AC standard must be enforceable; voluntary (and thus partial) compliance will not achieve the HP2030 target.

One advantage of second proposed approach, the AC-elimination standard, is that samples only need to be collected at one location (post-chill), which has benefits from the associated cost savings in time, materials, and laboratory resources. However, an AC-elimination standard can only be effective if FSIS requires mandatory compliance with the standard. If adoption of the standard is voluntary, the overall goal of a 25% reduction in *Salmonella*-positive samples cannot be achieved.

Conclusions

This quantitative risk assessment for *Salmonella* in chicken and chicken products incorporates *Salmonella*

enumeration and genomic data to evaluate the comparative risk of chicken products containing higher levels of *Salmonella* and more virulent serotypes. The risk assessment also provides evidence of the public health benefit of controlling *Salmonella* at various points throughout production. Available data enabled FSIS to quantify the public health benefits for controlling *Salmonella* on chicken at rehang, during processing, and the direct effect of FSIS enforcement of final products standards.

Microbial Profile of Chicken Products

- Data analyses indicate that current and past chicken product samples have low *Salmonella* levels at the final product stage (<1cfu/unit), with 99.73% of carcasses, 99.83% of parts, and 97.12% of comminuted chicken products produced in the U.S. having a *Salmonella* level below 1cfu/mL (or cfu/g for comminuted product).
    - It is rare for consumers to be exposed to a serving from chicken product that has at least 10 cfu of *Salmonella* per gram.
- The most frequent higher virulence *Salmonella* serotypes in chicken products, based on genomic and outbreak data are: Enteritidis, Typhimurium, I 4,[5], 12:i:-. Together these serotypes make up between 24%and 33% of *Salmonella* positive chicken product samples.

Final Product Standards

- Risk Per Annum: Of the annual 125,115 U.S. foodborne salmonellosis cases attributed to consuming chicken products, 83% are from chicken parts, 6% from comminuted chicken, and 11% from consumption of meat from whole carcasses.
- Risk Per Serving*:* The per serving risk posed by a poultry product containing *Salmonella* levels at or above 10 cfu per milliliter or gram and containing higher virulence *Salmonella* serotypes is much higher than risk from average servings of chicken products.
    - Servings from production lots of raw chicken carcasses that test positive for *Salmonella* at levels of 10 cfu/mL or greater with a serotype of public health significance are 2,000-fold more likely to cause illness than the average across all chicken servings.
    - Servings from production lots of raw chicken parts that test positive for *Salmonella* at levels at or above 10 cfu/mL with a serotype of public health significance are 1,100-fold more likely to cause illness than the average across all chicken parts servings.
    - Servings from production lots of comminuted chicken that test positive for *Salmonella* at or above 10 cfu/g with a serotype of public health significance are 590-fold more likely to cause illness than average across all comminuted chicken servings.
- Public Health Impact: Between 1,000 and 2,200 illnesses per year would be prevented as a direct result of not allowing chicken product exceeding 10 cfu/mL or /g into commerce based on FSIS' verification sampling program.
    - 1,000 illnesses would be prevented by a carcass standard, 200 illnesses would be prevented by a parts standard, and 1,000 illnesses would be prevented by a comminuted product standard.

- As carcasses are the farthest upstream product, the effect of a carcass final product standard will reduce illnesses from secondary products.

- Additional benefits from parts and comminuted product standards are likely, but the data to assess the interaction of all three standards are not available.

- Estimating Public Health Benefits*:* The public health benefits estimated from establishing enforceable final product standards is limited to the direct benefits of FSIS' current level of verification sampling and testing of chicken products (e.g., 5 samples per month).

- Additional Benefits: Public health benefits are anticipated from industry response to enforceable final product standards, but information on industry behavior is needed to quantify these potential benefits.

Preharvest Control

- Maximum Benefit: Maximum public health benefits from a verification program could be achieved by eliminating higher virulence *Salmonella* serotypes in flocks, which could prevent up to 55,000 human illness each year (44% illness reduction). Reducing the proportion of higher virulence serotypes in each flock by half could prevent up to 27,000 human illnesses each year (22% illnesses reduction).

- Sampling Program: A sampling program at rehang could reduce illnesses by 5,000 to 50,000 (7-40%) annually. The flock diversion count would be equally significant.

- Proof-of-Concept: There is potential benefits from this approach, but the lack of data describing the effectiveness of preharvest interventions is a major source of model uncertainty. These results should serve as proof-of-concept approach to illustrate the strengths and potential weaknesses of a flock diversion risk management option.

Process Control

- Over the last 15 years, the chicken carcass industry has achieved an additional 1- 2log reduction in AC from rehang to post-chill. This has corresponded to a 59% reduction in volume-adjusted *Salmonella* prevalence at post-chill.

- While the correlation between AC-level reduction and post-chill *Salmonella* prevalence is weak, it is a useful metric of process control.

    - The remaining room for improvement is through an enforceable, rather than voluntary, 2.5log to 3log industry-wide AC reduction standard.

- Lower volume production establishments (slaughtering less than 10 million carcass per year) are underperforming the average industry wide reduction and have potential to improve.

- AC elimination at post-chill is now correlated to post-chill *Salmonella*. Such a removal of all organisms on final products is akin to pasteurization.

Cross-Cutting Issues

- Combination of Scenarios: It was not possible to model the effects of multiple risk management options in sequence (Risk Management Question #4).

    - Combining final product standards with process control scenarios was not possible because the correlation between indicator organisms and *Salmonella* does not extend down to the resolution of *Salmonella* levels and serotypes.

    - Combining receiving guidelines with the other scenarios under consideration was not supportable given the receiving data gaps that have been outlined.

    - However, it is reasonable to expect that industry risk-managers respond to a final product standard by taking actions to improve process control or reduce live-bird contamination, thus reducing the likelihood that their products would fail such standards. These actions, while not modeled in this risk assessment given the paucity of data, would likely greatly enhance the number of illnesses reduced.

- Data Sharing: Sharing data is useful for enhancing the characterization of *Salmonella* levels and benefits of controlling this pathogen in final poultry products. FSIS supports continued efforts to work collaboratively with our industry and other stakeholders to share data. Industry data has the potential to enhance the findings presented here and reduce uncertainty in the scenarios modeled in this risk assessment.

Chapter 1 Introduction

The United States Department of Agriculture (USDA) Food Safety and Inspection Service (FSIS) protects the public's health by ensuring that meat, poultry, and egg products are safe, wholesome, and properly labeled. The Agency is committed to reducing foodborne infections associated with FSIS-regulated products, in particular *Salmonella* illnesses attributable to poultry. *Salmonella* in poultry products remains a significant food safety concern in the United States (U.S.). The U.S. Centers for Disease Control and Prevention (CDC) estimates *Salmonella* infection is responsible for over 1 million illnesses, 26,500 hospitalizations, and 420 deaths in the U.S. every year (Scallan, 2011).Of these illnesses, an estimated 66 percent are from food (Beshearse, 2021), with over 17 percent from eating chicken (IFSAC, 2022). All salmonellosis cases impose an estimated $3.7 billion in economic burden in a typical year. Almost 90 percent of this burden, $3.3 billion, is due to deaths; 8 percent, $294 million, is due to hospitalization; and the remaining 2 percent is due to non-hospitalized cases (Hoffmann, 2021).

On October 17, 2021, FSIS announced that it is mobilizing a stronger, and more comprehensive effort to reduce *Salmonella* illnesses associated with poultry products by establishing an enhanced food safety framework to address *Salmonella* contamination on poultry (FSIS, 2022c). Central to this effort is a quantitative microbial risk assessment (QMRA) that provides information addressing risk management questions on the predicted public health impact of controlling the prevalence, levels, and/or specific serotypes of *Salmonella* on chicken presented for slaughter, on chicken carcasses throughout the slaughter process, and/or in final chicken products.

This risk assessment addresses the following risk management questions:

1. What is the public health impact (change in illnesses, hospitalizations, and deaths) achieved by eliminating at receiving a proportion of chicken contaminated with specific levels of *Salmonella* and/or specific *Salmonella* subtypes?

2. What is the public health impact (change in illnesses, hospitalizations, and deaths) achieved by eliminating final product contaminated with specific levels of *Salmonella* and/or specific *Salmonella* subtypes?

3. What is the public health impact of monitoring/enforcing process control from rehang to post-chill? Monitoring could include analytes such as *Enterobacteriaceae*, Aerobic Plate Count, or other indicator organisms, analysis could include presence/absence or levels and the monitoring could also include variability of actual result versus expected result, log reduction, absolute sample result, or other individual establishment specific criteria.

4. What is the public health impact of implementing combinations of the risk management options listed above?

1.1 Policy Context

Pathogen reduction performance standards have been applied to meat and poultry slaughter establishments since the inception of the Pathogen Reduction and Hazard analysis And Critical Control Point (PR/HACCP) rule (FSIS, 1996b). Under performance standards—which are two-class attribute sampling plans—each establishment is subjected to a series of sampling events. FSIS uses these *Salmonella* sampling results to assess establishment performance during a reference period of one

completed 52-week moving window based on a 3-category system. Establishments at or below half of the performance standard over the previous moving window are placed in Category 1, those that meet the standard in that period are placed in Category 2, and those that fail the standard in the previous moving window are placed in Category 3. FSIS posts on its website the category status of individual establishments for pathogen reduction performance standards for *Salmonella* in young chicken carcasses, young turkey carcasses, raw chicken parts, and not-ready-to-eat (NRTE) comminuted chicken and turkey products, based on FSIS verification sampling results. Public dissemination of establishment categorization has been shown to serve as a market-based incentive to encourage establishments to reduce *Salmonella* contamination in failing establishments (Ollinger, 2020).

Analyses of the effectiveness of current FSIS *Salmonella* performance standards indicates there has been an overall reduction in the occurrence of *Salmonella* on meat and poultry products (Williams, 2020).

1.2 Purpose and Scope

The overall purpose of this risk assessment is to assess—at different points in the chicken slaughter process—the public health benefit of various risk management options.

These risk management options were modeled to evaluate their potential effect on the entire U.S. chicken industry, rather than any single slaughter and processing establishment or subset of establishments. Whenever possible and appropriate, these options were evaluated for three chicken products: carcasses, parts, and comminuted chicken. **Table 8** summarizes a basic description of these product categories and some features that are important to this risk assessment; further details are provided in Appendices B and C.

Each risk management question addresses different areas of the chicken slaughter process: receiving, process control, and final product. It was not possible to model the public health benefit for each chicken product in each of these areas. Specifically, the first risk management question only applies to carcasses because receiving is the initial step of the slaughter process. Assessment of process control—Risk Management Question 3—requires two points of sampling to adequately model changes to the industry. As such, process control is only assessed for carcasses, and is not directly examined for parts and comminuted product. Similarly, as will be discussed in detail, assessment of the public health impact of elimination of product with specific *Salmonella* subtypes is only possible with data from two points in the slaughter process to adequately model changes to the industry. As such, the impact of serotype is only assessed for carcasses, but not for parts or comminuted products.

Table 8: Description of chicken product assumptions used in the risk assessment model. These model assumptions are based the current FSIS sampling frequency, product definitions, and laboratory methodology (FSIS, 2021; FSIS, 2022b).

Chicken Product	Definition	Lot Size	FSIS Sampling Unit
Carcass	Young chicken carcasses	1 flock (~46,000 birds)	Rinsate (30 mL)
Parts	Legs, Breasts, Wings	1 day of production	Rinsate (30 mL)
Comminuted	Ground Chicken	1 day of production	Ground (325 g)

Analysis of all risk management scenarios in this risk assessment is predicated on the assumption that the industry will maintain the overall pathogen reductions that have been achieved in the past.

1.3 Conceptual Model

Figure 4 illustrates the conceptual model used to determine the public health benefits of the risk management options above. The three-component model: (1) slaughter and processing, (2) growth and die off, and (3) public health outcomes, is sufficiently flexible to describe the U.S. chicken industry, and targeted enough to answer the specific risk management questions.

Scenarios (that is, risk management options) for receiving guidelines, process control monitoring, and final product standards serve as inputs to the overall model, and the estimates for annual illnesses prevented are the outputs for each scenario, whenever appropriate. For process control standards, illnesses prevented could not be estimated due to the weak correlation between indicator organisms and pathogens.

To provide support for this risk assessment, FSIS entered into a Cooperative Agreement with the University of Maryland's Joint Institute for Food Safety and Applied Nutrition (UMD-JIFSAN), in partnership with EpiX Analytics. With UMD-JIFSAN, FSIS gained a partner to assist in facilitating voluntary industry data-sharing in a confidential and secure manner. EpiX Analytics developed two dose-response models that describe the probability of illnesses given ingestion of a *Salmonella* level dose: one model for higher virulence serotypes and one for lower virulence serotypes. EpiX Analytics' approach uses whole genome sequencing (WGS) data to cluster serotypes (i.e., serocluster) based on virulence gene markers and then estimates the difference in infectivity between these two seroclusters based on epidemiological data for foodborne illnesses associated with *Salmonella* from CDC National Outbreak Reporting System (NORS) and CDC Foodborne Diseases Active Surveillance Network (FoodNet). The EpiX Analytics dose-response models are based on earlier work (Teunis, 2022; Teunis, 2010).

Confidence in Effect

To ensure confidence in the results of this risk assessment—and for these results to serve as scientific support for regulatory rulemaking—it is imperative that the public health impact of any change in FSIS' current approach can be attributed specifically to the adoption of a given risk management option, rather than by chance alone.

FSIS data indicates that all chicken flocks enter the processing environment with some *Salmonella* contamination (see section **3.4**), and analysis of FSIS two-point chicken carcass sampling indicates that when *Salmonella* is detected in a flock, multiple serotypes are present (section **3.5**). Therefore, it is follows that there is no public health benefit from diverting an average chicken lot from commerce. As a consequence, a key feature of this risk assessment is the determination of which of these scenarios are diverting product lots with higher-than-average risk.

1.4 Report Organization

This report begins with a description of the current status of *Salmonella* contamination across the chicken industry. We begin with a description of the public health context of *Salmonella* and its serotypes and introduce a novel virulence-gene informed clustering of these serotypes (**Chapter 2**); full details of the *Salmonella* serotype clustering are summarized in **Appendix A**. In the *Salmonella* Microbial Profile (**Chapter 3**), for each of the three product categories analyzed (carcasses, parts, and comminuted), *Salmonella* prevalence and level distributions are provided. The full details of the

methods and data used in these descriptions are in Appendices B and C. The key data features of regulatory importance are then described, including how production volume corresponds to contamination. In **Chapter 4**, salmonellosis surveillance data, chicken consumption estimates, and an analysis of the baseline probability of illness are discussed. The chapter ends with a description of the risk per serving for the main threshold levels under consideration.

The remainder of the document is dedicated to estimating the public health benefit of the various risk management scenarios as summarized in the risk management questions. These scenarios fall into three categories: final product standards (**Chapter 5**), receiving guidelines (**Chapter 6**), and process control monitoring (**Chapter 7**). Each chapter starts with a description of scenarios that address the risk management question under consideration and a summary of key modeling assumptions, followed by a description of the modeling method used and, finally, presentation and discussion of results.

Chapter 5, final product standards, contains the majority of the risk characterization, while the hazard characterization (i.e., dose-response model) is summarized in **Chapter 3**, serotype clustering, and **Chapter 4**, risk per serving. The same risk characterization approach is used to assess the public health benefit of final product standards and receiving guidelines (**Chapter 6**). In **Chapter 7**, process control monitoring is evaluated and has a separate risk characterization. A summary of overall results and recommendations for data needs based on limitations identified in the development of this risk assessment is included in **Chapter 8**.

Theory and data analysis details are provided in the Appendices. Full descriptions of data used in this risk assessment and analysis of quantitative PCR methods under current *Salmonella* conditions are available in **Appendix B**, and data used in the dose-response development is described in **Appendix A**[3]. A separate description of the bioinformatics that were used for the dose-response development has been provided to ensure the transparency requirements of the Information Quality Act (Section 515 of Public Law 106-554) in these *Bioinformatics Supplemental Materials* (available here). Details of the population description methodology are available in **Appendix C**.

1.5 Model Approach

The first goal of this risk assessment is to define a probabilistic model that explains the current state of pathogen contamination in the U.S. chicken carcass population at the rehang and post-chill locations of the slaughter process. For all risk management scenarios, FSIS considered the guidance it received from the National Advisory Committee on Microbiological Criteria for Foods (NACMCF) related *to Enhancing Salmonella Control in Poultry Products* referred to as the NACMCF 2023 response hereafter(NACMCF, 2023).

For the receiving scenarios assessed within the 'slaughter and processing' component of the model, FSIS utilized Agency rehang sampling data. This data source was the best available at the time this risk assessment was conducted, as FSIS does not collect data on live birds as the Agency does not have

[3] To provide support for this risk assessment, FSIS entered into a Cooperative Agreement (FSIS-02152022) with the University of Maryland's Joint Institute for Food Safety and Applied Nutrition (UMD-JIFSAN), in partnership with EpiX Analytics. With UMD-JIFSAN, FSIS gained a partner to help obtain industry data in a confidential and secure manner. EpiX Analytics provided expertise in dose-response modeling. As a part of this Agreement, EpiX Analytics used genomics to classify serovars into groups (clusters) based on virulence similarities and developed dose-response models for the serovar clusters. Details of the EpiX Analytics methodology are provided below in the report entitled "*Using genomics to identify nontyphoidal Salmonella serovars of concern and estimating dose-response models amenable to risk assessments in poultry."*

regulatory jurisdiction in the poultry preharvest environment. Further, at this time, the scientific community has not established a standard sampling method for *Salmonella* on live birds. Efforts are ongoing to enhance FSIS data with industry-supplied data through FSIS' Cooperative Agreement, as described previously.

For the process control scenarios assessed within the 'slaughter and processing' component of the model, the impact of a reduction in indicator organisms between rehang and post-chill on end-point *Salmonella* prevalence was assessed. Given the weak correlation between indicator organisms and pathogen prevalence (Williams, 2015), and that the dose-response model constructed for this risk assessment is *Salmonella* serotype informed and level-based, the public health impact from the process control scenarios are assessed separately.

For the final product scenarios assessed within the 'slaughter and processing' component of the model, a probabilistic model for *Salmonella* contamination at post-chill describes the potential human exposure from servings derived from whole carcasses. Data collected at the end of the other two major production processes, specifically where carcasses are fabricated into parts and at the end of the grinding process, are used to describe potential consumer exposure to *Salmonella* from chicken parts and comminuted chicken. The effect of various risk management options can then be assessed by adjusting the parameters of these population models in accordance with the anticipated effect of different risk management options.

We summarize the effects of the myriad of pathways contaminated product may follow from the end of processing through commerce and preparation using an attenuation distribution. This attenuation distribution captures the variability associated with mixing, partitioning, growth, cooking and serving size processes between production and consumption. A lognormal attenuation distribution ($\mu = -5.00\ log10, \sigma = 1.91\ log10$) was calibrated previously for chicken using a single distribution for *Salmonella* contamination, a general *Salmonella* dose-response function from WHO-FAO (FAO/WHO, 2000), and a prior estimate for total *Salmonella* illnesses attributed to chicken (Ebel, 2015). The log10 mean of this distribution (-5log10) is consistent with a scenario where a raw chicken product is properly handled to avoid growth of *Salmonella*, then subjected to cooking to achieve a minimum internal temperature of 165 ºF, which the Agency recommends as the final cooking temperature and has determined will deliver at least a 7log10 reduction of *Salmonella (FSIS, 2017)*, and consumed in a serving size of 100 g (adding 2log10 to the 7log 10 reduction). Lacking alternative estimates, this default attenuation distribution is used across analyses of carcasses, parts, comminuted chicken and for *Salmonella* serotypes. Similarly, this attenuation distribution is used in the derivation of the dose-response models for different virulence-based *Salmonella* serotype clusters (**Appendix A**). Nevertheless, FSIS also conducted a sensitivity analysis to explore the effects of alternative attenuation distribution parameters on the estimated effectiveness of risk management options.

The results of this risk assessment are likely different from other recently published and forthcoming poultry risk assessments (Lambertini, 2019; Oscar, 2021). Prior risk assessments are more akin to attribution studies that calculate the effect of removing all product that has a specific risk characteristic (e.g., specific serotypes or levels above a specified threshold value). The removal of all servings with the specified risk characteristic would require the testing of all servings, so the previously published estimates are seen as aspirational upper bounds for the number of illnesses that could be prevented. Much of the focus of this study will focus on how different risk management options would aid FSIS' ability to correctly identify product with a specified risk characteristic, and then ensure that affected product is either rendered safer for human consumption or removed from commerce. Thus, the goal of

the risk assessment can be restated as: determining what fraction of all illnesses associated with product having the risk characteristic will be prevented because of actions taken in response to data collected by FSIS.

To estimate the direct impacts of a new FSIS regulation, this risk assessment pays particular attention to modeling FSIS' sampling process and testing methods in a manner consistent with how the regulation is to be implemented in practice. This will include consideration of FSIS' product frequency, sample unit size (e.g., pounds), testing methods (*Salmonella* detection or level), and measurement uncertainty.

1.6 Introductory Tables and Figures

Given the length of the document and the complexity of the models developed, some introductory summary tables and figures are provided to aid the reader. As previously discussed, **Figure 4** is a schematic representation of the risk assessment model. **Table 9** outlines the required information and assumptions used in each of the three scenario analyses. **Table 10** summarizes the interpretation of the risk management questions and which scenarios were successfully implemented. **Table 11** contains the model parameters and variables used to evaluate the final product standard options.

Figure 4: The conceptual risk assessment model.

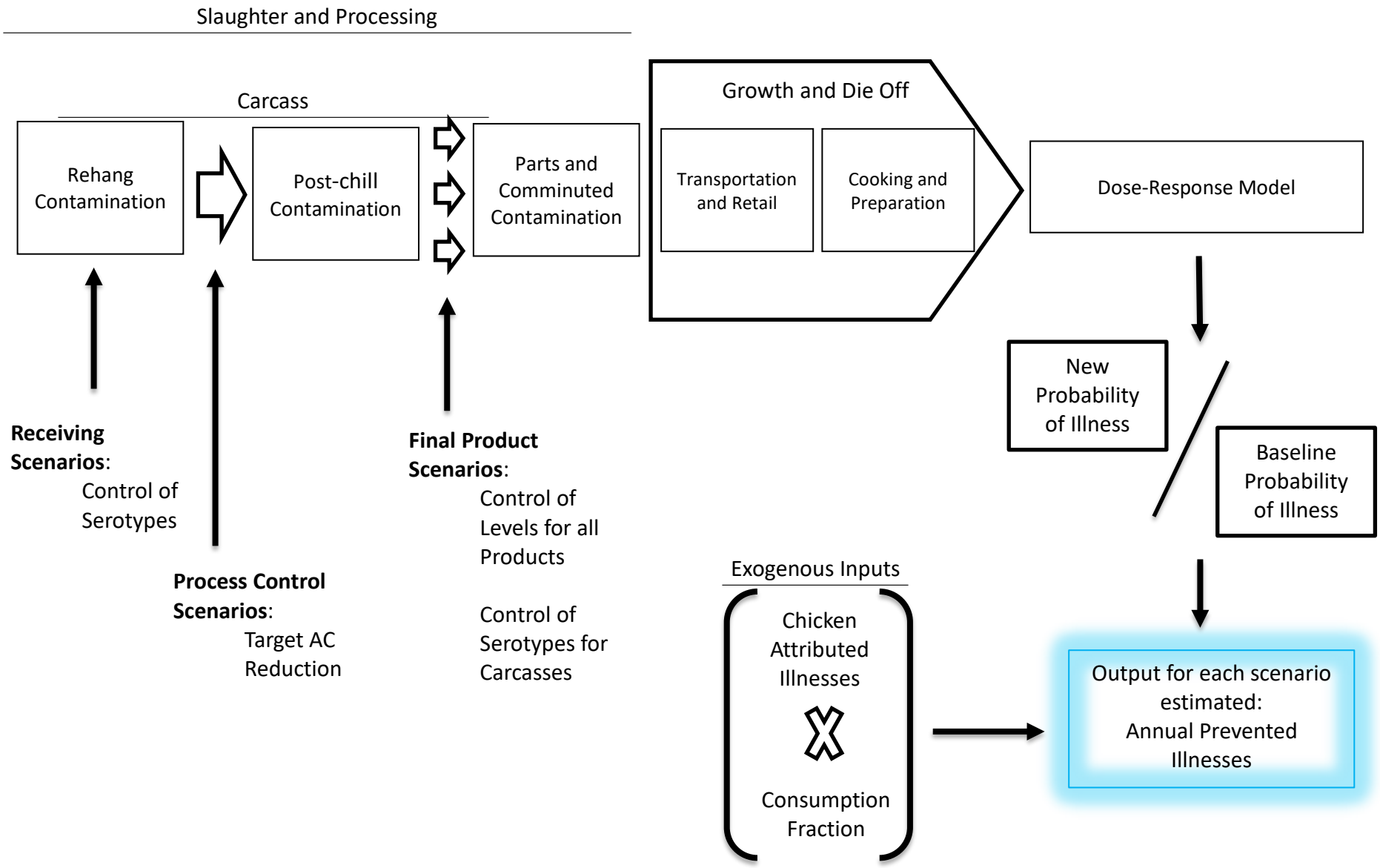


Table 9: Risk assessment information and assumptions.

Information Needed	Assumption Used	Supporting Data or Information
***Salmonella* in Poultry Baseline Profile**		
Salmonella Microbial Profile	FSIS data are representative of chicken slaughter and processing establishments under FSIS jurisdiction	FSIS establishment level PR/HACCP, Exploratory Sampling, and Microbial Baseline data
	All flocks contain some *Salmonella*. This risk assessment does not assume that all birds have *Salmonella*.	Analysis of FSIS 2022 Exploratory Data with two samples per flock – see section **3.5** . (Cox, 2020; Obe, 2023; Rasamsetti, 2023; Thompson, 2018)
	Multiple *Salmonella* strains are present in flocks.	
	Salmonella serotypes can be clustered into two groups: higher virulence serotypes (labeled C1) and lower virulence serotypes (labeled C2).	See **Chapter 2** for the FSIS summary of the EpiX Analytics report in **Appendix A**. For transparency, FSIS has developed a separate description to the EpiX Analytics serotype clustering that clarifies the bioinformatics approach taken by EpiX Analytics (available here).
	Epidemiological data can be used to categorize overall serocluster virulence: higher virulence serocluster (labeled C1) and lower virulence serocluster (labeled C2).	See **Appendix A** and attached **Bioinformatics Supplemental Materials** (available here).
	Flocks contain a dominant serocluster.	Contingency table analysis of FSIS 2022 Exploratory Data in section **5.1** subsection **Describing Serotype Mixtures in Flocks**.
Production Volume Data		FSIS establishment-level production volume data.
Estimated number of human *Salmonella* illnesses attributable to chicken consumption	125,115 chicken-associated *Salmonella* illnesses per year.	This value is calculated as the product of the total number of CDC FoodNet cases per year (7,600), the share of these cases that are foodborne (66 percent) and of domestic origin (89 percent), the under-diagnosis multiplier for *Salmonella* (24.3)(Ebel, 2012c), dividing by the FoodNet catchment area (15 percent), and multiplying by the portion IFSAC attribution to chicken (17.3 percent based on data through 2020).
Final Product Standards		

Determination of lot concentration	An accurate method will be used to determine the threshold status of tested lots.	This information need was highlighted in the NACMCF 2023 final report (NACMCF, 2023) and methods for evaluating enumeration data for accuracy are presented in **Appendix B Data and Data Analysis**.
FSIS adulterant testing procedure and frequency	FSIS sampling and lab methodology for carcasses, parts, and comminuted continue as currently utilized.	FSIS Microbiological Lab Guidebook (FSIS, 2022b) and Sampling Instructions (FSIS, 2021)
	FSIS will continue current sampling frequency	Small establishments: 24 samples/year, Large establishments: 60 samples/year
	Lot size will remain as currently used by industry in PR/HACCP documentation	Carcasses: 1 flock Parts: 1 day of production Comminuted : 1 day of production
Fate of diverted product	Consumer demand for raw chicken products will be met by the industry, so every removed lot will be replaced by another lot in the aggregate.	This assumption is considered reasonable because of the high consumer demand for prepared chicken products, see introduction to **Chapter 5** for more detail.
	There is no public health benefit from diverting random lots.	The removed lot and the lot replacing it will both be of the same risk on average.
Impact of individual product standards	Carcass standards have impact on *Salmonella* illnesses from all chicken product	The more than 9 billion young chicken carcasses slaughtered annually are the primary source material for parts and comminuted products.
	Illness attributed to parts and comminuted products are assumed to be proportional to product availability.	83% of chicken is sold as parts and 6% as comminuted product.
Salmonella growth and die-off after slaughter and processing.	The effect of mixing, transportation, storage, cross contamination, cooking and handling is described by an attenuation distribution.	A lognormal attenuation distribution ($\mu = -5.00\, log\, 10\,, \sigma = 1.91\, log\, 10$) was calibrated previously for chicken using a single distribution for *Salmonella* contamination, a general *Salmonella* dose-response function from WHO-FAO and a prior estimate for total *Salmonella* illnesses attributed to chicken (Ebel, 2015).

	This effect is described by the same attenuation distribution for all three products.	
Consumption of higher doses of *Salmonella* is associated with a higher probability of illness.	Dose-response relationship used in QMRA.	Teunis' dose-response model using outbreak data is the primary underpinning of this theory as applied in this document (Teunis, 2022; Teunis, 2010; Teunis, 2008).
Receiving Guidelines		
Contamination at Flock Receiving	Rehang data is descriptive of incoming contamination.	The best available data source of near incoming contamination is the FSIS 2022 exploratory rehang sampling.
Process Control		
Utility of indicator organisms to monitor process control		The weak correlation between post-chill *Salmonella* prevalence and AC reductions from rehang to post-chill are analyzed.
Illness Reduction and Prevalence	A reduction in *Salmonella* prevalence results in a proportional reduction in salmonellosis.	(FSIS, 2015)

Table 10: Interpretation of risk management questions and table of scenarios.

Risk Management Question	Scenario Description	Product	Range of Scenarios	Public Health Metric
What is the public health impact (change in illnesses, hospitalizations, and deaths) achieved by eliminating at receiving a proportion of chicken contaminated with specific levels of Salmonella and/or specific Salmonella subtypes?	Flocks contaminated with a higher virulence serotype never arrive at rehang, i.e., FSIS verifies preharvest control at rehang testing.	Chicken Carcasses	Two scenarios were modeled: reducing the overall prevalence of higher virulence serotypes and reducing the mix of higher to lower virulence *Salmonella* in flocks.	Annual illnesses prevented. Resolution does not extend to death and hospitalization estimates, but these are among the annual illnesses.
	Flocks contaminated with a higher virulence serotype are diverted on the basis of rehang testing	Chicken Carcasses	Two sampling programs were considered: 1 sample collected per flock and 4 samples collected per flock, with 0 to 3 allowable test positive higher virulence serotype samples.	Annual illnesses prevented. Resolution does not extend to death and hospitalization estimates, but these are among the annual illnesses.
What is the public health impact of monitoring/enforcing process control from rehang to post-chill? Monitoring could include analytes such as Enterobacteriaceae, Aerobic Plate Count, or other indicator organisms, analysis could include presence/absence or levels and the monitoring could also include variability of actual result versus expected result, log reduction, absolute sample result, or other individual establishment specific criteria.	Compliance with a target log reduction in AC from rehang to post-chill.	Chicken Carcasses	Mandatory: 2.9log10 reduction of AC Voluntary: 3.3log10 reduction of AC	The HP2030 target of a 25% salmonellosis reduction.
	Compliance with a target AC elimination from rehang to post-chill	Chicken Carcasses	Mandatory: At least 35% of post-chill samples have no AC Voluntary: the target cannot be achieved.	The HP2030 target of a 25% salmonellosis reduction.
What is the public health impact (change in illnesses, hospitalizations,	Level-based Threshold Standards:	Chicken Carcasses, Parts,	Standards at concentrations from	Annual illnesses prevented.

and deaths) achieved by eliminating final product contaminated with specific levels of Salmonella and/or specific Salmonella subtypes?	Lots are diverted if a regulatory sample tests above a predetermined level (cfu/g or cfu/mL). The terms level and concentration are used interchangeably throughout the document.	and Comminuted Product	1cfu/2,600g to 100cfu/g were simulated for chicken producst.	Resolution does not extend to death and hospitalization estimates, but these are among the annual illnesses.
	Serotype Standard: Lots are diverted if a regulatory sample tests positive for a serotype of higher virulence	Chicken Carcasses	*Serotype standards <u>could only</u> be modeled for chicken carcasses.* Serotyping at multiple points in the slaughter process is necessary to ensure reliable serotype distribution within a lot – these scenarios cannot be modeled when only the final product is tested.	Annual illnesses prevented. Resolution does not extend to death and hospitalization estimates, but these are among the annual illnesses.
What is the public health impact of implementing combinations of the risk management options listed above?	No combinations of scenarios could be implemented.	None	None	None

Table 11: Table of model parameters and variables for final product standard and receiving guidelines estimates.

Description	Parameter/ variable	Value/model	Units
log10 mean final product concentration	μ	carcasses= -4.51187 parts= -3.017 comminuted= -3.70	cfu/mL cfu/mL cfu/g
log10 std dev final product concentration	σ_x	carcasses= 1.623983 parts= 1.027 comminuted= 1.949	cfu/mL cfu/mL cfu/g
conversion of rinse sample conc per mL to conc per gram	g	$\log10\left(\frac{400}{0.14}\right)+\log10\left(\frac{1}{4\times454}\right)=0.1968$	mL/g
log10 mean initial final production concentration in grams	μ_x	carcasses and parts = $\mu + g$ comminuted = μ	cfu/g
log10 mean of attenuation distribution	μ_a	-5	g/serving
log10 std dev of attenuation distribution	σ_a	1.91	g/serving
concentration threshold	T	policy input	cfu/mL or cful/g
fraction of lots that pass	ω	$\Phi\left(\frac{T-\mu}{\sigma_x}\right)$	proportion
dose distribution at consumption	d	$10^{\text{Normal}\left(\mu_x+\mu_a,\sqrt{\sigma_x^2+\sigma_a^2}\right)}$	cfu/serving
dose-response functions for clusters 1 and 2	R_1, R_2	polynomials	proportion
prob of illness from cluster 1 or 2	$P(ill\|C1/C2)$	$\int R_{1/2}(d)f(d)\partial d$	proportion per serving
dose distribution given that lot passes	$d_{x\le T}$	$10^{TruncNormal(T,\mu,\sigma_x)}\times10^{Normal(\mu_a+g,\sigma_a)}$	cfu/serving
proportion of *Salmonella* from cluster 1	c	0.2	proportion

prob of illness among passing lots from C1 *Salmonella*	$P(ill\|pass,C1)$	$\int \mathrm{R}_1(\mathrm{d}_{x\le T})h(d_{x\le T})\partial \mathrm{d}_{x\le T}$	proportion per serving
prob of illness among passing lots from C2 *Salmonella*	$P(ill\|pass,C2)$	$\int \mathrm{R}_2(\mathrm{d}_{x\le T})h(d_{x\le T})\partial \mathrm{d}_{x\le T}$	proportion per serving
prob of illness among all passing lots	$P(ill\|pass)$	$\mathrm{c}\times P(ill\|pass,C1)+(1-c)\times P(ill\|pass,C2)$	proportion per serving
prob of illness among failing lots from C1 *Salmonella*	$P(ill\|fail,C1)$	$\frac{\left[P(ill\|C1)-\omega\times P(ill\|pass,C1)\right]}{1-\omega}$	proportion per serving
prob of illness among failing lots from C2 *Salmonella*	$P(ill\|fail,C2)$	$\frac{\left[P(ill\|C2)-\omega\times P(ill\|pass,C2)\right]}{1-\omega}$	proportion per serving
prob of illness among all failing lots	$P(ill\|fail)$	$\mathrm{c}\times P(ill\|fail,C1)+(1-c)\times P(ill\|fail,C2)$	proportion per serving
baseline prob of illness across all servings	$P_{baseline}(ill)$	$\omega\times P(ill\|pass)+(1-\omega)\times P(ill\|fail) = \mathrm{c}\times \mathrm{P}(\mathrm{ill}\|\mathrm{C1})+(1-\mathrm{c})\times \mathrm{P}(\mathrm{ill}\|\mathrm{C2})$	proportion per serving
total lots produced per year	L	carcasses=203,718 parts=104,505 comminuted=8545	lots
total lots tested per year	n	carcasses=9635 parts=14,192 comminuted=1991	lots
share of failing lots that are diverted	α	$\frac{\mathrm{n}}{L}$	proportion
number of illnesses before policy	λ_{ill}	carcasses=125,115 parts=103,845 comminuted=7507	illnesses/year
new prob of illness after policy	$P_{new}(ill)$	$\omega\times \mathrm{P}(\mathrm{ill}\|\mathrm{pass})+(1-\omega)\times\left[\alpha\times \mathrm{P}_{\mathrm{baseline}}(\mathrm{ill})+(1-\alpha)\times \mathrm{P}(\mathrm{ill}\|\mathrm{fail})\right]$	proportion per serving
illnesses prevented by policy	I_{avoid}	$\left(1-\frac{P_{new}(ill)}{P_{baseline}(ill)}\right)\lambda_{ill}$	illnesses/year

Chapter 2 Identifying *Salmonella* of Greatest Concern

2.1 Public Health Context

This risk assessment leveraged FSIS' *Risk Profile for Pathogenic Salmonella Subtypes in Poultry* (available here) to identify *Salmonella* serotypes in chicken and turkey (individually and in aggregate) linked to foodborne illness. This independently peer-reviewed risk profile provided a comprehensive review of the scientific literature and foodborne illness data to identify certain *Salmonella* serotypes in poultry linked to foodborne illness. The genus *Salmonella* is classified on biochemical reactions, surface protein antigen profiles and DNA sequence. Currently, there are 2 recognized species, *enterica* and *bongori,* in which there are about 2,500 serotypes. The Kauffman-White Scheme was the original typing scheme used to describe serotypes, based on somatic (O) antigens, capsular Vi antigens, flagellar (H) antigens and lipopolysaccharides (Yan, 2004).

Salmonella subtypes are a group of *Salmonella* organisms with the same attributes. *Salmonella* serotypes are a subtype defined by a combination of O- and H- antigens (*i.e.*, a serogroup is a subtype with the shared attribute of O- and H- antigens) (Bauer, 2014). Thirty-two *Salmonella* subtypes (28 serotypes and 4 serogroups) can be attributed to human salmonellosis from consuming chicken and turkey products.

Evidence suggests that exposure to *Salmonella* subtypes of concern can cause severe or debilitating human health outcomes, including acute gastroenteritis, bacteremia (bacteria in the blood), and focal infections (persistent infection of an organ or region) resulting in hospitalization or chronic disease lasting beyond one year. The domestic foodborne hospitalization rate for *Salmonella* is about 2% and the fatality rate is about 0.04% for all *Salmonella* (Scallan, 2011). Antibiotic-resistant *Salmonella* infections pose a risk of treatment failure in the case of invasive disease and have been associated with severe outcomes.

2.2 Clustering Serotypes by Virulence Gene Markers

As part of the Cooperative Agreement between FSIS and UMD-JIFSAN, EpiX Analytics categorized *Salmonella* serotypes into two clusters derived from a machine learning algorithm using *Enterobacteriaceae* virulence factors.

High resolution genomic analyses have recently evolved as a result of the development of new computationally intensive approaches (Karanth, 2022) (Njage, 2019; Wheeler, 2018) (Chen, 2022) to assess a number of *Salmonella* strains or subtypes and the underlying genetic variability. To describe a broader range of serotypes, EpiX Analytics employed a genomics-based approach to group *Salmonella* serotypes based on a virulence gene profile of 193 *Enterobacteriaceae* virulence factors. To that end, seroclusters (i.e., groups of serotypes) could be constructed based on genetic similarities, and then validated via epidemiological characteristics. The resulting seroclusters could then be considered to develop refined dose-response model relationships. In the following sections, key components of the methods are summarized. Additional details and results of the EpiX Analytics methodology are provided in their report entitled "*Using genomics to identify nontyphoidal Salmonella serovars of concern and*

estimating dose-response models amenable to risk assessments in poultry[4].”

FSIS has also developed a Bioinformatics Supplemental Materials (available here) to describe the genomics-based clustering to ensure the transparency requirements of the Information Quality Act (Section 515 of Public Law 106-554), which is required for risk assessments used to inform rulemaking, such as this *Salmonella* in poultry risk assessment, are met. In addition to describing the EpiX Analytics approach, this appendix outlines general best practices in bioinformatics, contextualizes the approach used by EpiX Analytics within the discipline, and provides limitations of the approach and future directions that are of interest to FSIS.

Clustering approach

Clustering methods may group *Salmonella* serotypes in a variety of ways; however, in the approach performed by EpiX Analytics, the clusters were driven by the presence/absence of virulence factors (VFs) that are informative for clustering *Salmonella* serotypes into defined groupings. This clustering method relied on genes lost or gained in the isolate data curation (i.e., predicting open reading frames and gene annotation of isolate assemblies) as opposed to phylogenetic similarity measured by single nucleotide polymorphisms (SNPs), core genes, or O-antigen genes. Moreover, clustering was agnostic to the biological function or role of individual virulence factors as well as point mutations or insertions/deletions of genes that can modify gene function resulting in public health risk as illustrated by the emergence of *Salmonella* Reading (Miller, 2020). Nonetheless, further analysis into the biological function of each virulence gene can be utilized in future iterations and enhance virulence clustering models and classification schemes.

Over 40,000 pre-assembled isolates from human and animal sources (poultry and beef/bovine) in the U.S. and VFs from the *Enterobacteriaceae* family were compiled from public databases. VFs from the Enterobacteriaceae family were considered as these are more peripheral markers that may correspond to pathogenesis and affords the opportunity to include VFs commonly passed through horizontal gene transmission, while also providing the ability to find differences between serovars that the core genome would not uncover. Each pre-assembled *Salmonella* isolate was subsequently annotated to determine the virulence gene profile (presence/absence). Virulence genes that were present in the majority of isolates (>95%) as well as limited gene presentation (i.e., <10 total isolates) were removed from further analysis. Hence, 193 genes available for the clustering analysis included 57 *Salmonella* VFs, 94 E. coli VFs, 10 Shigella VFs, and 32 Yersinia VFs. The full list of these virulence factors along with additional descriptive characteristics are provided in FSIS’ Bioinformatics Supplemental Materials (available here).

The genetic similarity between all isolates was then estimated via an unsupervised random forest (URF) approach balancing computational expense and performance. To that end, 10,000 trees were simulated with 60 features (VFs) randomly selected as candidates for each split to maintain an appropriate level of efficiency without loss in the predictive power of the algorithm. Following the unsupervised random forest simulation, an isolate proximity matrix was estimated by averaging the distance between terminal nodes for each isolate across all trees. This result would imply the isolate relatedness; i.e., isolates in the same terminal nodes are more similar to each other. It is important to note that this averaging across trees can potentially overestimate the low values and underestimate the high values. Finally, the

[4] Description of the previous iteration of EpiX Analytic’s serotype clustering are available at Fenske, G. J., Pouzou, J. G., Pouillot, R., Taylor, D. D., Costard, S., & Zagmutt, F. J. (2023). The genomic and epidemiological virulence patterns of *Salmonella* enterica serovars in the United States. *PLoS One*, *18*(12), e0294624. https://doi.org/10.1371/journal.pone.0294624 .

isolates were grouped into *k* clusters using hierarchical clustering (Ward's method) and non-parametric bootstrapping to assess the stability (Jaccard stability and serotype switching). Results of cluster assignment based on how the majority of isolates classified (i.e., best cluster) are depicted in **Table 12**. The default labelling of the clusters (1, 2, 3, …) are determined by algorithm, however, results for k=2, 3, and 4 also exhibit decreasing associated risk as the numeric label increases in each *k* scenario (Risk multiplier section below).

Table 12: Best cluster assignment for 42 *Salmonella* serotypes resulting from EpiX Analytics' analysis.

Serotype	2 Clusters	3 Clusters	4 Clusters
Muenchen	1	1	1
I 4,[5],12:i:-	1	1	1
Typhimurium	1	1	1
Newport	1	1	1
Berta	1	1	1
Enteritidis	1	1	1
Litchfield	1	1	1
Saintpaul	1	1	1
Dublin	1	1	1
I 4,[5],12:b:-	1	1	1
Blockley	1	1	1
Hadar	1	1	1
Kentucky	2	3	4
Infantis	2	2	3
Schwarzengrund	2	2	2
Montevideo	2	2	2
Reading	2	2	2
Heidelberg	2	2	2
Anatum	2	2	2
Javiana	2	2	2
Cerro	2	2	2
Thompson	2	2	2
Braenderup	2	2	2
Agona	2	2	2
Senftenberg	2	2	2
Uganda	2	2	2
Mbandaka	2	2	2
Mississippi	2	2	2
Muenster	2	2	2
Johannesburg	2	2	2
Meleagridis	2	2	2
Oranienburg	2	2	2
Bareilly	2	2	2

Give	2	2	2
Lubbock	2	2	2
Brandenburg	2	2	2
Albany	2	2	2
Norwich	2	2	2
Alachua	2	2	2
Panama	2	2	2
Kiambu	2	2	2
Poona	2	2	2

Comparison of clustering results

To test whether other genomics-based clustering methods produced similar groupings, FSIS compared the k=4 cluster results from the most abundant serovars in the risk assessment with clusters obtained using reference-free SNPs (Timme, 2013), core genome approach (Worley, 2018), and O-antigen groupings (Grimont) in **Table 13**. The higher virulence cluster 1 serotypes, as defined by EpiX Analytics, generally clustered together in Timme cluster A2 (a major lineage of Clade A) and Worley cluster A (i.e., Clade A lineage). Timme cluster A2 contained 45 different serovars overall and included several cluster 2 serotypes, which demonstrated a high level of diversity as well as pathogenicity. The majority of Timme Clade B clusters (B1-B4) are monophyletic with unique SNPs.

Additionally, the broad grouping of lower virulence cluster 2 serotypes were further discretized in these approaches. Using the O-antigen grouping classified serotypes from both clusters into different groupings: D1 included Dublin and Enteritidis from cluster 1 as well as Javiana from cluster 2; C2-C3 contained Hadar, Muenchen, and Newport from cluster 1 and Kentucky from cluster 3; B included I 4,[5],12:i:- and Typhimurium from cluster 1 and Heidelberg, Reading and Schwarzengrund from cluster 2.

Risk multiplier

The clusters were validated by linking them to epidemiological data (i.e., documented outbreaks attributed to poultry sources with consideration of prevalence in animal sources from FSIS poultry sampling programs). In this sense, the relative risk estimate is skewed towards strains to which a poultry consumer is likely to be exposed. FSIS constructed **Figure 5** to graphically describe the process for estimating risk associated to each serocluster

Table 13: Comparing seroclusters developed under different approaches with EpiX Analytics' k=4 cluster scenario.

***Salmonella* Serotype**	**URF or SRF**	**EpiX Analytics k=4**				**reference-free SNPs (Timme 2013) N=156**						**Core Genome (Worley 2018) N=445**		**O-antigen (WHO Formulary)**					
	N > 500	C1	C2	C3	C4	B1	B2	B3	B4	A1	A2	B	A	E1	C1	K	D1	C2-C3	B
Enteritidis	5510	x									x		x				x		
Typhimurium	3421	x									x		x						x
Newport	2740	x									x		x					x	
I 4,[5],12:i:-	987	x									x		x						x
Dublin	697	x									x		x				x		
Saintpaul	612	x									x		x						x
Muenchen	607	x									x		x					x	
Hadar	558	x									x							x	
Schwarzengrund	1528		x						x			x							x
Reading	1299		x									x							x
Javiana	971		x				x					x					x		
Heidelberg	728		x								x		x						x
Anatum	673		x								x		x	x					
Cerro	591		x								x		x			x			
Thompson	549		x								x		x		x				
Braenderup	525		x								x		x		x				
Infantis	5604			x		x							x		x				
Kentucky	6413				x					x			x					x	

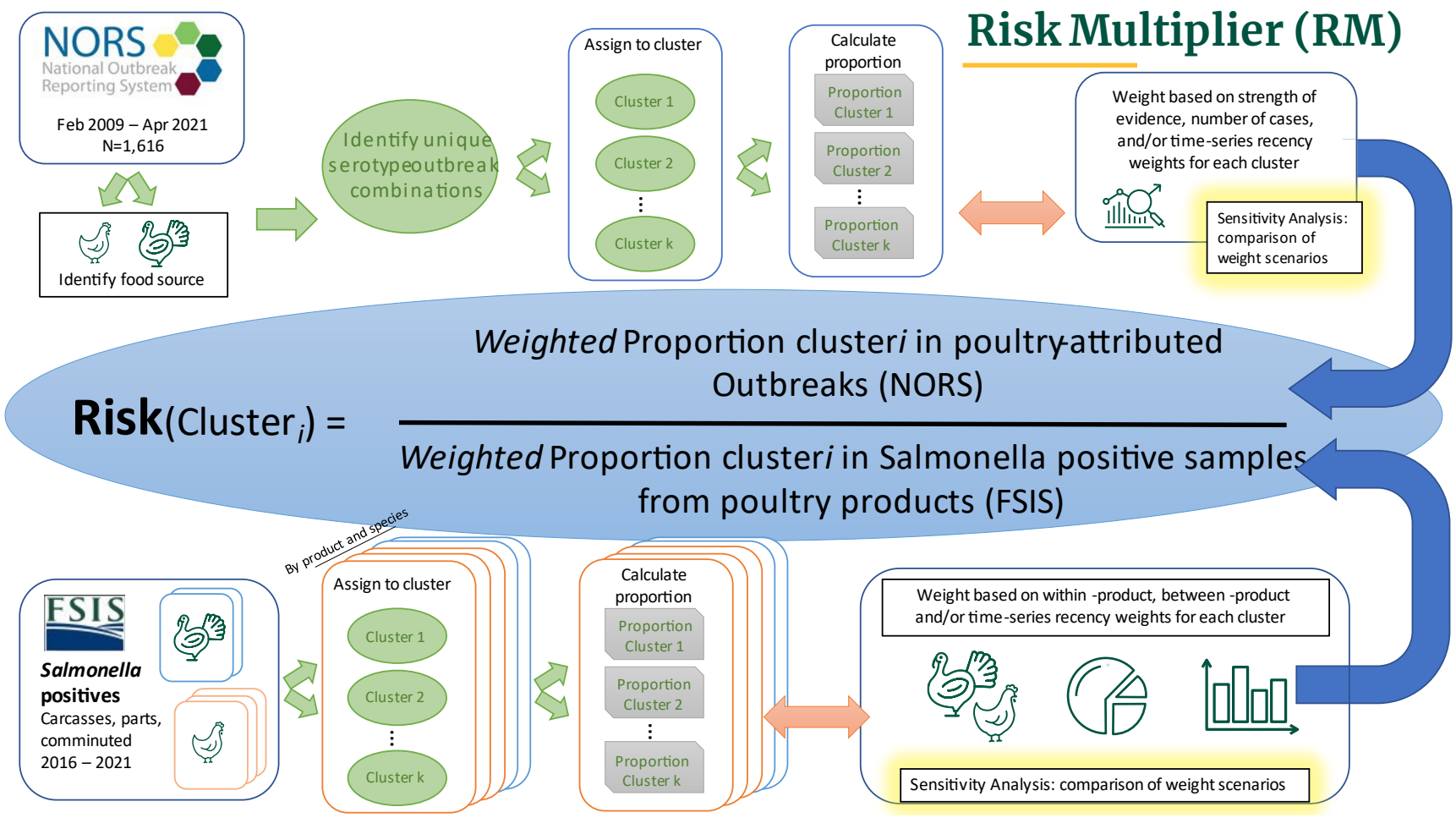


Figure 5: FSIS diagram of risk multiplier estimation for each serocluster.

There were 1,616 outbreaks initially considered based on data obtained from CDC NORS between 2009 and 2020. Within these outbreaks, 216 unique serotype-outbreaks combinations were identified and filtered down to outbreaks that were likely attributed to poultry based on Interagency Food Safety Analytics Collaboration (IFSAC) classification and text mining of ingredients associated to the outbreak origin sources. Strains with higher association to poultry-attributed outbreaks typically grouped together in cluster 1 for all *k* clustering scenarios.

FSIS conducted a preliminary exploratory analysis to illustrate the temporal dynamic of outbreaks for cluster 1 (C1) serotypes (**Figure 6**) and cluster 2 (C2) serotypes (**Figure 7**). Cluster 1 serotypes consistently comprised the dominating proportion of poultry-attributed outbreaks since 2013 with the majority of these outbreaks associated with Enteritidis. Two of the twelve cluster 1 serotypes (Litchfield and Dublin) were not associated with any poultry-attributed outbreaks. On the other hand, cluster 2 serotypes, Infantis and Reading, outbreaks have been increasingly observed more recently, but not to the same extent (i.e., number of outbreaks).

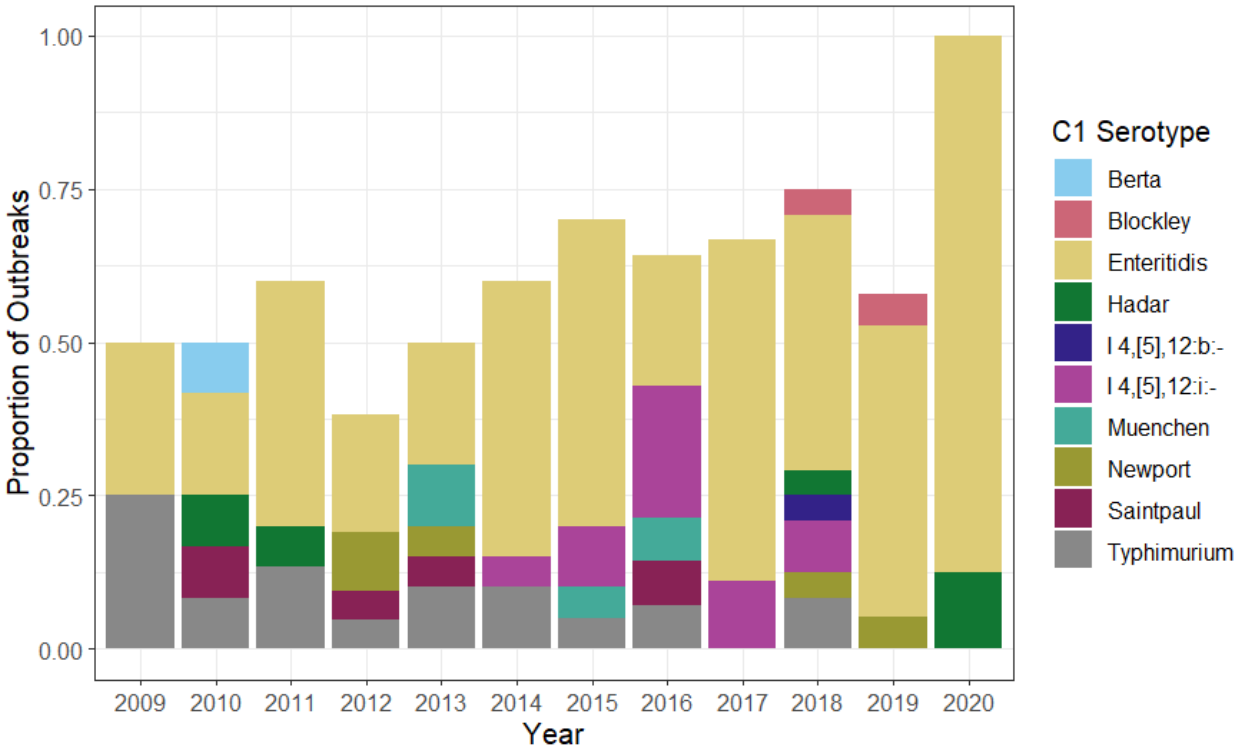


Figure 6: Composition of cluster 1 poultry-attributed outbreaks across time.

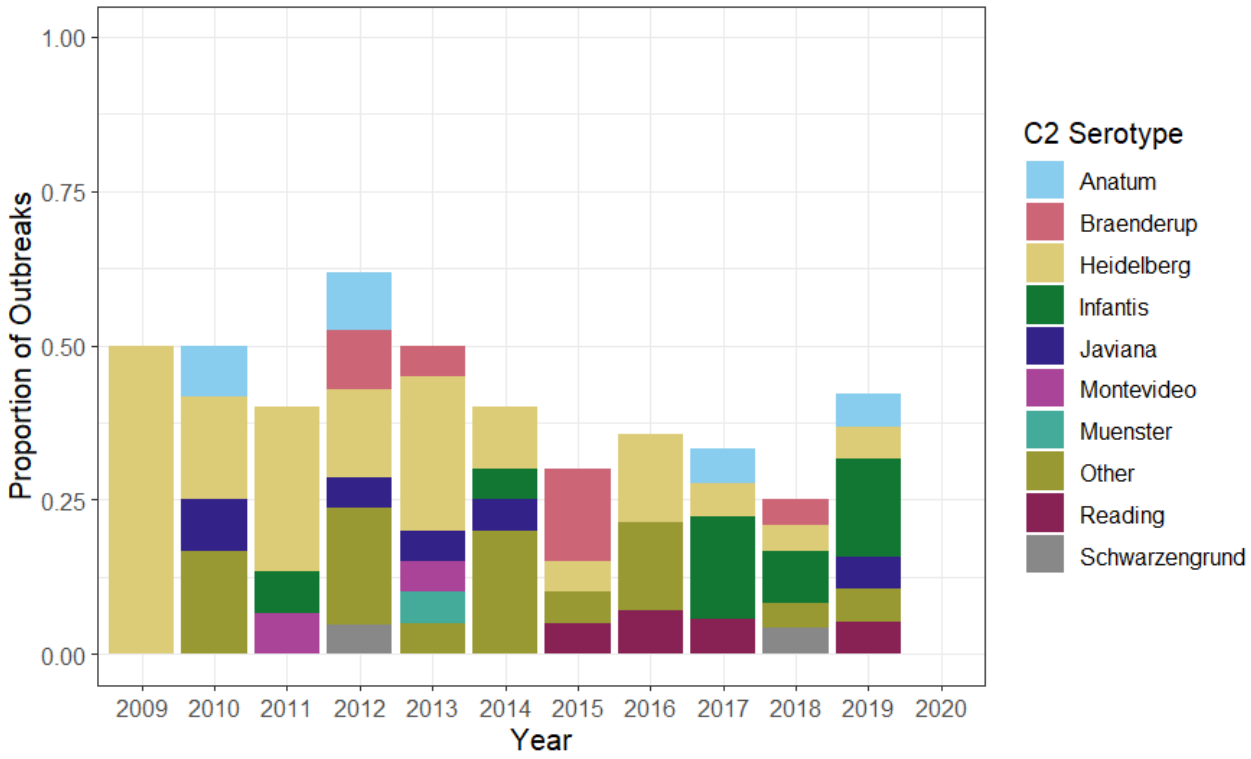


Figure 7: Composition of cluster 2 poultry-attributed outbreaks across time.

To elucidate the impact of more contemporary strains, additional transformations of the outbreak proportion should be considered (e.g., time-series component, number of primary cases, severity/strength of evidence). Of particular interest is a time-series component, as recent outbreaks are more representative of the current status of foodborne illness compared to more historical data. Following the approach described in (Batz, 2021), a recency weighting was used to capture the time-series component and provide one factor to encapsulate the shifting dynamic of outbreaks. Poultry-associated outbreaks older than 5 years (i.e., prior to 2017) were subjected to an exponential decay function with a decay parameter defined as 5/7 (0.7142).

Figure 8 illustrates the increasingly dominant picture of poultry-attributed outbreaks associated with cluster 1 serotypes compared to the broad range of cluster 2 serotypes as well as how the weights shift across time. Although there was a relative balance of cluster 1 and cluster 2 poultry-attributed outbreaks prior to 2015, these have less influence or weight on the risk multiplier estimation. Moreover, EpiX Analytics considered several additional factors to estimate the proportion in outbreaks component of the risk multiplier (**Table 14**); the complete derivation is described in **Appendix A**. Nonetheless, these epidemiological dynamics highlight the notion that risk multipliers must be continuously assessed to describe current risk to public health.

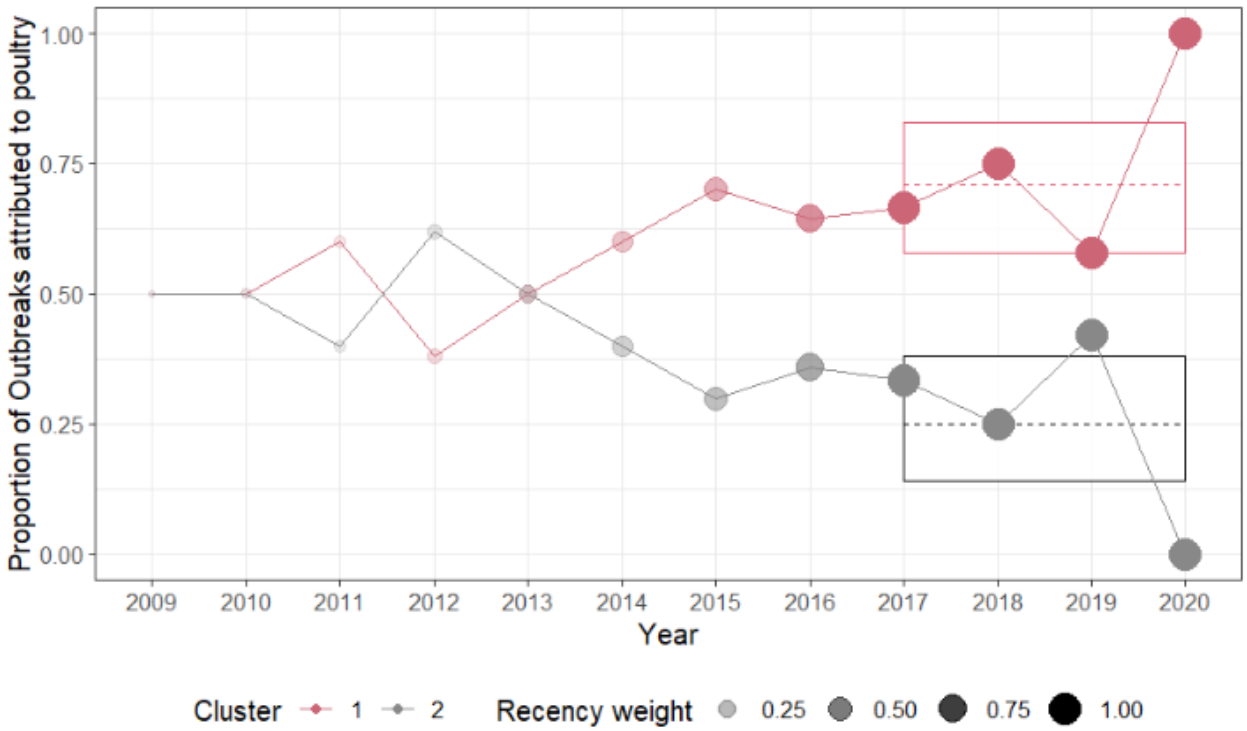


Figure 8: Time-series weighting scenario on the overall poultry-attributed outbreak proportion with the risk multiplier numerator estimation (dashed line) including confidence bounds (rectangles) overlaid on biased recent timeframe (2017-2020).

The risk multiplier denominator is derived from 2016-2021 FSIS regulatory sampling programs with consideration of product (i.e., chicken or turkey) and commodity (e.g., carcass, parts, comminuted). That is, to estimate the current status of *Salmonella* in poultry by cluster, annual production volumes were used to determine within-product and -commodity weights as well as between-product and -commodity weights based on general product availability and consumption rates, and a time-series (i.e., exponential decay) function (Batz, 2021) was applied to emphasize recent data compared to historical (prior to 2017, as previously considered). In particular, the following weights were incorporated: 83% chicken parts, 6% comminuted chicken, 11% chicken carcasses, 75% turkey carcasses, 25% comminuted turkey, and an overarching 5/1 chicken to turkey ratio. These weights dictate a commanding influence of chicken parts, followed by turkey and chicken carcasses, and finally, comminuted product across time. Additional details on the product type break down by cluster are described in **Chapter 3**.

The majority of *Salmonella* in poultry detected belongs to cluster 2, however, the majority of *Salmonella* in poultry-attributed outbreaks stem from cluster 1 serotypes (**Table 14**). Cluster 1 had a relative risk of 2.1 (95% CI 1.7-2.5) whereas cluster 2 was approximated as 0.38 (95% CI 0.21-0.58). Isolates that could not be assigned to a cluster are also presented.

Table 14: Risk multiplier estimation including the 95% confidence interval for k=2 seroclusters derived by EpiX Analytics.

	Cluster 1	**Cluster 2**	**Not Assigned**
Proportion in outbreaks (numerator)	0.71 [0.58; 0.83]	0.25 [0.14; 0.38]	0.039 [0.012; 0.081]
Proportion in poultry (denominator)	0.33 [0.31, 0.35]	0.66 [0.64; 0.68]	0.010 [0.006; 0.017]
Risk multiplier	**2.1 [1.7; 2.5]**	**0.38 [0.21; 0.58]**	**3.9 [1.1; 9.1]**

A description of the risk multiplier calculation is described in EpiX Analytics' report, **Appendix A**.

Sensitivity analysis

Different scenarios were considered in the derivation of these serocluster risk multipliers. The baseline case (presented in **Table 14**) considers various weights to evaluate/balance contributions from chicken and turkey outbreak characteristics including a mixed-effects model, a time-series component to differentiate recent information from historical data, and proportional cluster attribution rates. **Table 15** summarizes the different modeling scenarios assessed by EpiX Analytics during risk multiplier calculations for comparison with the baseline. It is readily observed that the associated risk was mostly consistent in each model except in the cases: (1) only using turkey data or (2) removing the time-series. Using turkey data only unnecessarily constrains the information feeding into the model. Out of the 216 unique serotype-outbreak combinations, only 44 were definitively attributed to turkey with a majority occurring prior to 2017. At the same time, within-product weights to estimate the proportion in poultry (i.e., denominator) considers 75% turkey carcasses to 25% comminuted turkey.

As described in the FSIS *Quantitative Risk Assessment for Salmonella in Raw Turkey and Raw Turkey Products* (available here), turkey carcass data is limited (roughly <10 detections annually), and thus, serotype proportions fluctuate dramatically. Additional data is required to appropriately refine the weighting scenario to consider turkey alone under this approach. In the case where the time-series is removed, the risk multiplier becomes heavily biased to historical data and does not accurately represent the changing dynamic of serocluster risk to public health.

In short, including turkey data does not change the results of this chicken *Salmonella* serocluster risk analysis. On the other hand, turkey data is not sufficient to determine risk, and chicken data is necessary for a turkey analysis.

Table 15: Risk multiplier sensitivity analysis of select scenarios considered by EpiX Analytics.

	Cluster 1	**Cluster 2**
Baseline*	2.1 [1.7; 2.5]	0.38 [0.21; 0.58]
Outbreak counts transformation	2.0 [1.6; 2.4]	0.39 [0.24; 0.55]
Estimated Primary cases transformation	2.0 [1.3; 2.5]	0.51 [0.22; 0.84]
No recency weighting	1.8 [1.4; 2.1]	0.53 [0.36; 0.70]
Recency weight starting to decrease after 1 year	2.4 [1.8; 2.9]	0.32 [0.14; 0.58]
Turkey only	1.7 [0.77; 3.0]	0.65 [0.22; 1.10]
Chicken only	2.2 [1.8; 2.6]	0.32 [0.16; 0.52]
Do not weight different products	2.4 [1.9; 2.9]	0.36 [0.21; 0.56]
Outbreaks Definitively or Probably attributed to poultry	2.1 [1.8; 2.4]	0.39 [0.24; 0.55]
Use best Cluster	2.1 [1.7; 2.5]	0.38 [0.21; 0.59]

The Choice of Two Seroclusters

When considering the number of clusters to use, the serotypes in the higher virulence cluster 1 (e.g., Enteritidis and Typhimurium) remained the same across the choices of 2, 3 or 4 clusters. Serotypes in this cluster have an estimated relative risk of 2.1 (i.e., the risk of illness is 2.1 times higher than the probability of illness, prior to knowing that the strain belonged to cluster 1). In the 2-cluster model, the relative risk for the lower virulence cluster 2 is 0.38. This leads to a large difference in the probability of illnesses between the clusters, which is estimated to be 2.1/0.38=5.5. Adding a third cluster did not change the serotypes in cluster 1 but did divide cluster 2 into two low-virulence clusters, with the lowest virulence cluster 3 consisting primarily of Kentucky. Adding a fourth cluster resulted in a new cluster consisting of Infantis. The relative risk for the Infantis cluster was 0.31, so the influence of Infantis would be less in the 4-cluster model than the 2-cluster model. While the relative risk for the cluster consisting of Kentucky in either the 3- or 4-cluster models is low, a conservative assumption is to include Kentucky with the other lower virulence serotypes. This choice is justified by noting that majority of Kentucky isolates from U.S. poultry samples are of the Group 1 variety. Nevertheless, the more virulent Kentucky Group 2 (Soltys, 2021) has been recently isolated from chicken samples in the U.S. (Thompson, 2018), so these findings should be revisited periodically to determine if *Salmonella* Kentucky maintains its low-virulence status.

Table 16: Dose-response model multipliers for 2, 3, and 4 *Salmonella* seroclusters.

	Multipliers for k = 4 (Estimate [bootstrap 95% CI])			
	Cluster 1	**Cluster 2**	**Cluster 3 (Infantis)**	**Cluster 4 (Kentucky)**
Multiplier	2.1 [1.7;2.5]	0.81 [0.44;1.30]	0.31 [0.0095;0.89]	0.01 [0.000;0.094]

Inclusion of Salmonella Infantis in cluster 2

As discussed, the genomic data and associated analysis classify Infantis as a lower virulence serotype, of lower risk than any serotype other than Kentucky. As the initial FSIS Key Performance Indicators (FSIS,

2022a) did include Infantis on the basis of the best available human illness data at that time, FSIS chose to further evaluate this serotype clustering.. Specifically, FSIS considered evidence linking human Infantis cases to consumption of chicken contaminated with Infantis, which would imply that increases in human illnesses associated with Infantis are driven by the increasing proportion of Infantis identified on broiler chickens (McMillan, 2022; Williams, 2022). A subanalysis was also conducted to compare the changes in salmonellosis cases reported to FoodNet, with the change in the proportion of *Salmonella* positive samples whose serotype was Infantis.

The trend for Infantis salmonellosis cases was estimated using the Infantis case rate data between 1996 and 2019 acquired from FoodNet Fast (CDC, 2022a). A penalized B-spline model (Powell, 2016) was fitted to these data to derive the trend in case rate per 100,000. The trend in the proportion of Infantis isolates associated with broilers is derived using a compositional data model (Aitchison, 1982; Faes, 2011) fitted to the FSIS carcass testing data.

The human case data exhibits an initial decline between 1996 and 2004, though the decline is not statistically significant. Beginning in roughly 2005, the case rate begins to increase, with the average case rate nearly tripling between 2010 and 2019. The annual rate of increase is almost perfectly linear during this time period.

The FSIS data demonstrates that Infantis was a minor serotype in broilers from 1998 through roughly 2012 and generally not associated with human cases (Shah, 2017). After 2012, there is a rapid increase in the proportion of *Salmonella* positive samples with Infantis, with the first multidrug resistant variants identified in 2014 (Tate, 2017). Between 2016 and 2019, this fraction roughly triples, with many of the newer isolates having similar multidrug resistant characteristics (Tyson, 2020; Williams, 2022).

The comparison of these trends (**Figure 9**) demonstrates a roughly six-year lag between the increase in human cases and contamination in broilers, which suggests that other reservoirs and modes of transmission (other than broilers) were responsible for the initial increase in cases of salmonellosis. In the later years, if the consumption of chicken were a major contributor to sporadic cases of salmonellosis, one would expect that the rapid increase in the Infantis observed in broilers would have resulted in some additional increase in the case rate observed in FoodNet. While chicken consumption is likely to be related to some of the cases, the continued linearity in the case rate suggests that the contribution of chicken is a minor component.

In summary, *Salmonella* large increases in *Salmonella* FoodNet case rates are observed for Infantis that do not coincide with the changes in chicken, leading to the conclusion that the increase in human Infantis cases is not related to the consumption of Infantis contaminated chicken. Further, Infantis is a lower virulence serotype, as demonstrated in the EpiX Analytics clustering, that is currently outcompeting higher virulence serotypes, such as Enteritidis and Typhimurium in chicken. Therefore, Infantis should be considered together with the other lower virulence *Salmonella* serotypes.

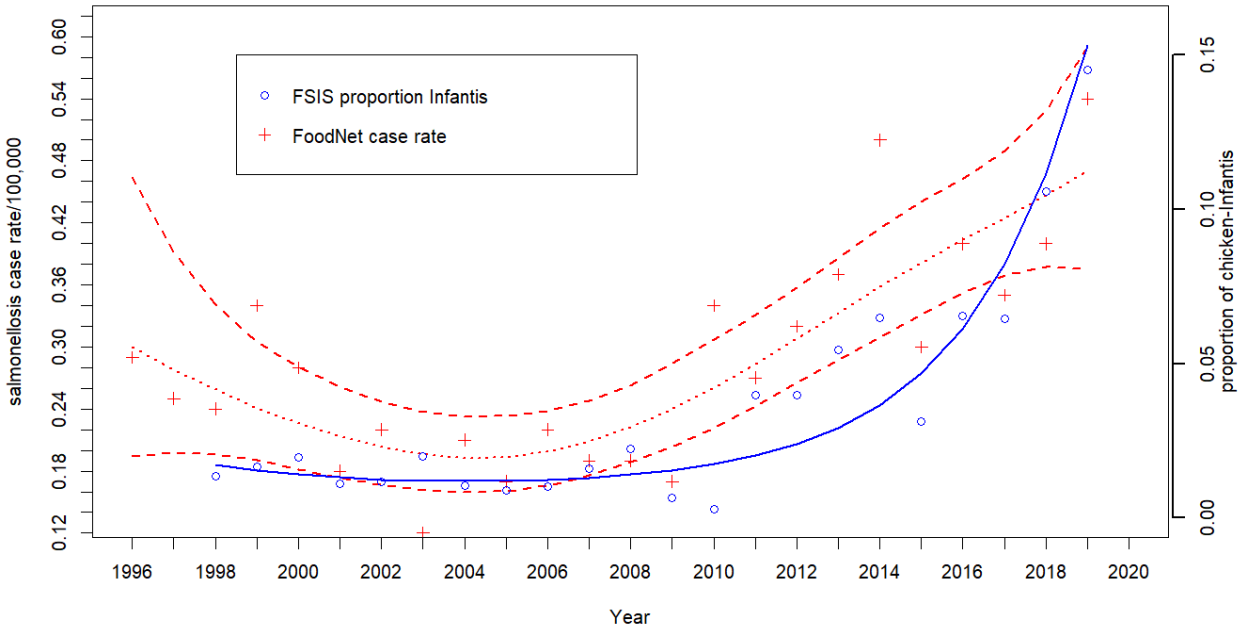


Figure 9: Comparison of changes in the Infantis-related cases in FoodNet with changes in the occurrence of Infantis amongst *Salmonella*-positive broiler samples. Confidence intervals are not provided for the FSIS data because the properties of compositional data (i.e., statistics that are constrained to sum to unity).

Dose-response models

Dose-response models, developed by EpiX Analytics, were approximated for the k=2 serocluster result. The higher virulence cluster 1 dose-response model was estimated using outbreak data and employing a beta-Poisson model of infection for a given dose as derived in (Teunis, 2010; Teunis, 2008). The risk multipliers (**Table 14**) were then used to scale the relative risk of illness from exposures to each cluster. That is, cluster 1 dose-response was developed using data from the literature on Enteritidis and Typhimurium (two primary serotypes in cluster 1) and scaled a second dose-response model for the lower virulence cluster 2 based on the risk multiplier ratios. Finally, a polynomial regression was fit to the dose-response functions for swift implementation in the risk assessment model.

The polynomial approximation of the dose-response models was used to estimate some useful illness doses. For higher virulence cluster 1 *Salmonella* serotypes, the ID50, the dose at which 50% of individuals in an exposed population will experience symptomatic illness, is approximately 2000 cfu. For lower virulence cluster 2 *Salmonella* serotypes, the ID50 is not attained, with at most 40% of an exposed population becoming ill at doses higher than 1 billion cfu. There is a 1 in 100 hundred probability of illness at 1 cfu of higher virulence *Salmonella* per serving. While for the lower virulence serotypes, the dose response model estimates a 0.002 probability of illness at 1 cfu *Salmonella* per serving. For comparison, the FAO/WHO *Salmonella* dose-response model estimated a 13 percent chance of becoming ill if ingesting an average dose of 100 organisms (FAO/WHO, 2002). Even at the level of 1 organism ingested, there was still a non-zero chance of illness (0.25%).

Full details of the dose-response model development are described in EpiX Analytics' report, **Appendix A**.

2.3 Serotypes of Public Health Significance

The higher virulence serotypes which appear most frequently in FSIS chicken samples (Enteritidis, Typhimurium, and I 4, [5],12:i:-) are summarized in **Table 17** and referred to as "serotypes of public health significance" in this risk assessment. The portion of FSIS PR/HACCP *Salmonella* positive samples that are sequenced as a serotype of public health significance is 24% for chicken carcasses, 33% for chicken parts, and 29% for comminuted chicken product.

Table 17: Higher virulence *Salmonella* serotypes in FSIS PR/HACCP poultry sampling. An X indicates the serotype is among the top 10 FSIS serotypes for that product. The average percent of *Salmonella* positive samples that are higher virulence or top 3 higher virulence positive are also included.

Higher Virulence Serotypes	Chicken Carcass 2016-2021 N = 2,602	Chicken Parts 2016-2021 N=6,437	Comminuted Chicken 2016-2021 N=2,860
Berta			
Blockley			
Dublin			
Enteritidis	X	X	X
Hadar	X		
I 4,[5],12:b:-			
I 4,[5],12:i:-	X	X	X
Litchfield			
Muenchen			
Newport			
Saintpaul			
Typhimurium	X	X	X
Top 3 Higher Virulence Serotypes	24%	33%	29%
All Higher Virulence Serotypes	26%	35%	32%

Chapter 3 *Salmonella* Microbial Profile

The first goal of this risk assessment is to define a probabilistic model that explains the current state of pathogen contamination in the U.S. chicken carcass population. The effect of various risk management options can then be assessed by adjusting the parameters that describe the population in accordance with the anticipated effect of different risk management options. Probabilistic models are constructed for boiler carcasses, chicken parts, and comminuted chicken. These models are combined into a single model that describes *Salmonella* levels per gram for all meat from chicken products.

The microbial data used in the *Salmonella* model is summarized in **Table 18**. Unless otherwise stated the data used for subsequent analyses of chicken carcasses are the FSIS PR/HACCP carcass *Salmonella* verification samples paired by flock with FSIS *Salmonella* exploratory program rehang samples. This carcass data includes samples consisting of a single carcass randomly chosen from a flock at both rehang and then again at post-chill. No attempt is made to choose the same carcass at each location. Each pair of samples was obtained from a single flock or production lot and the sample was collected using a carcass rinse consisting of 400 mL of neutralizing buffered peptone water (nBPW). Further details of the sample collection methodology and laboratory techniques used to assess the sample are available in the FSIS Poultry Exploratory Sampling Program Report (available here) and described elsewhere (FSIS, 2021; FSIS, 2022d).

Full descriptions of data used in this risk assessment are available in **Appendix B** and data used in the dose-response development in **Appendix A**. Details of the population description methodology are available in **Appendix C**.

Table 18: Description of main sources of data used in the risk assessment.

Product	Data	Limitations
Carcass Microbial Data	• Establishment-level FSIS PR/HACCP chicken carcass post-chill samples from the *Salmonella* verification program results from 2016 through October 31, 2022 (post-chill) • Establishment-level FSIS chicken carcass rehang samples from the *Salmonella* exploratory program results from April 2022 through October 31, 2022 (rehang) • FSIS establishment-level *Salmonella* data from young chicken microbiological baseline study from July 2007 through September 2008 (post-chill and rehang)	• One sample is collected per location per flock. • No attempt is made to sample the same carcass at rehang and post-chill. • No information is available for pre-harvest occurrence or levels of *Salmonella* or indicator organisms.
Parts Microbial Data	• Establishment-level FSIS PR/HACCP chicken parts samples from the *Salmonella* verification program results from 2016 through October 31, 2022 (Prevalence, Serotype) • FSIS establishment-level *Salmonella*	• There is only one sample per lot of parts product. This limits the ability to analyze serotype standards for these products

	data from chicken parts microbiological baseline study from January-August 2012 (Prevalence, Enumeration)	
Comminuted Microbial Data	• Establishment-level FSIS PR/HACCP comminuted chicken samples from the *Salmonella* verification program results from 2016 through October 31, 2022 (Prevalence, Serotype; Enumeration through February 2020)	• There is only one sample per lot of comminuted product. This limits the ability to analyze serotype standards for these products.

From the *Salmonella* testing data, it is possible to estimate the prevalence of test-positive carcasses, the prevalence of carcasses with a specific serotype or belonging to a virulence cluster, and an industry-wide distribution of *Salmonella* cfu/mL in the assay. These estimates are made for both rehang and post-chill. For chicken parts and comminuted chicken, the same estimates are provided for both products at the end of production.

3.1 *Salmonella* Prevalence

The prevalence of *Salmonella*-positive carcasses was estimated using methods described in the **Appendix C** and the 2022 exploratory paired carcass data and report. **Table 19** provides estimates for various statistics at rehang and post-chill, after accounting for the difference in establishment production volume.

Table 19: *Salmonella* prevalence at rehang and post-chill.

Sample Location	***Salmonella* Prevalence** $(\widehat{\mathbf{P}})$	**Standard Deviation** $\left(\sqrt{var[\widehat{P}]}\right)$	**95% confidence interval**
Rehang	0.655	0.0074	(0.641, 0.670)
Post-chill	0.031	0.0030	(0.026, 0.037)

The prevalence at rehang is higher than the proportion of positive samples reported in the previous FSIS microbiological baseline study, which was 0.458, but the previous microbiological baseline estimate was not adjusted for the differences in production volume across the population of establishments. The post-chill prevalence of 0.031 represents a 59% reduction from the volume-adjusted estimate of 0.075 in the previous microbiological baselines (FSIS, 2009b).

Production Volume and Salmonella

Since the beginning of regular sample collection by FSIS in the 1990s, the presence or absence of microbial contamination in the meat and poultry industry has been strongly related to the production volume of slaughter and processing establishments. In this risk assessment, high-volume is defined as producing more than about 10 million chicken carcasses per year, on a roughly weekly basis. Lower-

volume establishments have production volume ranges from about 20,000 to 10 million chicken carcasses per year.

These volume definitions do not align with FSIS' standard PR/HACCP definition of establishment size, but rather are based on analyses of FSIS' data that showed a clear separation between high-volume and lower-volume establishments. FSIS attempts to collect samples from high-volume establishments at a rate of approximately five times per month. Lower-volume establishments are sampled at a rate of approximately twice per month. This sampling frequency leads to a bimodal distribution of the number of samples collected at each establishment (**Figure 10**).

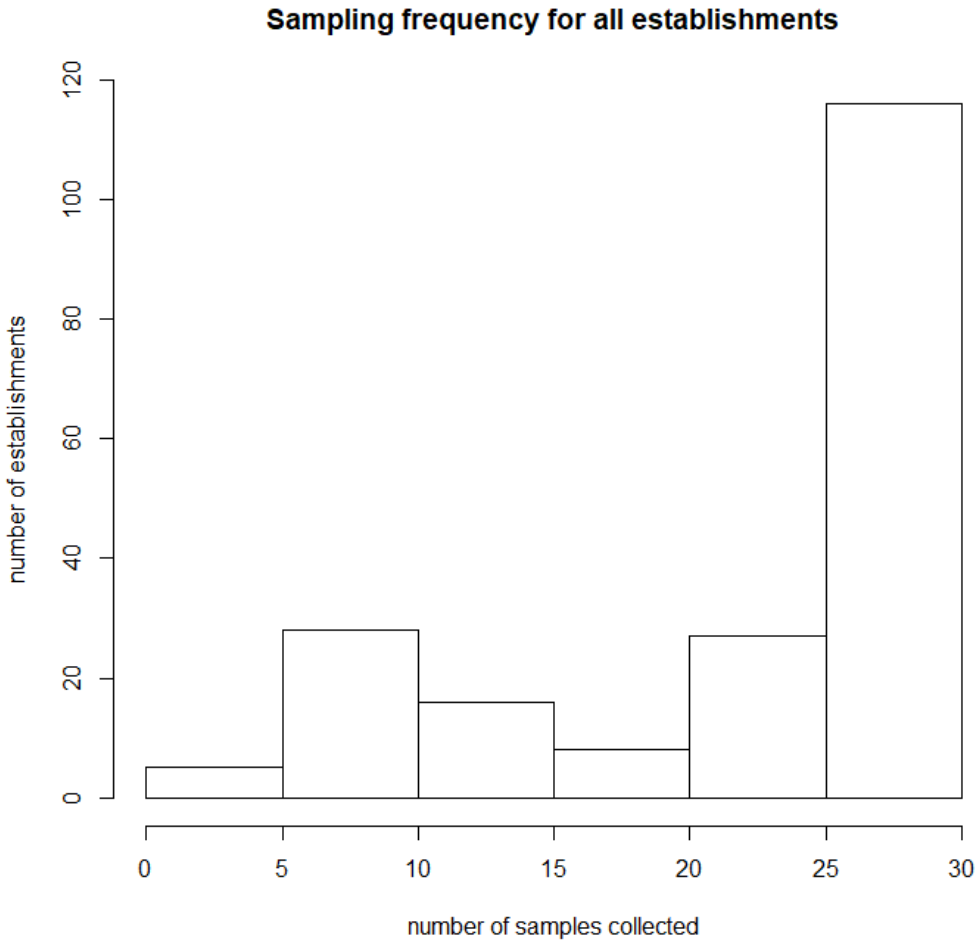


Figure 10: The distribution of the number of PR/HACCP *Salmonella* verification samples FSIS collects in each establishment per year. Higher-volume establishments are characterized by greater than 20 samples, while lower-volume establishments generally have fewer than 10.

For the poultry industry, *Salmonella* and *Campylobacter* occurrence is more frequent on products produced by lower-volume establishments. The opposite phenomenon is observed in the pork and beef industries, where a small number of large establishments account for the majority of the contaminated product reaching consumers (Williams, 2022).

Another factor that is likely to affect pathogen occurrence is establishment ownership, with some establishments being owned and operated by a single corporation. Industry-developed pathogen control programs for these establishments are likely coordinated across all establishments operated by the single corporation, so it is reasonable to expect no significant differences in performance for the collection of establishments owned by a single corporation. Given these differences in industry structure and their possible contribution to the occurrence of *Salmonella* on final products, the results of this risk assessment should be considered within the context of production volume and establishment ownership.

Analysis of the 2022 exploratory paired carcass data indicates nearly all of the corporately-owned establishments produce greater than 30 million carcasses per year and *Salmonella* is rarely found at post-chill in these establishments. Conversely, most lower volume establishments that are not part of a corporate structure. As a consequence, FSIS sampling does not have a high discriminatory power to find contamination differences among these establishments (**Figure 11**).

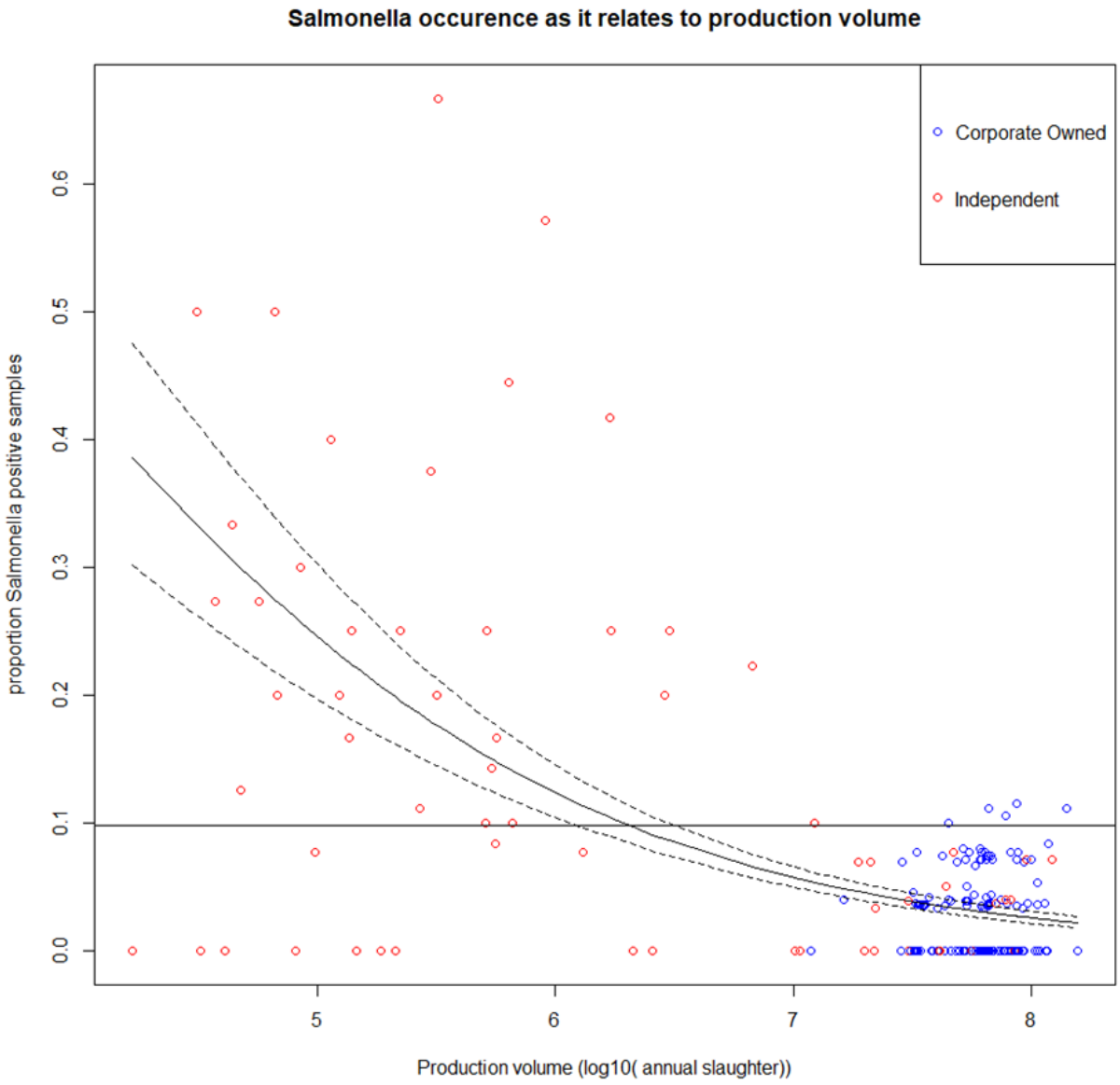


Figure 11: The relationship between production volume, ownership, and *Salmonella* occurrence. The horizontal line represents the maximum allowable proportion of *Salmonella*-positive samples under the current performance standard.

Homogeneity of Higher Volume Establishments

High volume establishments are homogeneous in nature. While **Figure 11** shows a significant negative relationship between *Salmonella* occurrence and production volume across the entire industry, no such significant relationship exists for the 158 establishments that slaughter greater than 10 million chicken carcasses per year (**Figure 12**). For this subset of the industry, which constitutes 99.6% of all chicken carcass production currently subject to FSIS pathogen reduction performance standards, the slope coefficient for the generalized linear model is not significant (p=0.4). Furthermore, 47% of higher volume establishment had no positive *Salmonella* samples. The fraction of establishments that had 1, 2

or 3 positives samples was 32%, 18%, and 3%, respectively.

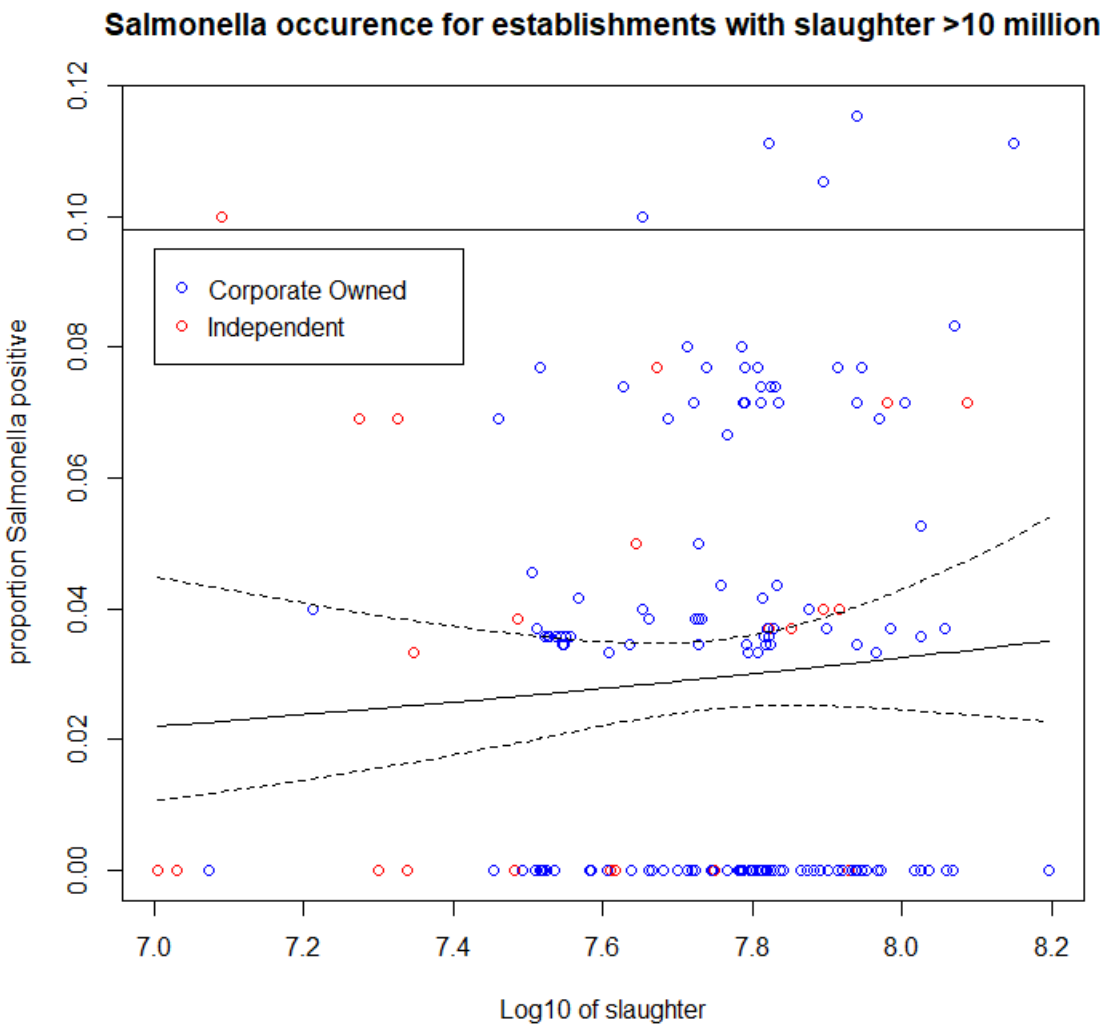


Figure 12: The proportion of *Salmonella* positive tests at post-chill for those establishments producing greater than 10 million chicken carcasses per year. *Salmonella* occurrence in these establishment is infrequent, with a large fraction of establishments having no positive samples in a given 12-month period.

3.2 *Salmonella* Levels

The parameter values for the lognormal level distributions derived from the rehang and post-chill 2022 exploratory paired carcass sampling data are provided in **Table 20**. The implied prevalence value is the probability of a 30 mL sample aliquot containing one of more viable *Salmonella* (i.e., 1 minus the cumulative distribution function at the theoretical limit of detection of x=1/30). The attenuation bias associated with the estimation of the level in each sample increases in magnitude as both the degree of measurement error and the proportion of samples below the limit of detection (LOD) of the assay increases. While the exact magnitude of the bias in each parameter is unknown, the direction of the

bias is consistent, with the $\hat{\mu}$ consistently overestimating the true μ parameter and $\hat{\sigma}$ underestimating σ the true parameter.

At rehang, the fraction of samples that test-positive and the implied prevalence were deemed to be close enough to the actual prevalence (0.59 vs 0.66) to not adjust for the parameters to account for the attenuation bias. In contrast, simulation results demonstrate that the magnitude of the attenuation bias in the estimated μ parameter of a lognormal distribution tripled from 0.2 to 0.5 when the true population μ decrease from -2 to -3.5 and the proportion of samples that were positive on the screen test decreased from 0.5 to 0.17 (Williams, 2012a). Given that the Poisson component of the Poisson lognormal distribution should account for some of the measurement error in the estimated levels, an adjustment of 0.4 logs was chosen for the $\hat{\mu}$ parameter and the $\hat{\sigma}$ was then adjusted using a nonlinear optimization routine so that the cumulative distribution at the limit of detection of the assay matches the fraction of samples that test positive $\hat{P}$.

The parameters of the level distribution at rehang are similar to those derived from the 2007-2008 FSIS young chicken microbiological baseline study $\hat{\mu}_{2007-8} = -1.41$ and $\hat{\sigma}_{2007-8} = 1.16$, suggesting that levels of *Salmonella* at rehang have not substantially changed in the last decade. In contrast, the level at post-chill distribution from the previous microbiological baseline was approximately 1 log lower ($\hat{\mu}_{2007-8} = -3.49$ and $\hat{\sigma}_{2007-8} = 1.33$), as was the case with the aerobic count (AC) data (page 139). Of note, the prevalence of *Salmonella* positive samples and the level distributions are similar to results presented in a recent bio-mapping study of a high-volume chicken carcass slaughter establishment (De Villena, 2022).

The bottom row of **Table 20** provides the parameters for the population lognormal distribution of total *Salmonella* on chicken carcasses. These parameters reflect the subsampling of the original 400ml rinse sample and the assumed removal rate of bacteria for rinse sampling of 0.14 (Lillard, 1988). The implied prevalence, which is defined as the mass of the lognormal distribution above the limit of detection of the assay, derived from these parameters indicates that 61% of all carcasses contain at least one viable *Salmonellae* at post-chill. This value is reasonable in light of a recent survey of retail meat establishments that found that 63.2% of enriched whole carcass samples were *Salmonella*-positive (i.e., an assay with a theoretical limit of detection of 1 viable *Salmonellae* on the carcass) (Talukder, 2022). Similarly high fractions of *Salmonella*-positive samples are reported in other studies that employ various whole carcass and whole part enrichment methods (Simmons, 2003a; Simmons, 2003b).

Table 20: Parameter values for the lognormal *Salmonella* level distributions derived from the 2022 exploratory paired carcass rehang and post-chill data. All rinse sample data are in units of cfu/mL. A final conversion is used to convert to units of cfu/g.

Sample Location	$\hat{\mu}$	$\hat{\sigma}$	Implied prevalence
Rehang	-1.210	1.416	0.588
Post-chill unadjusted	-4.112	1.174	0.014
Post-chill adjusted for attenuation bias	-4.512	1.624	0.031
Whole carcass **(**Conversion: $\hat{\mu}_{wc} = \hat{\mu} + \log_{10}(400/0.14)$ **)**	-1.056	1.624	0.612
Concentration/gram **(**Conversion: $\hat{\mu}_{gram} = \hat{\mu}_{wc} + \log_{10}\left(\frac{1}{454 \times 4}\right)$ **)**	-4.315	1.624	NA

The data and fitting methods used to determine the level distributions for parts and comminuted are listed in Appendices B and C. The shape of the distributions normalized to the per gram basis are given in **Figure 13**.

Table 21: Parameter values for the lognormal *Salmonella* level distributions derived for the parts and comminuted data.

Commodity	$\hat{\mu}$	$\hat{\sigma}$	Implied prevalence
Parts/mL	-3.017	1.027	0.067
Parts/gram	-2.820	1.027	NA
Comminuted/gram	-3.700	1.949	0.271

The location and shape of these distributions differ in a manner consistent with what would be biologically expected. For example, the partitioning and mixing of carcasses into parts should increase

the central tendency of the distribution because the parts sample is a pooled sample (Cowling, 1999) of carcasses that were both test-positive and test-negative. Similarly, the mixing of parts with differing levels should result in a more homogeneous distribution of contamination. Comminuted product has both a higher average level than the carcass distribution as well as having heavier tails. This is not unexpected because comminuted product can also be described as a pooled sample. The heavier tails for the distribution are also consistent with the product having a higher percentage of skin, which is a primary harborage site for *Salmonella* on carcasses (Rimet, 2019).

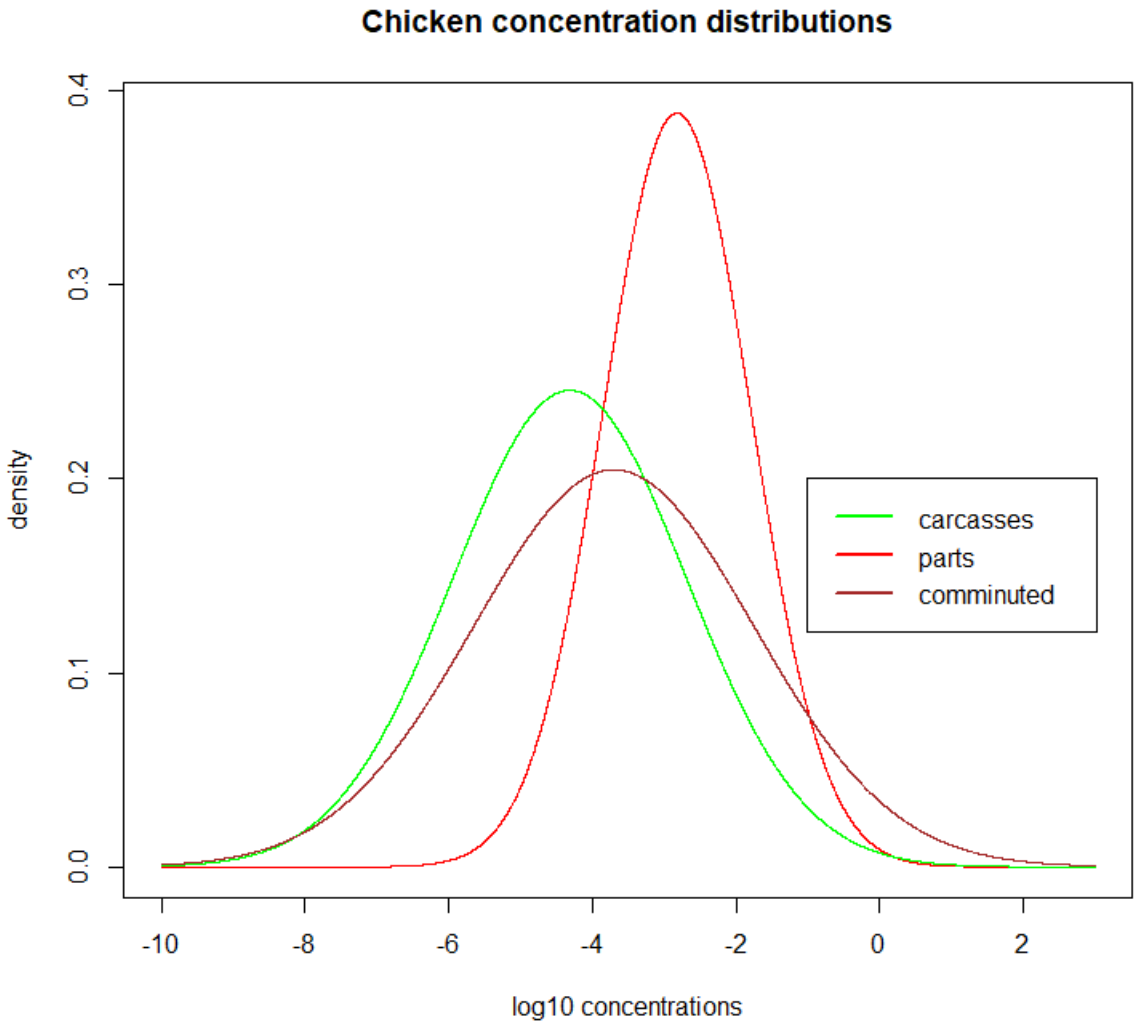


Figure 13: The three distributions represent the log10 transformed concentrations (that is, levels) of *Salmonella* for the three chicken products (i.e., carcasses, parts, and comminuted product).

To fit the dose-response model developed in **Appendix A**, it was necessary to determine the moments (i.e., expected value, variance, etc.) of a single mixture distribution describing the per gram *Salmonella* level from the weighted contribution of the distributions for the three products. The proportion of

servings consumed from whole carcasses, parts, and comminuted product are 0.11, 0.83 and 0.06[5], respectively. The first two moments of the mixture distribution are given by

$$\hat{\mu}_{all} = 0.11 \times \hat{\mu}_{carc} + 0.83 \times \hat{\mu}_{parts} + 0.06 \times \hat{\mu}_{comm} = -3.037$$

and

$$\hat{\sigma}_{all} = \sqrt{\begin{array}{l} 0.11 \times \left(\left(\hat{\mu}_{carc} - \hat{\mu}_{all} \right)^2 + \hat{\sigma}_{carc}^2 \right) + \\ 0.83 \times \left(\left(\hat{\mu}_{parts} - \hat{\mu}_{all} \right)^2 + \hat{\sigma}_{parts}^2 \right) + \\ 0.06 \times \left(\left(\hat{\mu}_{comm} - \hat{\mu}_{all} \right)^2 + \hat{\sigma}_{comm}^2 \right) \end{array}} = 1.28$$

Simulating the shape of log10-transformed realizations drawn from the three individual distributions demonstrates that the resulting distribution is unimodal and reasonably symmetric. The distribution is also reasonably approximated by a single normal distribution, with the greatest departure from normality being in the left tail, where the level less than 1 organism per million grams. For these reasons, a lognormal distribution with parameters $\hat{\mu}_{all} = -3.037$ and $\hat{\sigma}_{all} = 1.28$ was chosen to model the overall level per gram of *Salmonella* at the end of production.

[5] In previous work, these estimated fractions were 81%, 13% and 6% for parts, carcasses and comminuted product (USDA Food Safety and Inspection Service. (2015). Public Health Effects of Raw Chicken Parts and Comminuted Chicken and Turkey Performance Standards. https://www.fsis.usda.gov/news-events/publications/public-health-effects-raw-chicken-parts-and-comminuted-chicken-and-turkey .But subsequent data National Chicken Council. (2022). *How Broilers are Marketed*. Retrieved December 26 from https://www.nationalchickencouncil.org/statistic/how-broilers-are-marketed/ suggests the share of product marketed as whole carcasses has decreased. Therefore, we adjusted the carcass share down and increased the parts share, accordingly.

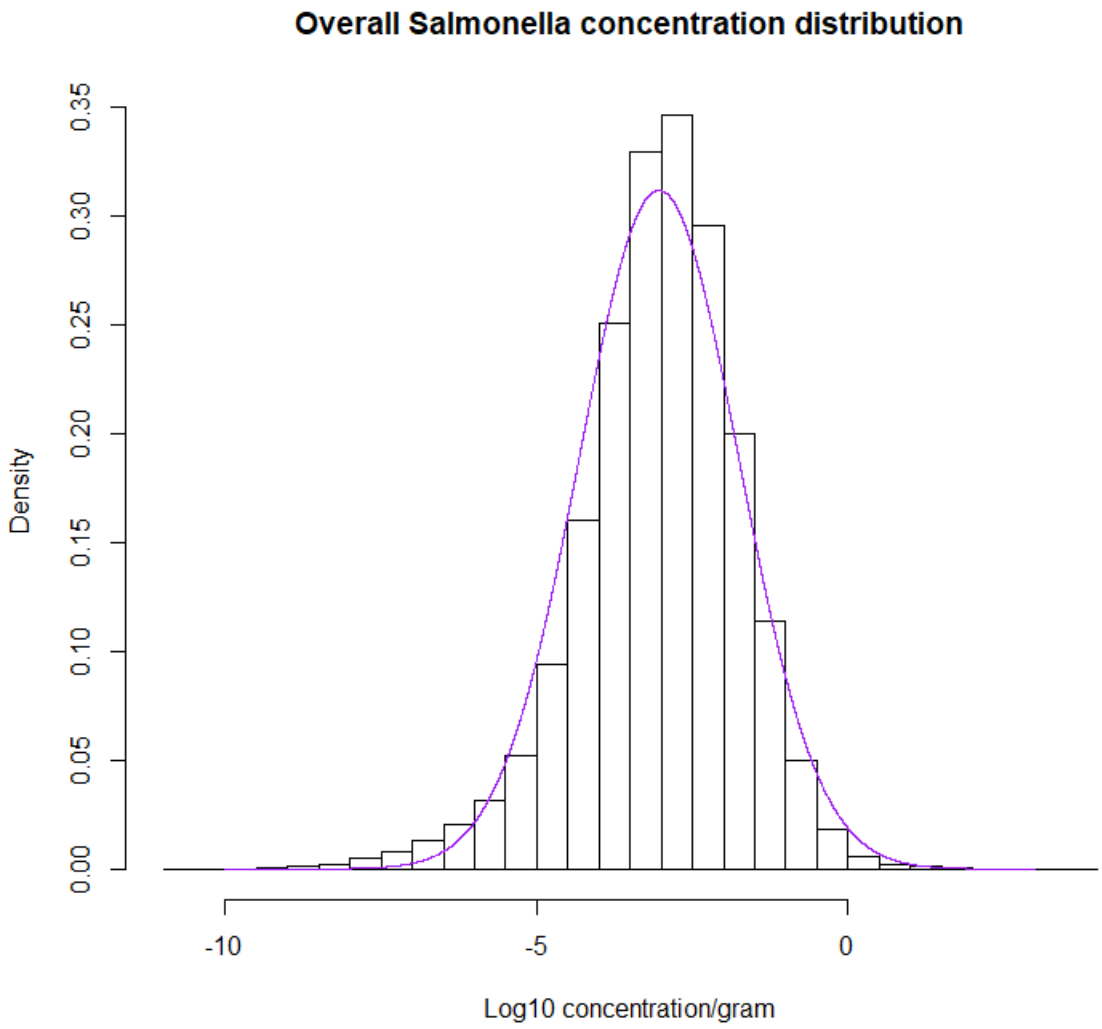


Figure 14: The Monte Carlo-based distribution describing the log10 transformed concentration (i.e., level) of *Salmonella* on all chicken products (i.e., average cfu/g level across all three product types).

3.3 *Salmonella* Serotypes

In samples that were confirmed positive for *Salmonella*, WGS was used to identify serotypes. **Table 22** highlights the top serotypes found by commodity from PR/HACCP sampling accumulated over a six-year period (2016-2021) and the top ten serotypes from the 2020 CDC FoodNet annual summary. Serotypes Kentucky, Infantis, Enteritidis, Typhimurium, and Schwarzengrund are consistently ranked in the top 5, regardless of commodity. These five serotypes were observed in over 85% of serotyped samples for each commodity (i.e., carcass; 91%, comminuted; 85%, and parts; 89%). Serovars Enteritidis and Typhimurium also appear in the top five serotypes from the 2020 CDC FoodNet annual summary, highlighting the possible impact of *Salmonella* in chicken on human illnesses. Furthermore, serotypes I 4,[5],12:i:- and Hadar were ranked as the fifth and sixth, respectively. These serovars were also most common laboratory-diagnosed infections in the 2020 FoodNet annual summary. Both I 4,[5],12:i:- and Hadar have appeared periodically in chicken commodities since 2016. Serotypes observed at rehang and

post-chill throughout the 2022 FSIS Exploratory Sampling Program are described in section **3.5** below.

Table 22: Top serotypes detected in PR/HACCP sampling programs compared to top serotypes in CDC FoodNet's 2020 annual summary. Parenthetical numbers indicate sample count.

Rank	Carcass (2016-2021) (N=2,602)	Carcass (Apr-Oct 2022) (N=235)	Comminuted (2016-2021) (N=2,860)	Parts (2016-2021) (N=6,437)	2020 CDC FoodNet
1	Kentucky (1,288)	Kentucky (128)	Infantis (1,011)	Kentucky (1,844)	Enteritidis
2	Infantis (363)	Infantis (38)	Enteritidis (599)	Enteritidis (1,586)	Newport
3	Enteritidis (335)	Typhimurium (25)	Kentucky (507)	Infantis (1,230)	Javiana
4	Typhimurium (254)	Enteritidis (23)	Typhimurium (206)	Schwarzengrund (586)	Typhimurium
5	Schwarzengrund (133)	Schwarzengrund (5)	Schwarzengrund (122)	Typhimurium (499)	I 4,[5],12:i:-
6	Heidelberg (49)	Braenderup (2)	Thompson (72)	Heidelberg (183)	Hadar
7	I 4,[5],12:i- (30)	Blockley (2)	Johannesburg (57)	Thompson (91)	Infantis
8	Braenderup (23)	Anatum (2)	Braenderup (49)	Johannesburg (66)	Muenchen
9	Thompson (18)	Alachua (2)	Heidelberg (44)	I 4,[5],12:i- (46)	Saintpaul
10	Hadar (10)	I 4,[5],12:i- (1)	I 4,[5],12:i- (29)	Braenderup (46)	Mississippi

**Shaded cells indicate higher virulence cluster serotype.*

Serotype Clustering

As part of the Cooperative Agreement between FSIS and the UMD-JIFSAN, EpiX Analytics categorized *Salmonella* serotypes into two clusters derived from a machine learning algorithm using virulence factors to estimate the genetic similarity between serotypes. For a full description of the method developed by EpiX Analytics, see **Appendix A**. A summary is provided by FSIS in **Chapter 2**. The most frequent serotypes detected in chicken for each cluster are summarized in **Table 23**. Cluster 1, which includes serotypes Enteritidis and Typhimurium, is the higher virulence grouping of serotypes, as compared to cluster 2. Analysis of most probable number (MPN) and qPCR poultry product *Salmonella* enumeration data (results not shown) did not indicate a significant difference in level of *Salmonella* by cluster.

Table 23: Summary of the five most frequent *Salmonella* serotypes in chicken by cluster.

"Higher virulence" Cluster 1	"Lower virulence" Cluster 2
Enteritidis	Kentucky
Typhimurium	Infantis
I 4,[5], 12:I:-	Schwarzengrund
Hadar	Heidelberg
Litchfield	Thompson

The post-chill occurrence of each serocluster is stable across commodity and time (**Figure 15**). The average annual proportion of cluster 1 is approximately 0.26 for carcasses, 0.32 for comminuted product, and 0.35 for parts. Data from the 2022 FSIS Exploratory Sampling program indicate post-chill serocluster proportions as roughly 0.2 in cluster 1, and correspondingly, 0.8 in cluster 2.

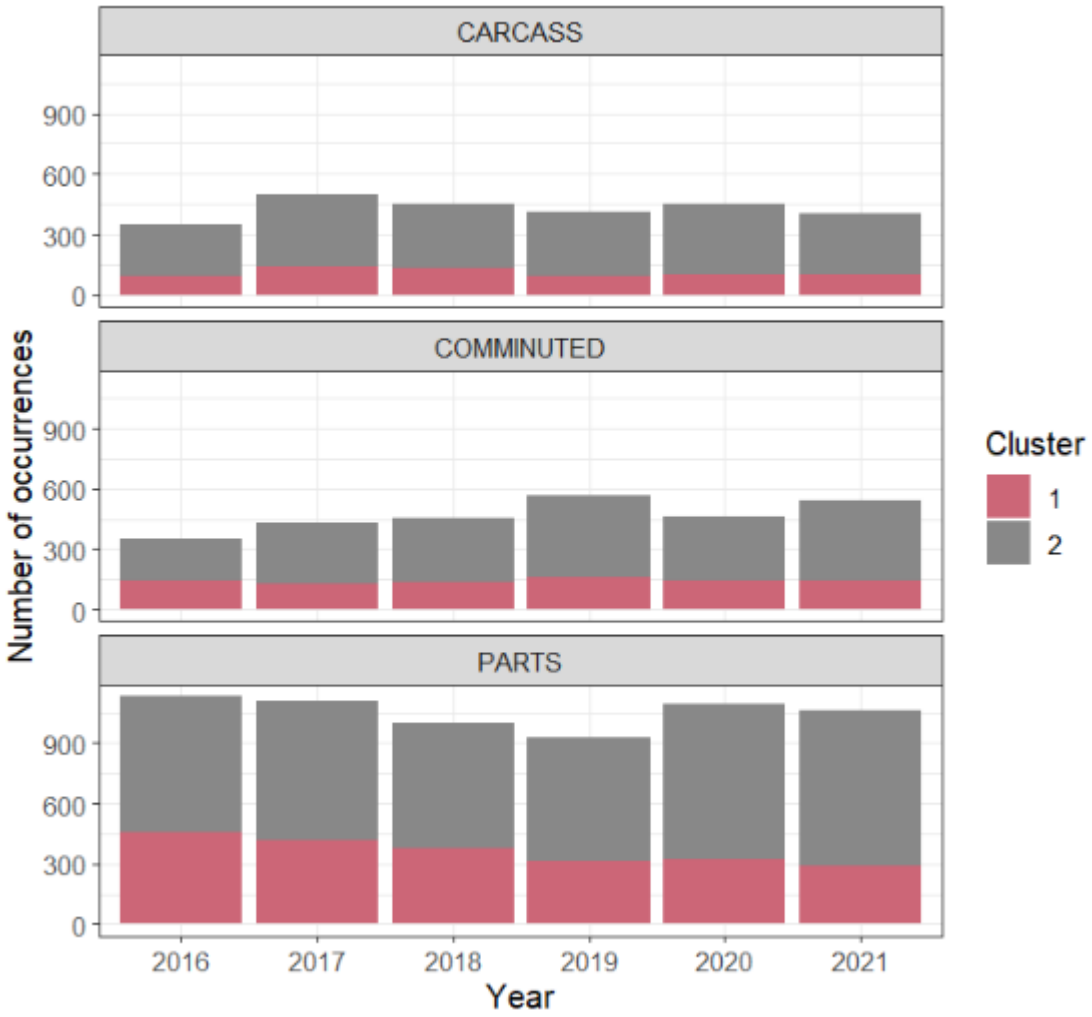


Figure 15: *Salmonella* seroclusters by poultry commodity across time.

3.4 *Salmonella* Contamination at Rehang

At present, FSIS lacks data on pathogen occurrence prior to rehang. A recent bio-mapping study (De Villena, 2022), two studies on the diversity of *Salmonella* serotypes (Altekruse, 2009b; Berrang, 2008; Berrang, 2009), and the within-flock contamination information (FSIS, 2022d) were combined to describe a reasonable depiction of *Salmonella* occurrence prior to rehang.

For the first analysis, data collected during a joint USDA, Agricultural Research Service (ARS) and FSIS study were used (Altekruse, 2009b; Berrang, 2008; Berrang, 2009). Sampling was conducted across 20 large chicken slaughter facilities in the U.S. At each facility, 40 carcass samples were collected at both

the rehang and post-chill locations of the slaughter process. To account for seasonal variations in *Salmonella* contamination, each establishment was visited quarterly. During each visit, 10 carcasses were collected at both rehang and post-chill using a 100 mL rinse sample. All samples belonged to a single flock or production lot. The same carcass was not sampled at each location because it was not feasible to identify individual carcasses after the chilling process. A total of 800 rehang and 798 post-chill samples were successfully analyzed (i.e., 2 post-chill samples could not be analyzed). No data on serotypes or levels of *Salmonella* are available.

Figure 16 shows the proportion of *Salmonella*-positive samples for all 80 flocks at rehang and post-chill. The average proportion of positives at rehang and post-chill was 0.71 and 0.21, respectively. At rehang, *Salmonella* was not isolated from 4 flocks and no establishment had more than one flock where no positives carcasses were found in a set of 10 samples. The sensitivity for detection of *Salmonella* within a flock with 10 samples provides approximately 95% confidence that the prevalence is less than 0.26, so 10 samples provide little confidence that the any of the four flocks were free of *Salmonella* contamination. Combining these results with the observation that the 100 mL rinse sample and 30 mL aliquot is only removing and testing roughly 4% of the *Salmonella* on the carcass, it is reasonable to assume that all flocks have some *Salmonella* contamination. A similarly high occurrence was also observed in a recent biomapping study of *Salmonella* in a single U.S. chicken slaughter establishment (De Villena, 2022).

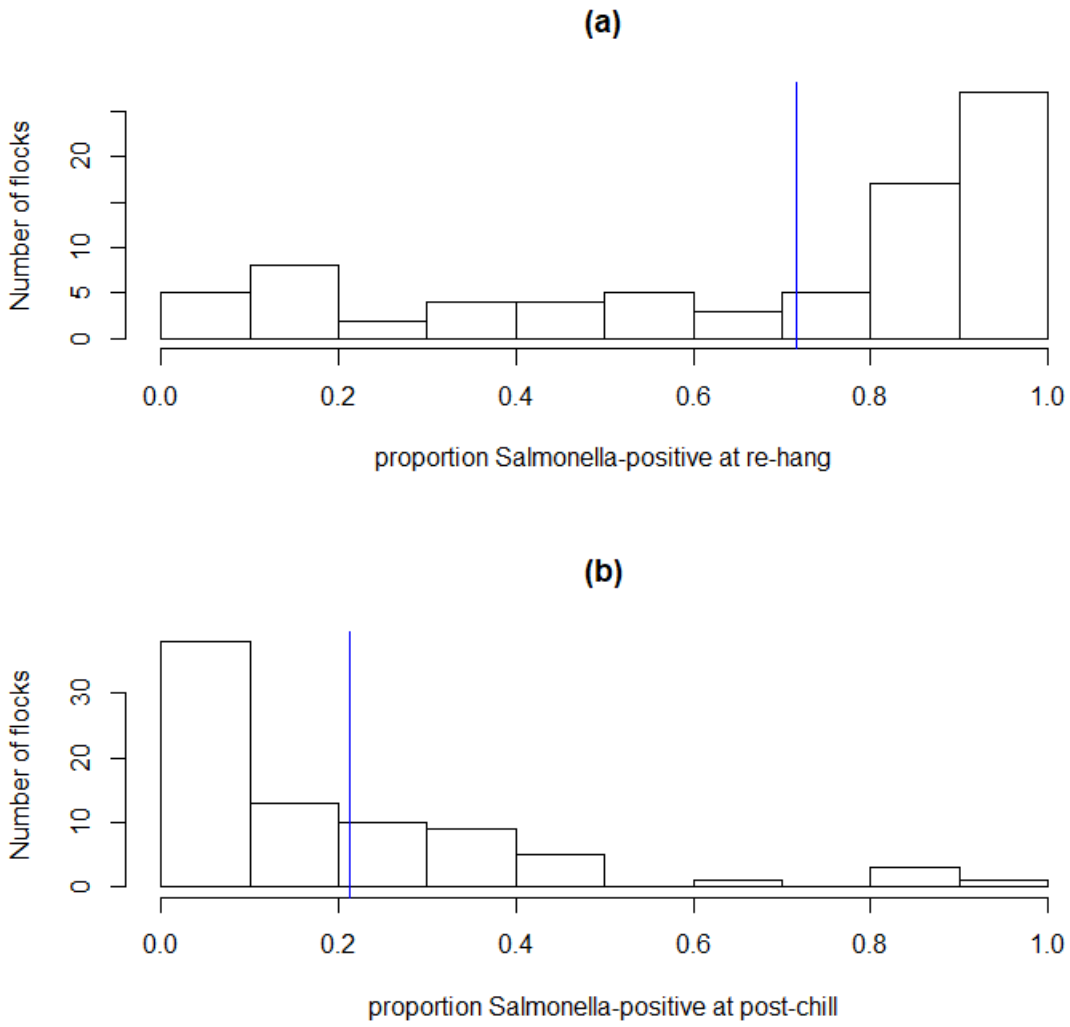


Figure 16: Proportion *of Salmonella* positive samples at rehang and post-chill from ARS-FSIS study.

3.5 Multiple Serotypes within Flocks

The standard practice for determining which *Salmonella* serotypes are present in a sample relies on a two-step process, where the first step is to enrich test-positive samples to grow individual colonies on selective media. The second step is to select a small number of these colonies to determine the serotype of the colony. On a limited number of occasions, FSIS has serotyped up to three colonies (FSIS, 1996a), but standard FSIS practice is to select a single colony. Using a single sample, it is possible to estimate the fraction of carcasses f_c with serotype c within an establishment or across the population. Nevertheless, lacking additional data, the only option when using a single colony is to assume that the identified serotype represents the dominant serovar in the sample.

Recent research has found that samples of individual chicken carcasses almost always contain multiple

Salmonella serotypes (Obe, 2023; Thompson, 2018). The proportion of any one serotype within a single sample varies from the sample contains a dominant (i.e., >90%) serotype to a mixture of multiple serotypes with none present in a fraction greater than 50%. A limitation of this risk assessment is the lack of data to characterize the degree to which the serotype identified in the sample represents a dominant serotype within the flock, or if the serotype assignment is a poor predictor of serotype composition of the flock because there isn't an overwhelmingly dominant serotype within the flock.

To estimate the degree to which flocks contain a dominant serotype, the following approach is used. It is assumed that if a flock contains a dominant serotype, then samples from a flock that is positive at both rehang and post-chill should have a high degree of agreement between the serotype identified at the two sampling locations. For example, FSIS considered a situation where serotypes are grouped into two clusters: higher virulence (cluster 1) and lower virulence (cluster 2). Additionally, we assumed the higher virulence class represents 20% of carcasses and 100 samples are collected across different flocks. At the most extreme case, we suppose flocks are only infected with a single serotype, as demonstrated in **Table 24**.

Table 24: Contingency table of serotypes at rehang to post-chill if only one serotype is possible.

		Post-chill	
		Serotype 1	Serotype 2
Rehang	Serotype 1	20	0
	Serotype 2	0	80

At the other extreme, assume that all flocks contain the same 20%-80% mixture of higher and lower virulence serotypes. In this case, the expected values for the contingency table are shown in **Table 25**.

Table 25: Contingency table of serotypes at rehang to post-chill all flocks have same 20%-80% mixture of serotypes.

		Post-chill	
		Serotype 1	Serotype 2
Rehang	Serotype 1	4	16
	Serotype 2	16	64

The FSIS 2022 exploratory paired carcass sampling used in this analysis contains a total of 168 paired samples that were positive at rehang and post-chill. **Figure 17** illustrates the correspondence of serotypes for each paired sample, with 62.5% (105) of sampling having matching serotypes at rehang and post-chill. Serotype Kentucky was detected most often as compared to other serotype (at rehang or post-chill), underscoring the abundance of Kentucky while also suggesting the likelihood that other

serovars are present but undetected. Serotypes Enteritidis and Typhimurium, which are considered more virulent, both match within a sample pair less than 50% of the time. This suggests that while these more virulent serotypes occur less frequently, they likely are present but undetected in a high percentage of all flocks.

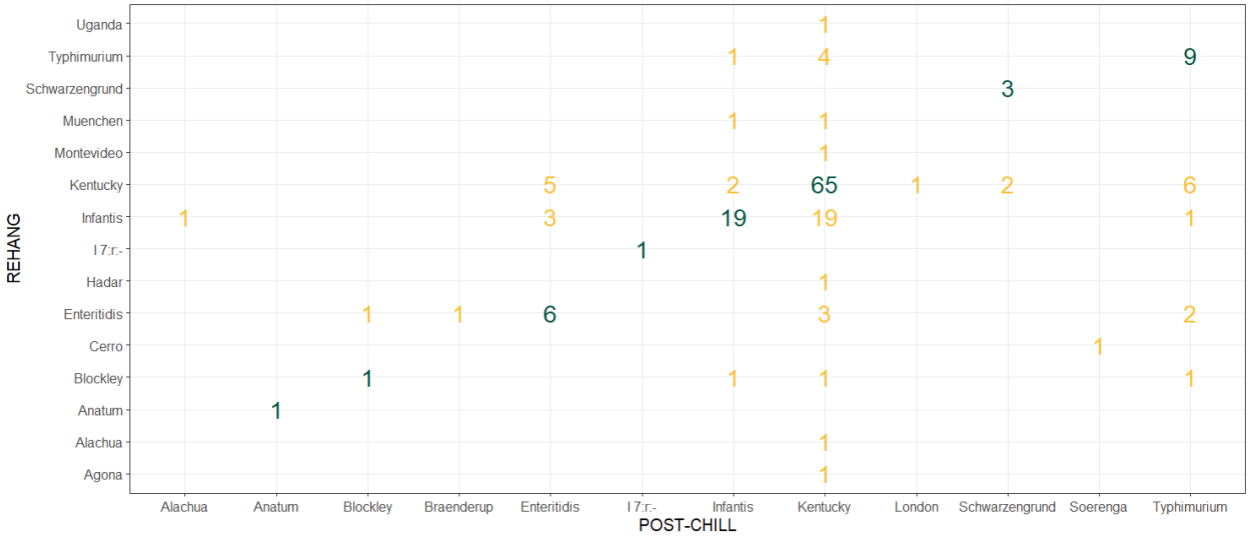


Figure 17: Correspondence of serotypes for each of the 168 paired FSIS 2022 exploratory carcass samples.

Table 26: Summary of paired *Salmonella* cluster results at rehang and post-chill: counts and frequencies.

		Post-chill		
		Cluster 1	Cluster 2	Total
Rehang	Cluster 1	20	14	34
	Cluster 2	15	119	134
	Total	35	133	168

		Post-chill		
		Cluster 1	Cluster 2	Total
Rehang	Cluster 1	0.119	0.089	0.202
	Cluster 2	0.083	0.708	0.798
	Total	0.208	0.792	1

Summarizing rehang and post-chill results by the two *Salmonella* serotype clusters for positive paired results (**Table 26**), we observe following:

1. the frequency of clusters 1 and 2 are nearly the same at both locations (cluster 1 at rehang: 0.202, cluster 1 at post-chill:0.208; cluster 2 at rehang: 0.798, cluster 2 at post-chill: .792)
2. a number of samples are discordant pairs (opposite clusters are found at rehang and post-chill (cluster 1 at rehang, but cluster 2 at post-chill: 0.089, cluster 2 at rehang, but cluster 1 at post-chill: 0.083) and,
3. by comparing to the "chance alone" contingency **Table 25** observe that the frequency of pairs with cluster 1 at both locations is larger than expected by chance alone (for a more detailed analysis see section **5.1**).

The first finding suggests that interventions between rehang and post-chill do not disproportionately influence the relative frequency of the clusters. The second finding suggests that both clusters do occur in sampled lots. The third finding suggests some positive correlation in the results between rehang and post-chill. The proportion of samples where a lower virulence serotype was isolated at both rehang and post-chill is similar to the expected value if all flocks had an identical proportion of lower and higher virulence serovars (i.e., 0.71 versus 0.63), whereas the result of the higher virulence clusters represent a greater departure (i.e., 0.12 versus 0.04). This suggests there might be a higher degree of clustering for the higher virulence serotypes.

Chapter 4 Baseline Exposure Assessment

4.1 Foodborne illness surveillance

Foodborne illness surveillance in the U.S. relies on a broad network of local and state health departments and the CDC. The CDC FoodNet conducts surveillance for nine laboratory-diagnosed infections, including, *Salmonella*, identified by culture or culture-independent diagnostic test for bacterial pathogens of samples from patients. The network was established in July 1995 and is a collaborative program among CDC, 10 state health departments, FSIS, and the Food and Drug Administration (FDA). The surveillance area includes 15% of the U.S. population (48 million persons). Personnel at each FoodNet site collect information about cases of infection and share that information with CDC through FoodNet's database.

CDC NORS includes data on illnesses resulting from contact with animals, environmental contamination, spread by person-to-person, waterborne transmission, and other enteric illness outbreaks. CDC also maintains the Foodborne Disease Outbreak Surveillance System (FDOSS) for collecting and reporting data about foodborne disease outbreaks in the U.S. In FDOSS, outbreaks are defined as the occurrence of $\geq$2 cases of a similar illness resulting from the ingestion of a common food (Gould, 2013). NORS data provide detailed food items considered as vehicles of the outbreaks and are more reliable to determine the causative contaminated food vehicles. Each of the implicated food vehicles has been grouped into one of 17 broad commodity classes (Painter, 2013a; Richardson, 2017)

Foodborne illness outbreaks attributed to *Salmonella*-contaminated foods provide the most robust data source available for the attribution of illnesses to different commodities due to the large number of outbreaks, relative to the other foodborne bacterial pathogens (IFSAC, 2022), the occurrence of *Salmonella* outbreaks across all 17 commodity classes (Painter, 2009; Richardson, 2017), and the general similarity between the characteristics of sporadic cases identified through laboratory surveillance and outbreak cases (Ebel, 2016).

Foodborne illness source attribution is the process of identifying which foods are the most important sources of selected major foodborne illnesses.

IFSAC produces annual estimates for *Salmonella*, among other pathogens. The annual estimates utilize a published statistical modeling approach to mitigate the influence of large outbreaks that might bias estimates. This approach also incorporates epidemiologic factors relevant to outbreak size, weights recent outbreaks more heavily than older ones, and quantifies uncertainty by estimating credibility intervals around estimates (Batz, 2021).

The implicated foods were divided into 17 categories for the analysis, and the method gives the greatest weight to the most recent five years of outbreak data (2016–2020). In the 2020 report, 17.3% of *Salmonella* illnesses were attributed to chicken.

Overall Burden of Salmonella Illnesses

Surveillance systems and surveys provide vital information about the burden of foodborne illness in the U.S., but they do not capture *every* illness. Because only a fraction of illnesses are diagnosed and reported, periodic assessments of the total burden of illness are required. CDC developed an approach to estimate the total number of foodborne illnesses from *Salmonella* and other priority pathogens

(Scallan, 2011). This approach utilizes data from CDC FoodNet and other surveillance databases and corrects for underreporting and under-diagnosis. The adjusted number is multiplied by the proportion of illnesses acquired in the U.S. (that is, not during international travel) and the proportion transmitted by food to yield an estimated number of illnesses that are domestically acquired and foodborne (Beshearse, 2021).

In recent years, CDC has worked to develop updated estimates of the burden of foodborne illness. As a part of this effort, new analyses have been conducted to revisit the multiplier used by CDC to determine the percent of *Salmonella* illnesses that are foodborne in nature. In Scallan (2011), an estimate of 94% was utilized, which was derived from based on FoodNet case-control study of sporadic illness and on outbreaks reported to the CDC from 1996-2006 (Mermin, 2004; Scallan, 2011). More recently, CDC conducted a structured expert judgement (SEJ) (Beshearse, 2021) to revisit the estimate of percent foodborne for *Salmonella* and many other pathogens. In this SEJ, the authors looked holistically at multiple pathways, including foodborne, waterborne, person-to-person, and animal contact. Based on this work, the authors determined that the percent of all *Salmonella* that were foodborne in nature was 66%. As such, this risk assessment utilizes this 66% foodborne estimates in its calculations of the total number of *Salmonella* illnesses prevented from the various risk management options.

4.2 Chicken Consumption

Data on the consumption of chicken in the U.S. were obtained from the NHANES. The NHANES program suspended field operations in March 2020 due to the 2019 coronavirus disease (COVID-19) pandemic. As a result, data collection for the NHANES 2019-2020 cycle was not completed and the collected data are not nationally representative. Therefore, data collected from 2019 to March 2020 were combined with data from the NHANES 2017-2018 cycle to form a nationally representative sample of NHANES 2017-March 2020 pre-pandemic data.

All NHANES participants are eligible for two 24-hour dietary recall interviews. The first dietary recall interview is collected in-person in the Mobile Examination Center and the second interview is collected by telephone 3 to 10 days later. In the 2017-March 2020 pre-pandemic sample, 12,634 participants provided complete dietary intakes for Day 1. Of those providing the Day 1 data, 10,830 provided complete dietary intakes for Day 2. The NHANES Day 1 weights, adjusted for non-response and daily variability, were used in the analysis (CDC, 2022b). **Table 27** summarizes the mean serving sizes of the chicken products used in this risk assessment. Additional details are in **Appendix B**.

Table 27: Mean serving size (g) for chicken overall, chicken parts, and comminuted chicken.

Chicken Product	Mean Serving Size (g)
Chicken (overall)	139.0
Parts	135.7
Comminuted	110.0

4.3 Empirical Baseline Probability of Illness

Using FSIS data, the size and composition of the chicken parts and comminuted industries is summarized in **Table 28**. The comminuted product industry represents smaller numbers of establishments, while the

chicken parts industry is more than twice the size of the chicken carcass industry, which slaughters all live birds and supplies the parts and comminuted industries.

The total number of units produced by each industry are affected by risk management decisions. For this analysis, FSIS defines flocks as susceptible to diversion for the chicken carcass industry, while an entire day of production is defined as susceptible to diversion for the parts and comminuted industries. In 2021, FSIS collected about one sample per week from larger establishments while smaller establishments were sampled less frequently.

It is estimated there are 125,115 chicken-associated *Salmonella* illnesses per year. This value is calculated as the product of the total number of CDC FoodNet cases per year (7,600), the share of these cases that are foodborne (66 percent) and of domestic origin (89 percent), the under-diagnosis multiplier for *Salmonella* (24.3)(Ebel, 2012c), dividing by the FoodNet catchment area (15 percent) and multiplying by the portion IFSAC attribution to chicken (17.3 percent). As previously discussed, these total cases are distributed across products by assuming the proportion of servings consumed (0.11, 0.83 and 0.06) is proportional to illnesses resulting from exposure to carcasses (whole chickens), parts and comminuted (ground) forms of chicken, respectively.

Using FSIS data, the total number of chicken carcasses produced in 2021 is 9.4 billion, with some 16% exported (NCC, 2021). Using USDA Economic Research Service (USDA-ERS) estimates for retail boneless chicken consumed in the U.S. (~2.4 pounds per carcass produced) and NHANES estimates for the average serving size for chicken (139 gram), we estimate 61 billion servings of chicken—in all its forms—are consumed in the U.S. each year. The ratio of total chicken illnesses to total chicken servings (2 per million) provides an empirical estimate of the risk of illness per serving (Hsi, 2015).

Table 28: Relevant parameters for the assessed products' industries.

Parameter	Chicken carcasses	Chicken parts	Comminuted chicken
No. establishments	206	484	74
No. units/year	203,718 lots	104,505 days	8,545 days
No. samples/year	9,635	14,192	1,991
Illnesses/year	13,763	103,845	7,507

4.4 Descriptive Estimates of Risk per Serving

Adequately answering the risk management questions necessitated the use of the virulence-adjusted dose-response models; the development of which is outlined in **Appendix A**. These models provides a description of risk of illness per serving for poultry products, beyond the empirical estimate described above. That description, summarized in **Table 29**, can be informative to risk managers, but was not the ultimate goal of the dose-response models development.

The scenarios in **Table 29** summarize the average initial concentration of FSIS-sampled product that

exceeds different concentration thresholds (note that all the average initial concentrations are cfu/g, but thresholds are cfu/mL for carcasses and parts). Following multiplication of this average initial concentration by the attenuation distribution, we can calculate the average dose per serving and integrate each dose-response function across the resulting distribution to calculate probabilities of illness per serving. We also predict the likelihood that lots will fail the different concentration thresholds.

The average initial contamination concentrations above a threshold and the average doses per serving increase as the concentration thresholds increase, but in a non-linear pattern. For example, the average contamination concentration at the 1, 10 and 100 cfu/mL thresholds for chicken carcasses is approximately 16, 9.7 and 6.8 times larger than the threshold, respectively. This declining relationship reflects the reduced likelihood of larger values in the right tail of the initial contamination distributions.

In contrast, the ratio of average dose per serving to average initial contamination concentration is the same for each concentration threshold (e.g., for comminuted product at the 100 cfu/g threshold, the ratio is 408/2572). This ratio, 0.16, is the expected value of the attenuation distribution ($e^{-5\times ln(10)+0.5\times(1.91\times ln(10))^2}$) that modifies the initial contamination value to account for the effects of mixing, partitioning, growth, attenuation (e.g., cooking) and serving size between production and consumption. It is notable that this expected value of the attenuation distribution represents the 98th percentile (approximately) of that extremely skewed distribution. For comparison, the median, 95th, 99th and 99.9th percentiles of the attenuation distribution are 0.00001 (i.e., 10^{-5}), 0.014, 0.28, and 8, respectively.

Similarly, the probabilities of illness increase with the concentration threshold. The increases are not linear because the average initial contamination above the threshold is not changing in a linear pattern and the dose-response functions are non-linear (particularly at doses above 1 cfu/g).

Table 29*:* Risk of illness per serving of poultry product based on the initial concentration of *Salmonella* in FSIS-sampled products.

		Initial concentration threshold (cfu/mL or g)				
Measurement	**Product type**	**0.003**	**0.033**	**1**	**10**	**100**
Avg initial concentration for failing lots (cfu/g)	Carcasses	0.48	1.65	16	97	682
	Parts	0.08	0.30	4	33	281
	Comminuted	17	37	163	582	2,572
Average dose consumed for average failing lot (cfu/serving)	Carcasses	0.08	0.26	3	15	108
	Parts	0.01	0.05	0.67	5	45
	Comminuted	3	6	26	92	408
Probability of illness per million servings*, higher virulence	Carcasses	230	539	2,243	6,124	15,980
	Parts	60	164	995	3,398	10,504
	Comminuted	2,353	3,632	8,000	14,859	28,487
Probability of illness per million servings*, lower virulence	Carcasses	39	95	417	1,193	3,292
	Parts	10	28	179	643	2,109
	Comminuted	438	690	1,581	3,047	6,113
Likelihood of Consumer Exposure to Raw Product at or above Initial Concentration	Carcasses	11%	3%	0.27%	0.03%	0.00%
	Parts	31%	7%	0.17%	0.00%	0.00%
	Comminuted	27%	13%	3%	1%	0.17%

* Given average initial concentration multiplied by attenuation distribution

Chapter 5 Final Product Standards

The second risk management questions states:

> What is the public health impact (change in illnesses, hospitalizations, and deaths) achieved by eliminating final product contaminated with specific levels of *Salmonella* and/or specific *Salmonella* subtypes?

Based on input from FSIS risk managers, the model interpretation of this risk management question is as an analysis of a test-and-hold final product standard based on specific levels of *Salmonella* and *Salmonella* subtype scenarios. While it can be postulated that any regulation that declares a pathogen an adulterant will have an indirect effect on the chicken industry's pathogen control measures, no data is available at this time describing the magnitude of that effect. As such, only the direct public health effects of the removal contaminated lots are modeled.

A major assumption of this modeling approach is that consumer demand for raw chicken products will be met by the industry, so every removed lot will be replaced by another lot overall. FSIS believes this assumption is reasonable because of the high consumer demand for prepared chicken products. For example, of the more than 9 billion broiler chickens produced annually in the U.S. (USDA-ERS, 2021), about 1 billion are thought to be consumed as whole carcasses. Of these 1 billion whole chickens, roughly 625 million are purchased as ready-to-eat (RTE) rotisserie chickens (Gasparro, 2018). Similarly, for parts and comminuted product, much of the increase in chicken consumption is attributed to meals prepared outside the home and is attributed to the consumption of chicken nuggets, chicken strips, and grilled chicken sandwiches and their rising popularity in fast food, restaurant, and institutional settings (USDA-ERS, 2017) where product is increasingly likely to be RTE. Given the increasing demand for these RTE forms of various chicken products (Research And Markets, 2023), it is reasonable to assume that every removed lot can be replaced by another lot to meet the demand for RTE chicken. However, the validity of this approach does limit the potential for additional indirect effects of the policy because given the small fraction of lots that are tested, it is also plausible that some producers will simply divert all tested lots to the RTE market, regardless of the final test result, rather than incur the cost and inconvenience of holding the product or investing in new technologies to mitigate contamination.

While this approach differs from other modeling approaches described in the scientific literature (Lambertini, 2019; Lambertini, 2021; Oscar, 2021), FSIS thinks this approach represents a more realistic assessment of the current chicken industry and, therefore, the identified public health benefits.

A key driver for the model approach outlined below are the two dose-response *Salmonella* models developed by EpiX Analytics for this risk assessment (**Chapter 2**, **Appendix A**). The stratification of *Salmonella* serotypes into higher and lower virulence clusters with a dose-response model fit to each cluster's epidemiological data, and the mixture of multiple serotypes in chicken flocks (section **3.5**) necessitates consideration of levels and serotypes in all of the analyzed scenarios. The model's handling of serotype standards is based on the available FSIS two-point chicken carcass data and, as such, cannot be used for parts and comminuted product where two-point data is not available.

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, and the testing method used for risk management option implementation should be considered when evaluating the results below, as discussed in the NACMCF 2023 response (NACMCF, 2023).

5.1 Modeling Approach

Hazard Characterization for Final Product Standards and Receiving Guidelines: Dose-Response Model

Two clusters of *Salmonella* serotypes were defined (see **Appendix A**) and dose-response functions were developed for each cluster. The first cluster consists, generally, of the more virulent *Salmonella* serotypes; in the following analysis we denote this cluster 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 denoted C2.

To develop the dose-response models, a lognormal distribution (*Log10Normal*(-3.037117, 1.279985)) was used that reflected the initial contamination of a mixture of the three raw chicken products – carcasses, parts and comminuted – according to their relative frequencies of consumption (see subsection **Chicken Consumption**). An attenuation distribution that encompassed all the effects of partitioning, mixing, growth, and attenuation that can occur between production of raw chicken and consumption of chicken servings was also defined (Log10Normal(-5,1.91) (Ebel, 2015)), as described above. Combining the initial contamination and attenuation distributions constituted an (log10) exposure distribution.

The derivation of the EpiX Analytics-developed dose-response model parameters for clusters 1 and 2 depends on maintaining the following relationship:

$$\frac{\int R_1(d)h(d)\partial d}{\int R_2(d)h(d)\partial d} = \frac{RR_1}{RR_2}.$$

The left side of this equation is a ratio of the outputs of integrating the dose-response functions for cluster 1 ($R_1(d)$) and cluster 2 ($R_2(d)$) across the exposure distribution, $h(d)$. The outputs of these integrals can be interpreted as the overall probability of illness per serving given the dose-response function. Importantly, this assumes that exposure distributions do not differ between the seroclusters.

The right side of this equation is a ratio of relative risk terms. The numerator of this ratio RR_1 is the expected increased probability of illness given exposure to C1 serotypes. This is estimated as the ratio of the fraction of outbreak illnesses attributed to C1 serotypes (e.g., 71 percent) to the fraction of chicken isolates determined to be C1 serotypes (e.g., 33 percent); this latter term serves as a proxy for the relative exposure probability. For example, the equation, $RR_1 = 2.15$ implies that a C1 exposure increases risk of illness 2.1 times some baseline risk. Similar reasoning for the denominator RR_2. concludes that the fraction of outbreak illnesses attributed to C2 serotypes (e.g., 25 percent) as compared to the fraction of chicken isolates determined to be C2 serotypes (e.g., 66 percent) is 0.38, or that a C2 exposure reduces risk of illness $1/0.38 = 2.63$ times some baseline risk. As explained in **Appendix A**, substantial uncertainty attends the estimation of these relative risk terms.

The ratio $\frac{RR_1}{RR_2} = \frac{2.15}{0.38} = 5.66$ indicates that the probability of illness per serving from C1 exposures is 5.66 times larger than the probability of illness per serving from C2 exposures. Therefore, the parameters for the two dose-response function must be selected to maintain this relative probability of illness.

e parameters of $R_1(d)$ and $R_2(d)$ are estimated using numerical techniques based on a simplifying assumption that, although the mean of the beta distribution underlying their beta-Poisson dose-response model differs between C1 and C2, the sum of those beta parameters must be equal. Given the complexity of this model – which uses a 2F1 hypergeometric confluent function of the second kind – the calculation of each integral is simplified using a polynomial expression such that, for example,

$$\int R_1(d)h(d)\partial d \approx \sum_d coef1_1 \times \ln(d_i+1) + coef2_1 \times \ln(d_i+1)^2 + \ldots + coef9_1 \times \ln(d_i+1)^9 \ .$$

Although the dose-response functions are developed based on an exposure distribution that encompasses all exposures to all chicken products, this risk assessment is concerned with distinguishing between the probability of illness from exposures to units that pass or fail standards imposed by FSIS on individual forms of the poultry products (e.g., chicken carcasses, parts, or comminuted chicken). Therefore, the dose-response functions are applied as described below in subsection **Using the Dose-Response Model**.

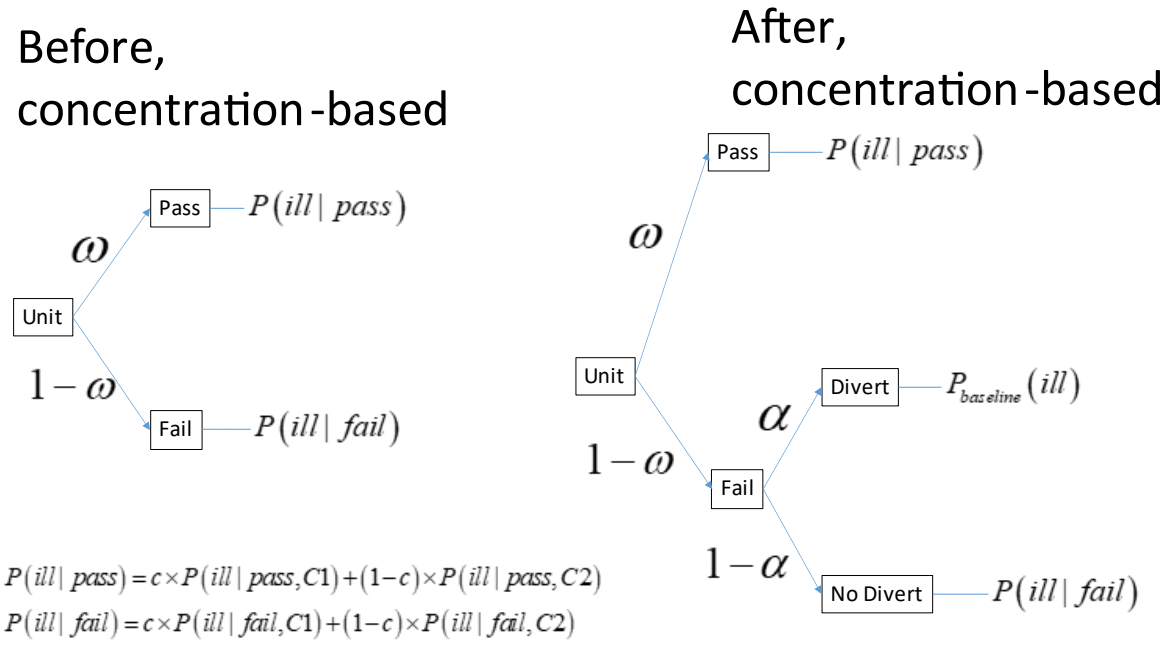


Figure 18: Schematic depiction of the possible pathways which product moves through before and after implementation of a concentration-based (i.e., level-based) diversion strategy.

Describing Serotype Mixtures in Flocks

As there are two serocluster specific dose-response curves, before any prediction of illnesses prevented can be made the mixture of C1 and C2 serotypes in chicken flocks is described.

Table 30: Summary of paired *Salmonella* cluster results at rehang and post-chill: counts (top) and frequencies (below).

		Post-chill		
		Cluster 1	Cluster 2	Total
Rehang	Cluster 1	20	14	34
	Cluster 2	15	119	134
	Total	35	133	168

		Post-chill		
		Cluster 1	Cluster 2	Total
Rehang	Cluster 1	0.119	0.089	0.202
	Cluster 2	0.083	0.708	0.798
	Total	0.208	0.792	1

The kappa statistic for the **Table 30** data is $\kappa = 0.47$; which indicates weak to moderate correlation (McHugh, 2012). Squaring the correlation coefficient to obtain the coefficient of determination $\left(\kappa^2 = 0.22\right)$. provides an estimate of the amount of variation in the dependent variable (the post-chill serotype) that can be explained by the independent variable (the serotype observed at rehang). Since both the post-chill and rehang serotype results are informative, there is merit to serotype-based risk management options, and more so to ones that rely on a paired serotype at rehang. In principle, both have the potential for additional reductions in illness.

ven these results, we propose a model that could explain these findings. We assume there are two types of sampled flocks that differ with respect to the frequency of each cluster (denoted Schemes 1 and 2). Furthermore, a parameter describes the frequency of Scheme 1 ($P\left(Scheme1\right)$) while the frequency of Scheme 2 is the complement of this parameter ($P\left(Scheme2\right) = 1 - P\left(Scheme1\right)$) . . Simplifying the notation (
Table 31), we can predict the frequency of paired sample outcomes. For example, the probability of getting cluster 1 results at both rehang and post-chill is the probability of a cluster 1 result in Scheme 1

multiplied by itself and the probability of Scheme 1 ($P(C1|S1)P(C1|S1)P(S1)$) plus the same cculation for Scheme 2.

Using a fitting algorithm that selects $P(C1|S1), P(C1|S2)$ and $P(S1)$. such that the difference is minimized between the observed frequency and predicted frequency for concordant pairs, we find $P(C1|S1)=0.875$, $P(C1|S2)=0.09$. and $P(S1)=0.146$.

Table 31: Model calculations to predict the frequencies in **Table 30**.

		Post-chill	
		Cluster 1	Cluster 2
Rehang	Cluster 1	$P(C1\|S1)P(C1\|S1)P(S1)+$ $P(C1\|S2)P(C1\|S2)P(S2)$	$P(C1\|S1)P(C2\|S1)P(S1)+$. $P(C1\|S2)P(C2\|S2)P(S2)$
	Cluster 2	$P(C1\|S1)P(C2\|S1)P(S1)+$ $P(C1\|S2)P(C2\|S2)P(S2)$	$P(C2\|S1)P(C2\|S1)P(S1)+$ $P(C2\|S2)P(C2\|S2)P(S2)$

These results suggest that the observed results can be predicted by assuming there are two types of flocks (i.e., schemes) with different mixtures of the clusters. The probability of a cluster 1 result is large for Scheme 1, and it is small for Scheme 2. Given this combination of serotype mixtures, we can assess what is the probability that a result came from Scheme 1.

To start, consider a situation where we test a flock at rehang and post-chill and find cluster 1 in each sample. We want to know the probability that the flock is from Scheme 1 (P(S1|C1,C1), the probability of Scheme 1 given rehang and post-chill results are cluster 1). From Bayes Theorem, we know

$P(S1|C1,C1)=\frac{P(C1,C1|S1)P(S1)}{P(C1,C1|S2)P(S2)+P(C1,C1|S1)P(S1)}$ =0.94. Therefore, we have high confidence that such a result came from a flock in Scheme 1. This finding suggests that a flock that tests positive for cluster 1 is very likely to have a high frequency of cluster 1 contamination.

What can we say about a single post-chill sample that tests positive for cluster 1? In this case, we calculate $P(S1 \mid C1) = \frac{P(C1 \mid S1)P(S1)}{P(C1 \mid S2)P(S2) + P(C1 \mid S1)P(S1)}$ =0.62. Therefore, we are less confident that such a result came from a flock in Scheme 1. Nevertheless, given the substantial differences in frequency of cluster 1 between the schemes, there is still a higher probability of Scheme 1 when we get a cluster 1 result than the underlying probability of that scheme (i.e., 0.146). These values are summarized in **Table 32**.

We can use these calculations to address the public health effects of diverting flocks that test positive for cluster 1 at post-chill with or without concordant results at rehang.

Table 32: Summary table of per flock serotype mixture probabilities.

Symbol	Definition	Value
$\boldsymbol{P}(\boldsymbol{C}1 \mid \boldsymbol{S}1)$	probability of a cluster 1 result in Scheme 1	0.875
$\boldsymbol{P}(\boldsymbol{C}1 \mid \boldsymbol{S}2)$	probability of a cluster 1 result in Scheme 2	0.09
$\boldsymbol{P}(\boldsymbol{S}1).$	probability of Scheme 1	0.146
	probability that the flock is from Scheme 1, given both rehang and post-chill test is cluster 1 positive	0.94
$\boldsymbol{P}(\boldsymbol{S}1 \mid \boldsymbol{C}1)$	probability that the flock is from Scheme 1, given the post-chill test is cluster 1 positive	0.62

Concentration-based Final Product Standard

This approach begins with the assumption that following equation predicts the distribution of illnesses prevented by a risk management option that changes the probability of illness from some baseline (Ebel, 2012a; Ebel, 2015; Williams, 2011a).

$$I_{avoided} = Poisson\left(\left(1 - \frac{P_{new}(ill)}{P_{baseline}(ill)}\right)\lambda_{ill}\right).$$

Next, we assume the baseline probability of illness is a function of units (lots/flocks → servings) that are determined to pass some level criterion (ω), as determined by the risk characterization (below).

$$P_{baseline}(ill) = \omega \times P(ill \mid pass) + (1-\omega) \times P(ill \mid fail)$$

We further assume the risk management option identifies some fraction of the units that fail the criterion (α) and, essentially, replaces them with random untested units. Then, the new probability of iness is

$$P_{new}(ill) = \omega \times P(ill \mid pass) + (1-\omega)\left[\alpha \times P_{baseline}(ill) + (1-\alpha) \times P(ill \mid fail)\right]$$

Using the Dose-Response Model

For units that pass (or fail) the criterion, the probability of illness depends on probabilities of illness based on one of two dose-response functions (i.e., R(d) for cluster 1 or 2; C1 or C2).

$$P(ill|pass) = c \times P(ill|pass, C1) + (1-c) \times P(ill|pass, C2)$$

$$P(ill \mid fail) = c \times P(ill \mid fail, C1) + (1-c) \times P(ill \mid fail, C2)$$

where c is the proportion of *Salmonella* that is in cluster 1. The default value for c is 0.2, keeping with the proportion observed in the 2022 FSIS Exploratory Sampling program.

In assessing exposure, we begin with an initial level $x \sim lognormal(\mu_x, \sigma_x)$ and an attenuation factor $a \sim lognormal(\mu_a, \sigma_a)$. Then the dose at consumption, (i.e., exposure distribution) is $d = x \times a$.

A failing unit is defined as having an initial level greater than or equal to some threshold level T (i.e., $x \geq T$). To determine the probability of illness, (i.e., risk characterization) for a passing or failing unit, conditioned on cluster type, we solve the following:

$$P(ill|pass, Cj) = \int_{x<T} R_j(d)\, h(d) \partial d$$

$$P(ill \mid fail, Cj) = \int_{x \geq T} R_j(d) h(d) \partial d$$

Actual solutions of the model proceed as follows.

1. For a particular product, solve $P_{baseline}(ill) = c \times \int R_1(d) h(d) \partial d + (1-c) \times \int R_2(d) h(d) \partial d$. In other words, the probability of illness per serving across all exposures is the weighted average of the probability of illness per serving across all exposures to clusters 1 and 2. This step can be accomplished using numerical integration because the exposure distribution, in log10, is simply the sum of two normal distributions (i.e., initial contamination and attenuation).
2. Using Monte Carlo simulation, sample from a truncated form of x, where its minimum is defined as negative infinity and its maximum is $log\ 10\ (T)$. Multiply this vector of initial levels less than the threshold by a vector of samples from the attenuation distribution to simulate exposure doses from passing units.
3. Use the simulated exposures for passing units to estimate $P(ill \mid pass) = c \times P(ill \mid pass, C1) + (1-c) \times P(ill \mid pass, C2)$..

4. e the components of steps 1-3 to solve for the probabilities of illness for failing exposures. For example,
   $\int R_1(d)h(d)\partial d = \omega \times P(ill \mid pass, C1) + (1-\omega) \times P(ill \mid fail, C1)$ i.e., the probability of illness from all C1 exposures is the weighted average of passing and failing exposures. Such an expression can be solved for $P(ill \mid fail, C1/C2)$.
5. The fraction of units passing, ω, is determined as the cumulative probability that $x < T$.
6. The parameter α is the fraction of failing units that are diverted and replaced by random units. If L is the number of units produced per year (e.g., production lots) and n is the total number of failing units tested per year, then $\alpha = \frac{n}{L}$ (i.e., if all units are tested, then all failing units will be diverted).

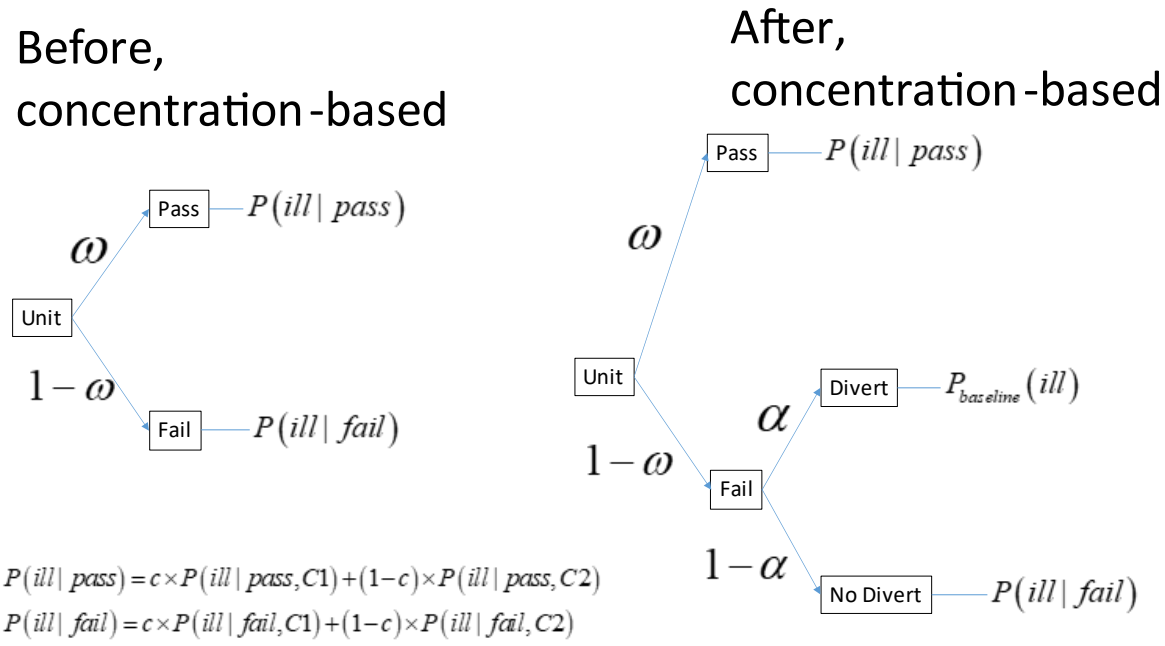


Figure 19: Schematic depiction of the possible pathways which product moves through before and after implementation of a concentration-based final product testing standard.

Serotype-based Final Product Standard

This approach begins with the assumption that there are two seroclusters (C1 and C2), but a tested unit will only be diverted if C1 is detected. If we detect and divert a C1 unit, a random unit replaces it.

In this construction, the baseline probability of illness per serving is

$$(1-\omega) \times (1-c) \times \left[k \times P(ill \mid fail, C1) + (1-k) \times P(ill \mid fail, C2) \right]$$

where h is the share of C1 *Salmonella* among failing units in which C1 was detected at post-chill, k is the share of C1 *Salmonella* among failing units in which C2 was detected at post-chill (explained below) and $P_{baseline.sero}(ill) \approx P_{baseline.}(ill)$.

We then assume the risk management option identifies some fraction of the units that fail the criterion and C1 is detected (α) and, essentially, replaces them with random untested units. Then, the new probability of illness is

$$P_{new.sero}\left(ill\right) = \omega \times P\left(ill \mid pass\right) + \left(1-\omega\right) \times c \times \alpha \times P_{baseline.sero}\left(ill\right) +$$
$$\left(1-\omega\right) \times c \times \left(1-\alpha\right)\left[h \times P\left(ill \mid fail, C1\right) + \left(1-h\right) \times P\left(ill \mid fail, C2\right)\right] +$$
$$\left(1-\omega\right) \times \left(1-c\right) \times \left[k \times P\left(ill \mid fail, C1\right) + \left(1-k\right) \times P\left(ill \mid fail, C2\right)\right]$$

We predict the annual illnesses prevented by this option as;

$$I_{avoided.sero} = Poisson\left(\left(1 - \frac{P_{new.sero}\left(ill\right)}{P_{baseline.sero}\left(ill\right)}\right)\lambda_{ill}\right)$$

To derive h and k, we assume there are two basic schemes among units that describe the frequencies of C1 among the *Salmonella* within the units. Scheme 1 has a relatively high frequency of C1 (and a complementary low frequency of C2) while Scheme 2 has a low frequency of C1 (and a complementary high frequency of C2).

As seen above, we described the estimation of the frequencies of C1 and C2 by scheme as;

Table 33: Description of the two derived mixtures in flocks of cluster 1 and cluster 2 *Salmonella*.

	Cluster 1	Cluster 2
Scheme 1	0.88	0.12
Scheme 2	0.09	0.91

For notation, we define the share of C1 in Scheme 1 as $S(C1, S1) = 0.88$. Similarly, we define the share of C1 in Scheme 2 as $S(C1, S2) = 0.09$.

Additionally, we illustrated previously the calculations for probabilities of each scheme based conditioned on whether a post-chill sample found C1 or C2. These probabilities are;

$$P(S1|C1) = 0.62 \text{ and } P(S1|C2) = 0.02$$

Using these estimates, we define $h = P\left(S1 \mid C1\right) \times S\left(C1, S1\right) + \left(1 - P\left(S1 \mid C1\right)\right) \times S\left(C1, S2\right)$. This parameter is the share of C1 *Salmonella* given that we observed C1 at post-chill; it is the weighted sum of the shares of C1 for each scheme where the weights are the conditional probabilities of each scheme given C1 was found. The complement of h $(i.e., 1-h)$ describes the share of C2 *Salmonella* given that

we observed C1 at post-chill.

Similarly, we define $k = P(S1 \mid C2) \times S(C1, S1) + (1 - P(S1 \mid C2)) \times S(C1, S2)$ and $(1-k)$ describes the share of C2 *Salmonella* given that we observed C2 at post-chill.

Graphical depictions on the movement of product before and after the implementation of serotype-based final product standards are provided in Figures **Figure 20** and **Figure 21**, respectively.

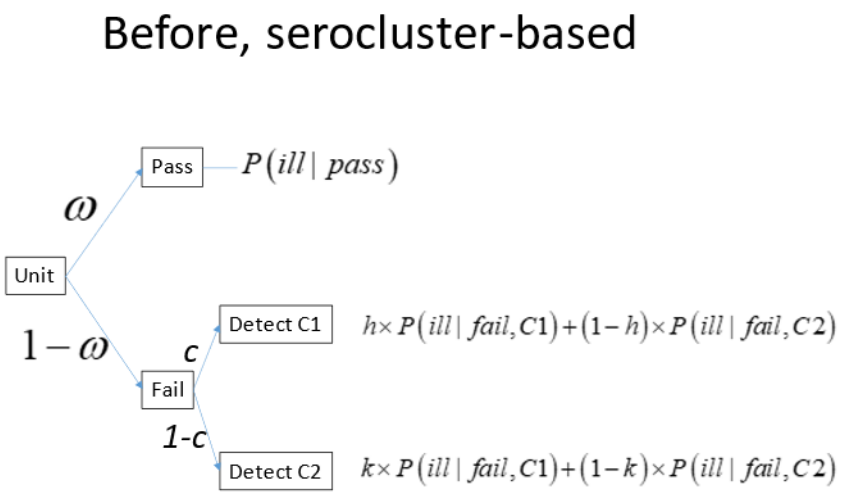


Figure 20: Schematic depiction of the 3 possible pathways through which all product moves prior to the implementation of a serotype-based final product standard.

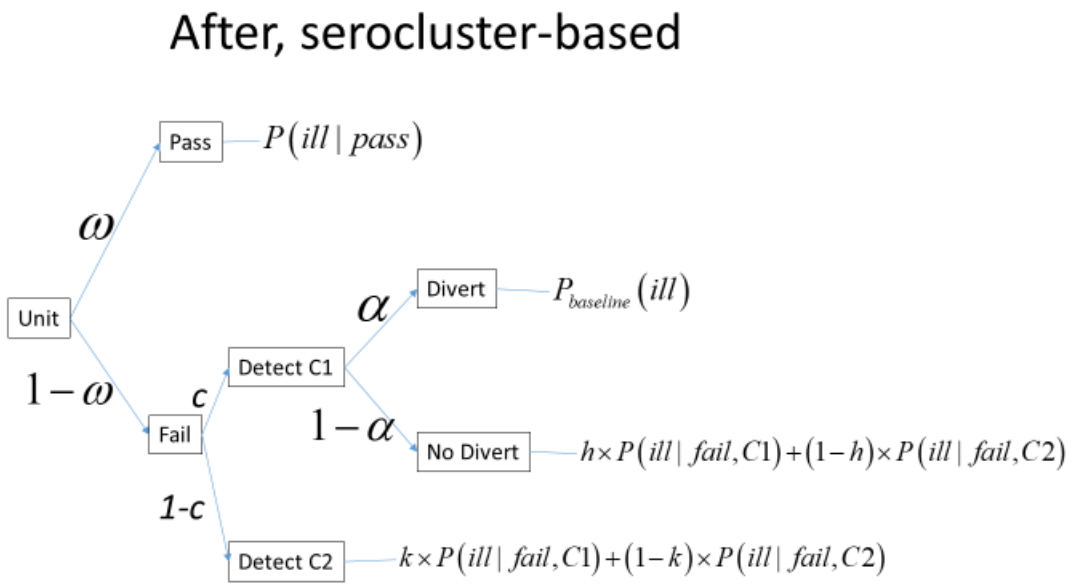


Figure 21:Schematic depiction of the 3 possible pathways through which all product moves after the implementation of a serotype-based final product standard.

Because of the absence of paired sampling data and complications in theoretic comparisons, this serotype modelling approach cannot be used for chicken parts or comminuted chicken performance standards. In fact, the concept of paired samples of chicken parts, or comminuted chicken, lots is difficult to conceive unless we assume parts or ground product lots are derived strictly from tested carcass lots. Interpreting such data is further complicated by the fact that the LOD for comminuted product is an order of magnitude lower than the carcass and parts LOD. That said, a serotype-based standard applied to carcass would necessarily, indirectly lower the amount of cluster 1 contaminated parts and comminuted. It is not possible to estimate the number of illnesses reduced via this pathway, but one can infer that additional reductions in human illnesses would occur beyond those assigned to carcasses.

5.2 Results

Using FSIS data, the size and composition of the chicken parts and comminuted chicken industries is summarized in **Table 34**. The comminuted product industry represent a smaller numbers of establishments, while the chicken parts industry is more than twice the size of the chicken carcass industry, which slaughters all live birds and supplies product to the parts and comminuted industries.

The total number of units produced by each industry are affected by risk management decisions. For this analysis, FSIS defines flocks as a “lot” susceptible to diversion for the chicken carcass industry, while an entire day of production is defined as a “lot” susceptible to diversion for the parts and comminuted industries. In 2021, FSIS collected about one sample per week from larger establishments while smaller establishments were sampled less frequently.

The fraction of non-compliant lots that can be diverted by testing (α) is the ratio of samples to units. This ratio is largest for the comminuted chicken industry and smallest for the chicken carcass industry.

It is estimated there are 125,115 chicken-associated *Salmonella* illnesses per year. This value is calculated as the product of the total number of CDC FoodNet cases per year (7,600), the share of these cases that are foodborne (66 percent) and of domestic origin (89 percent), the under-diagnosis multiplier for *Salmonella* (24.3)(Ebel, 2012c) and dividing by the FoodNet catchment area (15 percent), and multiplying by the portion IFSAC attribution to chicken (17.3 percent). These total cases are distributed across products by assuming the proportion of servings consumed (0.11, 0.83 and 0.06) is proportional to illnesses resulting from exposure to carcasses (whole chickens), parts and comminuted (ground) forms of chicken, respectively.

Using FSIS data, the total number of chicken carcasses produced in 2021 is 9.4 billion, with some 16% exported (NCC, 2021). Using USDA-ERS estimates for retail boneless chicken consumed in the U.S. (~2.4 pounds per carcass produced) and NHANES estimates for the average serving size for chicken (139.5 grams), we estimate 61 billion servings of chicken—in all its forms—are consumed in the U.S. each year. The ratio of total chicken illnesses to total chicken servings (2 per million) provides an empirical estimate of the risk of illness per serving (Hsi, 2015).

Table 34: Relevant parameters for the assessed products' industries.

Parameter	Chicken carcasses	Chicken parts	Comminuted chicken
No. establishments	206	484	74
No. units/yr.	203,718 lots	104,505 days	8,545 days
No. samples/yr.	9,635	14,192	1,991
Fraction of non-compliant units diverted (α)	0.05	0.14	0.23
Illnesses/yr.	13,763	103,845	7,507

Final product level-based standard

The results for final product level-based standards are used to illustrate the modeling mechanics for a chicken carcass standard that diverts test-positive lots based on a level threshold of 0.033 cfu/mL, which is the current limit of detection for FSIS's qualitative testing. This scenario presents the individual components of the model to demonstrate the general method of calculating point estimates for the fraction of illnesses prevented and the number of annual illnesses prevented. The baseline probability of illness – approximately 3 illnesses per million servings (**Table 35**) – is determined using numerical integration and the probability of illness among passing lots are estimates from 100 million Monte Carlo iterations.

Table 35: Results for a chicken carcass standard that diverts test-positive lots based on a *Salmonella* level limit of 0.033 cfu/mL (i.e., current LOD for qualitative assay).

Result	Symbol	Value
Fraction of lots compliant	ω	0.97
Probability of illness per serving from passing lots	$P(ill \mid pass)$	5.6×10^{-7}
Cluster 1	$P(ill \mid pass, C1)$	1.7×10^{-6}
Cluster 2	$P(ill \mid pass, C2)$	2.7×10^{-7}
Probability of illness per serving from failing lots	$P(ill \mid fail)$	8.6×10^{-5}
Cluster 1	$P(ill \mid fail, C1)$	2.5×10^{-4}
Cluster 2	$P(ill \mid fail, C2)$	4.4×10^{-5}
Baseline probability illness per serving	$P_{baseline}(ill)$	3.18×10^{-6}
Fraction of non-compliant lots diverted	α	0.047
w probability of illness per serving	$P_{new}(ill)$	3.06×10^{-6}
Fraction of illnesses prevented	$1 - \frac{P_{new}(ill)}{P_{baseline}(ill)}$	0.038
lnesses prevented per year	$\left(1 - \frac{P_{new}(ill)}{P_{baseline}(ill)}\right) \times \lambda_{ill}$	4700

The results illustrate the following:

- Only a small share of all lots are non-compliant with the standard (~3%).
- The probability of illness per serving among lots that pass the standard is about two orders of magnitude lower than that probability for failing lots.
- probability of illness per serving from C1 *Salmonella* is approximately one order of magnitude greater than that probability from C2 *Salmonella.*

The proportion of all lots that are diverted ($(1-\omega)\times\alpha$) is about 0.15%, but the result of this diversion is an overall reduction in illnesses of 3.8%. This effectiveness can be derived directly as the product of the fraction of lots diverted times the proportional difference in probability of illness per serving between the failing lots and the baseline probability of illness (i.e., $(1-\omega)\times\alpha\times\left(\frac{P(ill\mid fail)-P_{baseline}(ill)}{P_{baseline}(ill)}\right)=0.038$). In other words, replacing the failing lots with an average risk lot accounts for the disproportionate effect of this risk management option.

In the baseline, passing lots contribute about 20% of the total probability of illness while failing lots contribute about 80% $\left(\frac{\omega\times P(ill\mid pass)}{P_{baseline}(ill)}\approx 0.2\right)$. Similarly, we can determine that in the baseline, C1 contributes about 60% of the total probability of illness among passing or failing lots, while C2 contributes about 40% $\left(\frac{c\times P(ill\mid pass,C1)}{P(ill\mid pass)}\approx 0.6\right)$.

Increasing the level threshold necessarily increases omega (increases the fraction of units passing), which increases probability of illness among passing units and failing units (by including higher doses among passing and by removing lower doses among failing). But, because the exposure distribution is not affected by changing the level threshold (before any risk management options), the overall probability of illness per serving for the baseline should remain constant, regardless of the level threshold. This occurs because the rate of increase in the probability of illness among passing units is smaller than the rate of increase in the probability of illness among failing units as the level threshold increases. These relative effects can be appreciated by realizing that an increased level threshold allows into the passing population some higher risk units, but these units represent a small share of all passing units (so, the average for passing units does not increase much). In contrast, removing those same units from the failing population of units has a greater effect on the average for failing units.

To explore the reduction in illnesses associated with a level-based standard on final chicken carcasses, we used the model to predict effects across a range of level thresholds from 1 cfu/2600mL (-3.41 log10) to 100 cfu/mL (2 log10) (**Figure**). Illness reduction for each of these scenarios should be compared to a baseline of all chicken product illness because the 9.4 billion annual carcasses are the main source for all secondary chicken products (including parts and comminuted).

With the default assumption that failing lots would be diverted and replaced with an average lot with the baseline risk of illness, the illnesses prevented decreases steadily from about 4700 illnesses for the current LOD of 1 cfu/30mL (-1.48 log10) to nearly 450 illnesses for a threshold of 100 cfu/mL (2 log10). The peak effect of these predictions suggests that about 4 percent of all chicken-associated *Salmonella* illnesses (~125,000) might be prevented if this standard were imposed at a level threshold of 1 cfu/150mL (i.e., 5 times lower than the current LOD). Nevertheless, at a level threshold of 1 cfu/g, the number of illnesses prevented decreases to about 2450 (2% of all illnesses).

If we consider level thresholds less than the current LOD, the illnesses prevented initially increases slightly before declining (**Figure 22**). This pattern was unexpected, and its explanation reveals the importance of the assumption about replacement of diverted lots. When the level threshold is reduced, more lots become eligible for diversion, but the additional lots necessarily represent a lower risk of illness than the other failing lots. At some level threshold, we begin diverting lots whose risk of illness was actually lower than the average risk across all lots, and the effect is to moderate the overall reduction in illnesses. At an extremely low level threshold, we fail every lot (whose average risk is equivalent to the population's average) and simply replace those lots with others of equivalent risk, which produces no illnesses prevented.

This explanation is clearer when the results are compared with an alternative model that assumes diverted lots are replaced by passing lots (**Figure 22**). From this alternative model, we see little effect for level thresholds near or larger than the current LOD. For these scenarios, there is little effect between replacing diverted lots with an average risk lot or a passing lot. Nevertheless, as the level threshold is reduced below the current LOD, the alternative model demonstrates that illnesses prevented continues to increase. Therefore, assuming replacement of diverted lots with passing lots generates a pattern of illness reductions that is monotonically decreasing from a maximum of about 5800 (4.7 percent of 125,000 total illnesses, which is the fraction of all lots diverted, α) to zero.

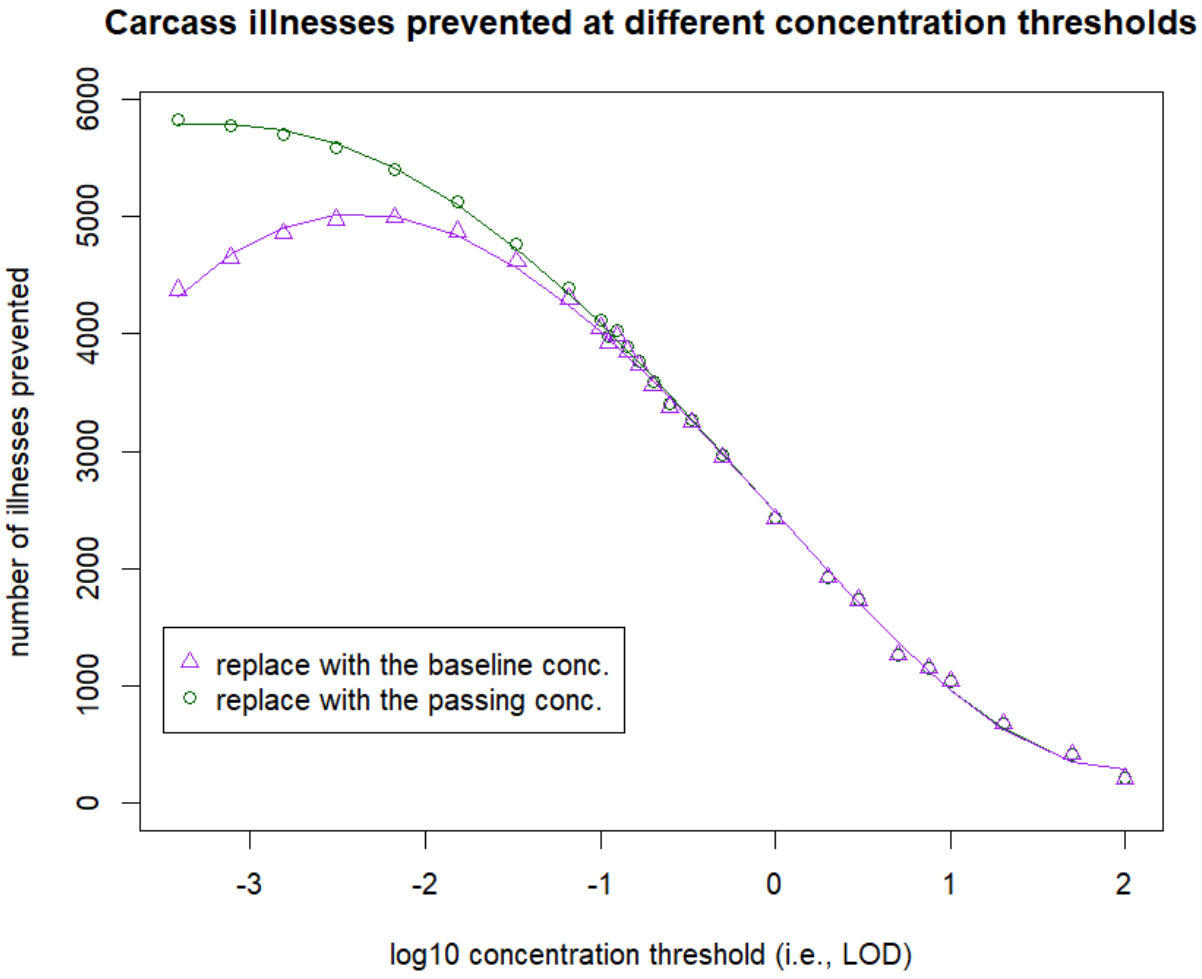


Figure 22: The predicted number of illnesses prevented across a range of *Salmonella* level thresholds is shown for a standard applied to chicken carcasses.

Chicken Parts

When simulating the chicken parts model, the baseline probability of illness is approximately 5 illnesses per million servings. This estimate is similar to our empirical risk estimate.

For chicken parts, using a level threshold equal to the current LOD (0.0333cfu/mL) is predicted to prevent about 7850 illnesses per year or 7.5 percent of the ~104,000 chicken parts illnesses each year (**Figure 23**). At a level threshold of 1 cfu/mL, the number of illnesses prevented is about 1400 (a 1.4 percent reduction). At a level threshold of 10 cfu/mL, the number of illnesses prevented is essentially zero. This lack of reduction at levels above 10 cfu/mL occurs because few samples were enumerated at the upper limit of detection for MPN levels in the parts microbiological baseline study (FSIS, 2013) of 11 cfu/mL, when the prevalence was roughly 4 times higher than the current estimate. Thus, the adjustment of the level distribution so that the prevalence matches the current estimate yields a level distribution with almost no mass at higher levels.

The alternative assumption that diverted lots of chicken parts are replaced by passing lots predicts that the maximum illnesses prevented approaches 14,000 (a 13.4 percent reduction) as the level threshold is reduced to 1 cfu/2600mL. The default model assumption that diverted lots are replaced by average risk lots suggests that illnesses prevented will start to decrease if the level threshold is less than about 1 cfu/100mL (-2 log10).

Relative to chicken carcasses, the high percent illness reduction for chicken parts at the LOD level threshold is explained mostly by a higher fraction of failing lots and a higher fraction of non-compliant lots diverted. For example, the fraction of lots failing $(1 - \omega)$ is 6.7 percent for parts, but 3 percent for carcasses. Similarly, the fraction of non-compliant lots diverted is 13.6 percent for parts and 4.7 percent for carcasses. Nevertheless, for a level threshold of 1 cfu/mL, the percent reduction in illnesses is similar for chicken carcasses (2 percent) and parts (1.4 percent). The differences in fraction of lots failing is a function of the underlying contamination distributions between parts and carcasses, but the difference in the fraction of non-compliant lots diverted is because parts diversion applies to a full day of production, while carcass diversion applies to individual flocks and large establishments typically process between 3 and 6 flocks per day.

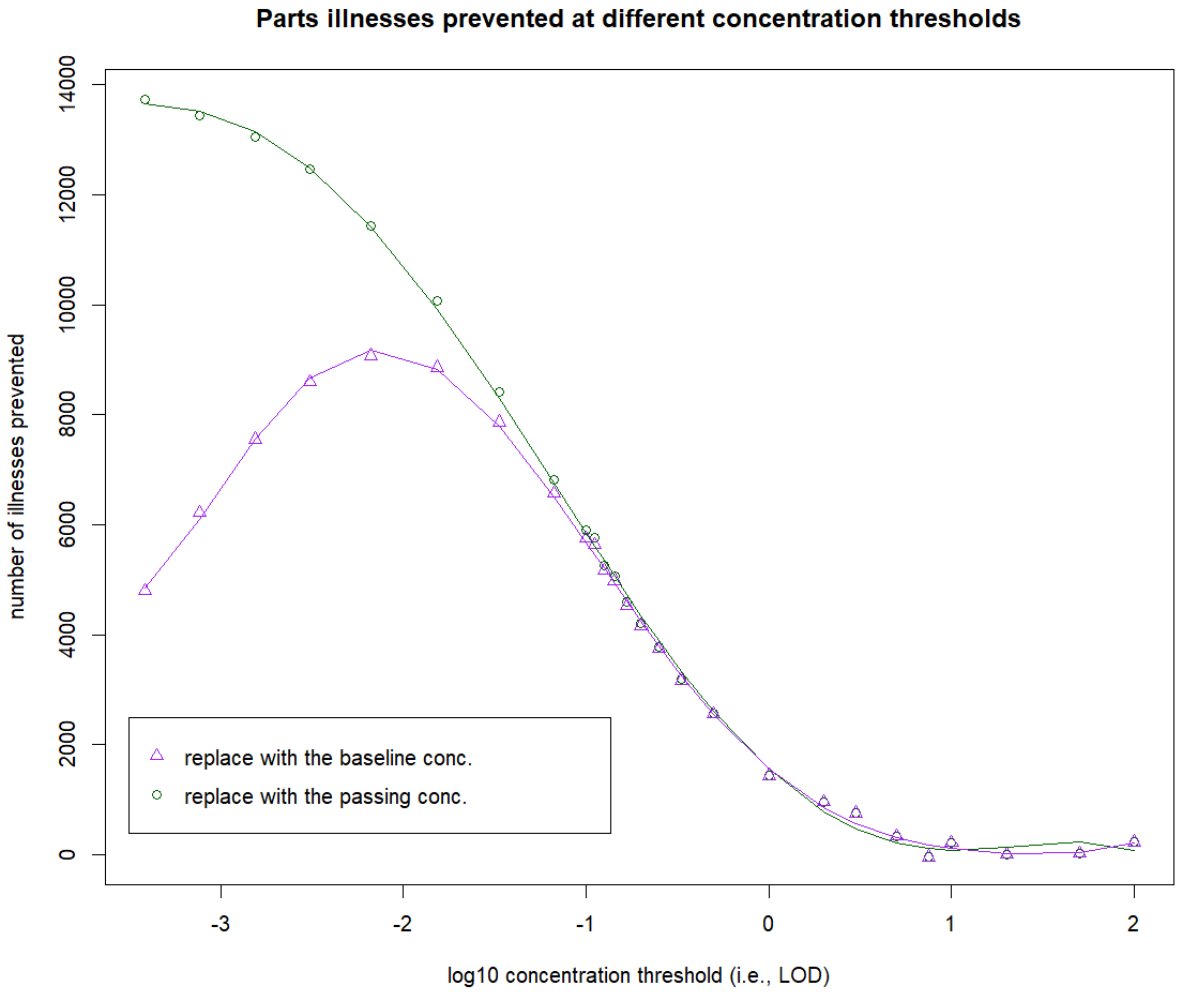


Figure 23: The predicted number of illnesses prevented across a range of *Salmonella* level thresholds is shown for a standard applied to chicken parts.

Comminuted Chicken

When simulating the comminuted chicken model, the baseline probability of illness is approximately 2.5 illnesses per 100,000 servings. This estimate is about 10 times larger than our empirical risk estimate of 2 illnesses per million servings and may reflect a truly higher risk for this product relative to intact chicken products, or it may indicate that one or more of the model inputs are specified incorrectly (see Sensitivity and Uncertainty Analysis for the effects of misspecification on the model's predictions).

For our default comminuted chicken model (i.e., replace diverted lots with average risk lots), using a level threshold equal to the current LOD (1 cfu/325g, -2.51 log10) is predicted to prevent about 1,250 illnesses per year or 17 percent of the ~7,500 comminuted chicken each year (**Figure 24**). Nevertheless, predicted illnesses prevented actually increases to 1,500 (a 20 percent reduction) when the level threshold increases to 1 cfu/10g (-1.0 log10). This increase in level threshold eliminates diverting lots

whose risk of illness is less than the average risk, thereby improving the effectiveness of diversion. At a level threshold of 1 cfu/g, the number of illnesses prevented is about 1,400 (a 19 percent reduction). At level thresholds of 10 cfu/g and 100 cfu/g, the number of illnesses prevented is 1,000 (a 14 percent reduction) and 600 (an 8 percent reduction), respectively.

The alternative assumption that diverted lots of comminuted chicken are replaced by passing lots predicts that the maximum illnesses prevented approaches 1,740 (a 23 percent reduction) as the level threshold is reduced to 1 cfu/2600 g.

In contrast to the chicken carcasses and parts results, the comminuted chicken results suggest that a level threshold above the current LOD might generate a larger reduction in illnesses. This difference in predictions can be explained by the difference in current LOD between comminuted chicken and the other two products. The comminuted chicken LOD (1 cfu/325g) is an order of magnitude lower than that for carcasses and parts (1 cfu/30mL). Therefore, the comminuted chicken LOD as a level threshold will classify more lots as non-compliant with initial levels closer to the mean level for the industry. Such a classification increases the likelihood, generally, that diverted lots pose a lower risk than the industry average.

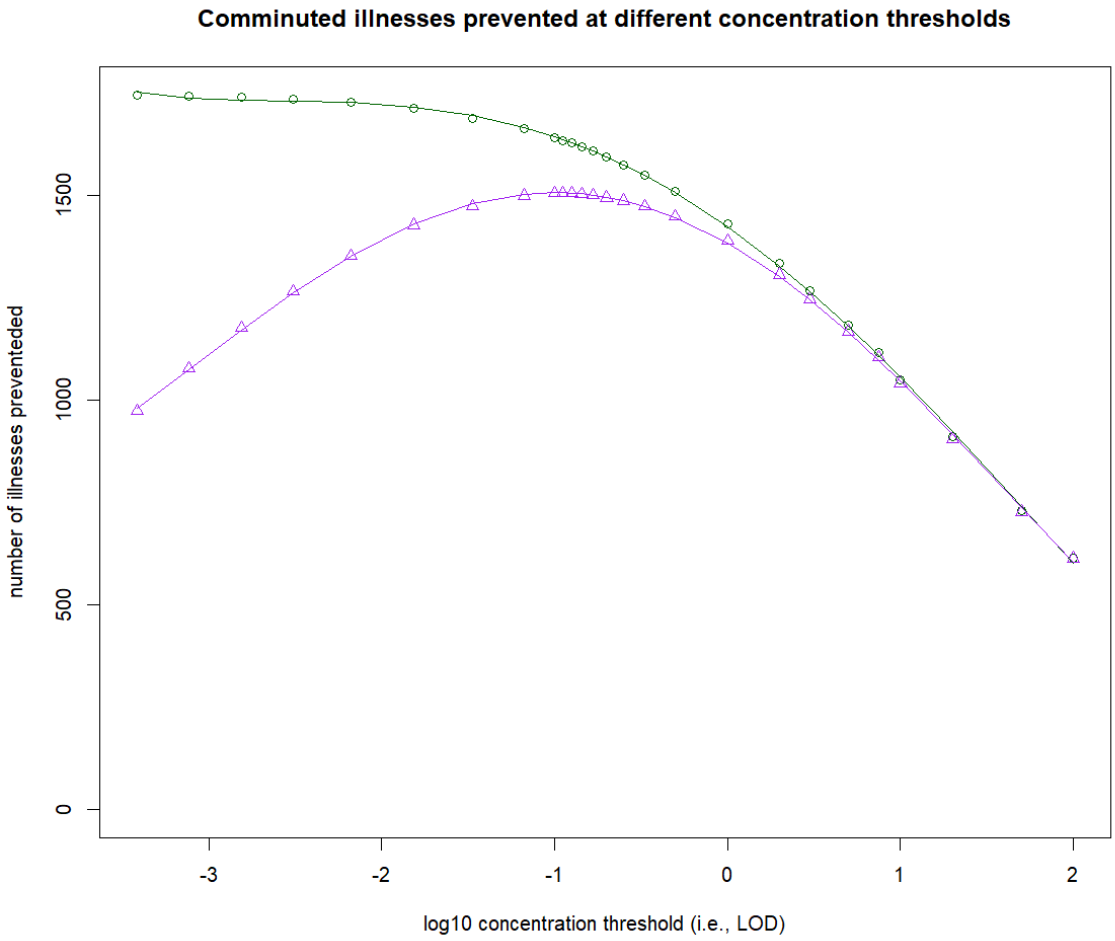


Figure 24: The predicted number of illnesses prevented across a range of *Salmonella* level thresholds is shown for a standard applied to comminuted chicken.

Final product Serocluster Standard

The intention of the final product standards based on serocluster virulence is to divert lots where the higher virulence serotypes are identified at post-chill. A summary of the key results are presented in **Table 36**.

Table 36: Results of a chicken carcass standard that diverts C1-positive lots (assuming an LOD of 0.033 cfu/mL).

Fraction of *Salmonella* in cluster 1	c	0.20
Share of C1 *Salmonella* among failing lots when C1 is detected	h	0.58
Share of C1 *Salmonella* among failing lots when C2 is detected	k	0.11
New probability of illness per serving	$P_{new.sero}\left(ill\right)$	3.14×10^{-6}
Fraction of illnesses prevented	$1-\frac{P_{new.sero}\left(ill\right)}{P_{baseline}\left(ill\right)}$	0.014
Illnesses prevented per year	$\left(1-\frac{P_{new.sero}\left(ill\right)}{P_{baseline}\left(ill\right)}\right)\times\lambda_{ill}$	1800

These results illustrate that a slight majority of *Salmonella* among lots that fail and C1 is detected are the C1 serotypes, while most of the *Salmonella* among lots that fail and C2 is detected are the C2 serotypes.

The share of all lots that are diverted ($(1-\omega)\times c\times\alpha$) is about 0.03%, but the result of this diversion is an overall reduction in illnesses of 1.7%. This effectiveness can be derived directly as the product of the fraction of lots diverted times the proportional difference in probability of illness per serving between the failing lots and the baseline probability of illness (i.e., $(1-\omega)\times c\times\alpha\times\left(\frac{\left[h\times P\left(ill\mid fail,C1\right)+\left(1-h\right)\times P\left(ill\mid fail,C2\right)\right]-P_{baseline}\left(ill\right)}{P_{baseline}\left(ill\right)}\right)=0.017$). In other words, replacing the failing lots with an average risk lot accounts for the disproportionate effect of this risk management option.

These point estimates represent expectations based on model inputs. Nevertheless, substantial variability – beyond the simple Poison variability implied in our first equation – is not captured in these estimates. For example, the composition (C1 vs. C2) of a lot that is detected varies at least by its scheme. Our point estimates have averaged the effect across all such lots, but this variability will have an effect in nature.

The point estimates also do not reflect their uncertainty. Yet, substantial uncertainty attends the true dose-response functions, the true initial contamination and attenuation distributions, the true mixtures of C1 and C2 across lots and most other inputs to this model. This uncertainty has yet to be explored in the context of this model. Nevertheless, none of the final product testing options come close to achieving the desired 25% reduction in *Salmonella* illnesses (i.e., a reduction of ~31,000 illnesses) that is necessary to meet the HP2030 goal.

Lots Diverted

At the request of FSIS risk managers, the risk assessment model and other data analyses contained in this document were used to develop the following estimates of annual lots diverted for the threshold scenarios under consideration. **Table 37** contains estimates for the main threshold scenarios run in the risk assessment, along with three most prevalent higher virulence serotype based diversion. The top three higher virulence *Salmonella* serotypes for chicken (Top 3) are Entertidis, Tymphiurium, and I,4, [5],12,i:-.

Table 37: Estimated annual product lots and pounds of product expected to be diverted under various concentration and serotype scenarios.

Products	Scenarios (cfu/mL or cfu/g)	Lots Diverted	Total Lots	Percent of Lots Diverted	Pounds of Product Diverted	Total Pounds Production	Percent of Total Weight Diverted
Chicken Carcasses	0.03	297	203,718	0.1%	64,758,176	44,418,875,500	0.1%
	0.03 + Top 3	71		0.03%	15,480,911		0.03%
	1	26		0.01%	5,669,066		0.01%
	1 + Top 3	6		0.003%	1,308,246		0.003%
	10	3[f]		0.001%	654,123		0.001%
	10 + Top 3	0.7		0.0003%	152,629		0.0003%
	100	0.3		0.0001%	65,412		0.0001%
	100 + Top 3	0.1		0.00005%	21,804		0.00005%
Chicken Parts	0.03	949	104,505	0.9%	334,791,787	36,867,666,665	0.9%
	0.03 + Top 3	313		0.3%	110,421,316		0.3%
	1	23		0.02%	8,114,026		0.02%
	1 + Top 3	8		0.01%	2,822,270		0.01%
	10	0.7		0.0007%	246,949		0.0007%
	10 + Top 3	0.2		0.0002%	70,557		0.0002%
	100	0.007		0.000007%	2,469		0.000007%
	100 + Top 3	0.002		0.000002%	706		0.000002%
Comminuted Chicken	0.003	540	8,545	6.3%	168,422,653	2,665,132,530	6.3%
	0.003 + Top 3	157		1.8%	48,967,327		1.8%
	1	57		0.7%	17,777,947		0.7%
	1 + Top 3	17		0.2%	5,302,195		0.2%
	10	16		0.2%	4,990,301		0.2%
	10 +Top 3	5		0.06%	1,559,469		0.06%
	100	3		0.04%	935,681		0.04%
	100 + Top 3	1		0.01%	311,894		0.01%

[f] The difference between this estimate and the 14 samples enumerated at or above 10 cfu/mL in the 2022 Exploratory Young Chicken Carcass Sampling Program is largely a function of the qPCR enumeration effect (see Appendix B) and the assumption of enumeration test accuracy in the model.

5.3 Sensitivity and Uncertainty Analysis

Techniques for approximation

For final product concentration standards, simplifications of our model can suggest that the proportional reduction in illnesses prevented by a standard might be approximated by assuming a linear dose-response function and/or ignoring the effect of attenuation between production and consumption.

The proportional reduction in illnesses is $1-\frac{P_{new}(ill)}{P_{baseline}(ill)}$ and the ratio

$$\frac{P_{new}(ill)}{P_{baseline}(ill)}=\frac{\int R(d_{new})h(d_{new})\partial d}{\int R(d_{baseline})h(d_{baseline})\partial d}$$

describes the relative probability of illness per serving after a risk management option effect (new) to before (baseline). Ultimately, an effective risk management option must change the exposure distribution such that, after integrating across a dose-response function, it reduces the probability of illness per serving relative to that probability before the option.

If the dose-response is approximately linear ($R(d) \approx \gamma d$), then we are left with a ratio of average doses per serving ($\frac{P_{new}(ill)}{P_{baseline}(ill)} \approx \frac{\tilde{d}_{new}}{\tilde{d}_{baseline}}$.) (Williams, 2011c). Given $d = x \times a$, where x is the initial contamination concentration random variable and α is an attenuation random variable that is independent from x and does not change after the policy, then the final simplification is a ratio of the average initial contamination concentrations (

$$\frac{P_{new}(ill)}{P_{baseline}(ill)} \approx \frac{E[x_{new}]E[a]}{E[x_{baseline}]E[a]} \approx \frac{\bar{x}_{new}}{\bar{x}_{baseline}}).$$

The generally low *Salmonella* doses implied by the chicken data suggests the assumption of a linear dose-response relationship is not unreasonable. Nevertheless, our model is complicated by the consideration of two dose-response functions (for C1 and C2) and a remixing of lots above and below some chosen threshold to determine the new initial contamination distribution.

If these approximations are reasonable, then the general effect of the concentration standards can be estimated without considering the specific dose-response functions or attenuation between production and consumption. A similar conclusion with respect to so-called "prevalence-based" standards was reported previously (Ebel, 2015).

To compare the approximate estimates to those generated by the full model, we first calculate the average concentration for the baseline using the log10 parameters in **Table 20** and **Table 21**. For example, for chicken carcasses $\bar{x}_{baseline} = e^{-4.512\times\ln(10)+0.5*(\ln(10)*1.624)^2} = 0.0334 cfu/mL$. Next, we calculate the conditional expected value for lots that are below a threshold concentration T. For μ and σ

in natural log units, $E\left[x \mid x \le T\right] = \dfrac{\bar{x}_{baseline} \times \Phi\left(\dfrac{\ln(T) - \mu - \sigma^2}{\sigma}\right)}{\omega}$, where $\Phi(\ )$ is the cumulative probability from a standard Normal distribution and ω is fraction of all lots that are below T (as defined previously). For chicken carcasses using a 1/30mL threshold, $E\left[x \mid x \le T\right] = 0.001 cfu / mL$. The conditional expected value for lots above a threshold is $E\left[x \middle| x \rangle T\right] = \dfrac{\bar{x}_{baseline} - \omega \times E\left[x \mid x \le T\right]}{1-\omega} = 1.05 cfu / mL$ for chicken carcasses at the same threshold.

Using these values, we can calculate our simplified replacement for $P_{new}(ill) \approx \bar{x}_{new} \approx \omega \times E\left[x \mid x \le T\right] + (1-\omega)\left[\alpha \times \bar{x}_{baseline} + (1-\alpha) \times E\left[x \mid x > T\right]\right] \approx 0.0320$, such that the approximate reduction in illness is about 4% (where α is the fraction of non-compliant establishments diverted as defined previously). This compares with 3.8% reported above for this same scenario using the full model.

Across a full range of concentration thresholds, the approximation for the proportional reduction in illnesses tends to be similar or somewhat larger than that estimated using the full model (**Figure 25**). The greatest difference is observed for comminuted chicken as the concentration threshold becomes larger. The differences suggest that the assumption of linearity in the dose-response function (applicable to the approximation) becomes less appropriate as effects of the diversion options are applicable to larger dose concentrations. As shown previously, using a linear approximation overestimates the probability of illness as dose increases (Williams, 2011c).

This analysis suggests that results estimated from a simplified model that only considers changes to the initial contamination distribution are comparable to estimates from a full model that simulates a) the full range of initial contaminations by passing and failing status, b) the modification of these initial contamination levels by an attenuation distribution, c) separate estimates of probability of illness given dose for two virulence clusters via dose-response functions. Such findings support the general idea that both the attenuation and dose-response functions have limited influence on the full model's estimates; i.e., the full model's results are not highly influenced by either attenuation or dose-response. Nevertheless, application of attenuation and dose-response are necessary for improved accuracy in estimates as the threshold increases.

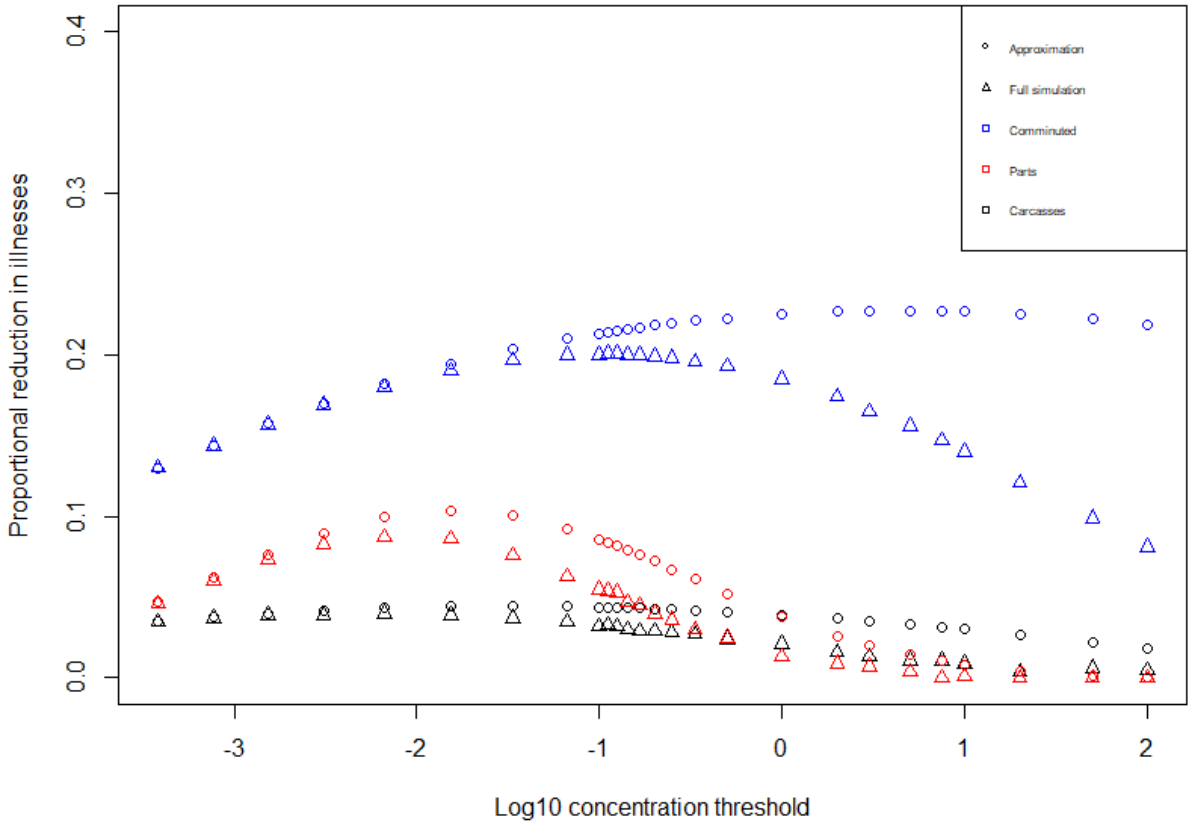


Figure 25: Approximate estimates of the proportional reduction in human illnesses – based solely on changes in the mean initial contamination concentration – across a range of concentration thresholds are compared with the full model's estimates for chicken carcasses and parts and comminuted chicken.

Sensitivity of effectiveness to model inputs

To conduct sensitivity analysis, we change individual model inputs – while holding others at their baseline values – and explore changes to the proportional public health effectiveness across a range of concentration thresholds. We perform these analyses in the comminuted chicken model because, given the greater effectiveness of the full range of concentration thresholds, the alternative scenarios approach stability using fewer iterations of the full model (~1 million) and the results are more easily observed graphically. Nevertheless, the general conclusions are applicable to the other products because the model structure is the same.

We examine the effects of increasing or decreasing the degree of attenuation between initial contamination and consumption by increasing or decreasing the negative mean of the log10 distribution from its default of -5 log10s. We examine the effects of increasing or decreasing the mean or standard deviation of the log10 initial contamination distribution by increasing or decreasing these parameters by 1 log10 unit. Such changes are well beyond the magnitude of uncertainty about the fitted parameters of the initial contamination distribution, but the general effect of changing these parameters is easier to observe by exaggerating the change.

To examine the effects of alternative dose-response functions, we use the lower- and upper-bounds for the C1 and C2 dose-response relationships (95 percent confidence limits in the uncertainty dimension – see

EpiX Analytics' Report on Dose-Response Model). For example, the lower bound percent reduction for a concentration threshold is estimated in the full model with default settings for all inputs except that the lower bound dose-response functions for C1 and C2 are used. Finally, we examine the effects of assuming the share of organisms in cluster 1 are 0% or 100% versus the baseline value of 20%.

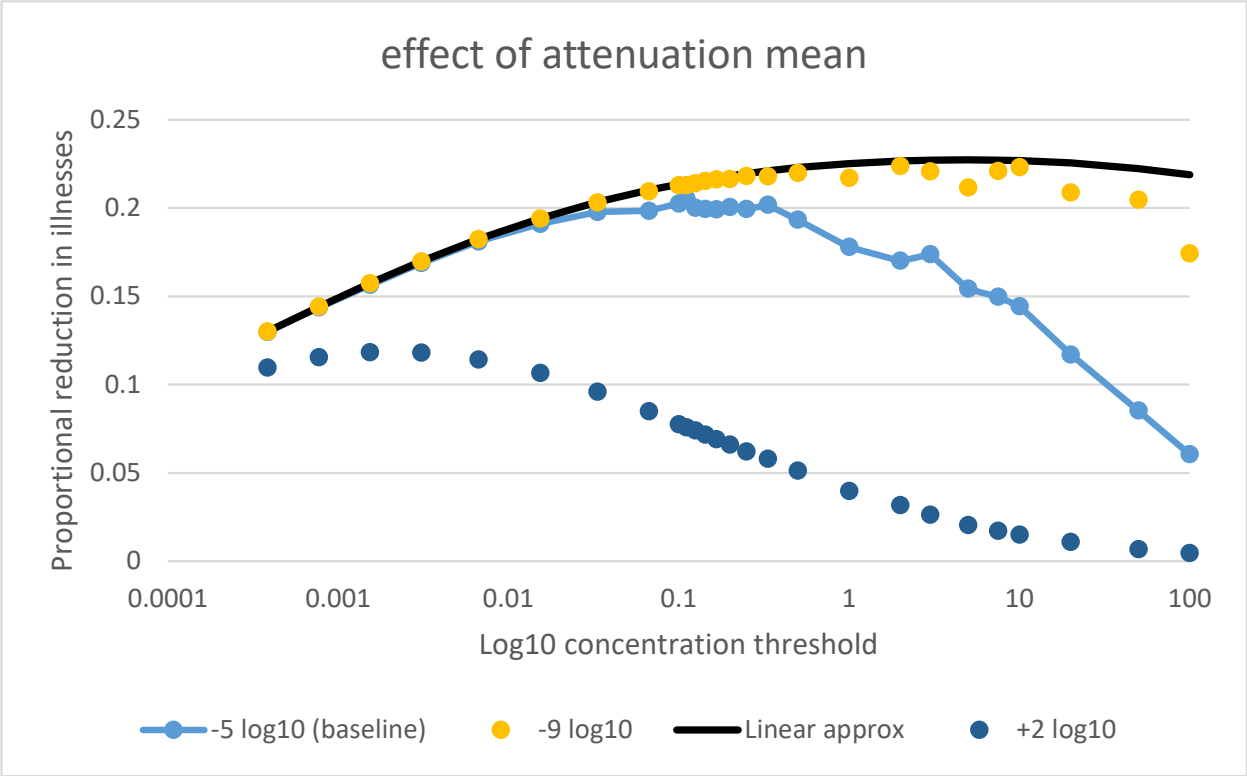


Figure 26*:* Effect of attenuation mean in comminuted chicken.

Changing the mean of the attenuation distribution to -9 log10 has the effect of substantially reducing the magnitude of the doses consumed. The predicted proportional reduction in illnesses for larger concentration thresholds approaches the effect calculated using the linear approximation where both the attenuation variable and the dose-response functions are ignored. Such a finding demonstrates that more attenuation of the initial contamination distribution generates low doses where the assumption of a linear dose-response relationship is most appropriate. Although this change in the attenuation distribution generates baseline probability of illness per serving estimates that are too low – therefore, inconsistent with empirical expectations – its approach to the linear approximation suggests that the approximation may represent an upper boundary of the effectiveness of concentration thresholds.

Changing the mean of the attenuation distribution to +2 log10 models the consumption of doses that are essentially unaltered from the initial contamination levels. For example, if the default -5 log10 is thought to represent a -7 log10 average reduction combined with a +2 log10 serving size, then this change only considers the serving size adjustment to the initial contamination. Generally, this change results in a lowered public health effectiveness – relative to the baseline model – across the range of concentration thresholds considered. It also illustrates the progressively important influence of the non-

linear dose-response functions on moderating the effect of increasing concentration thresholds. Although this change in the attenuation distribution generates baseline probability of illness per serving estimates that are too high – therefore, inconsistent with empirical expectations – it may represent a lower boundary of the effectiveness of concentration thresholds.

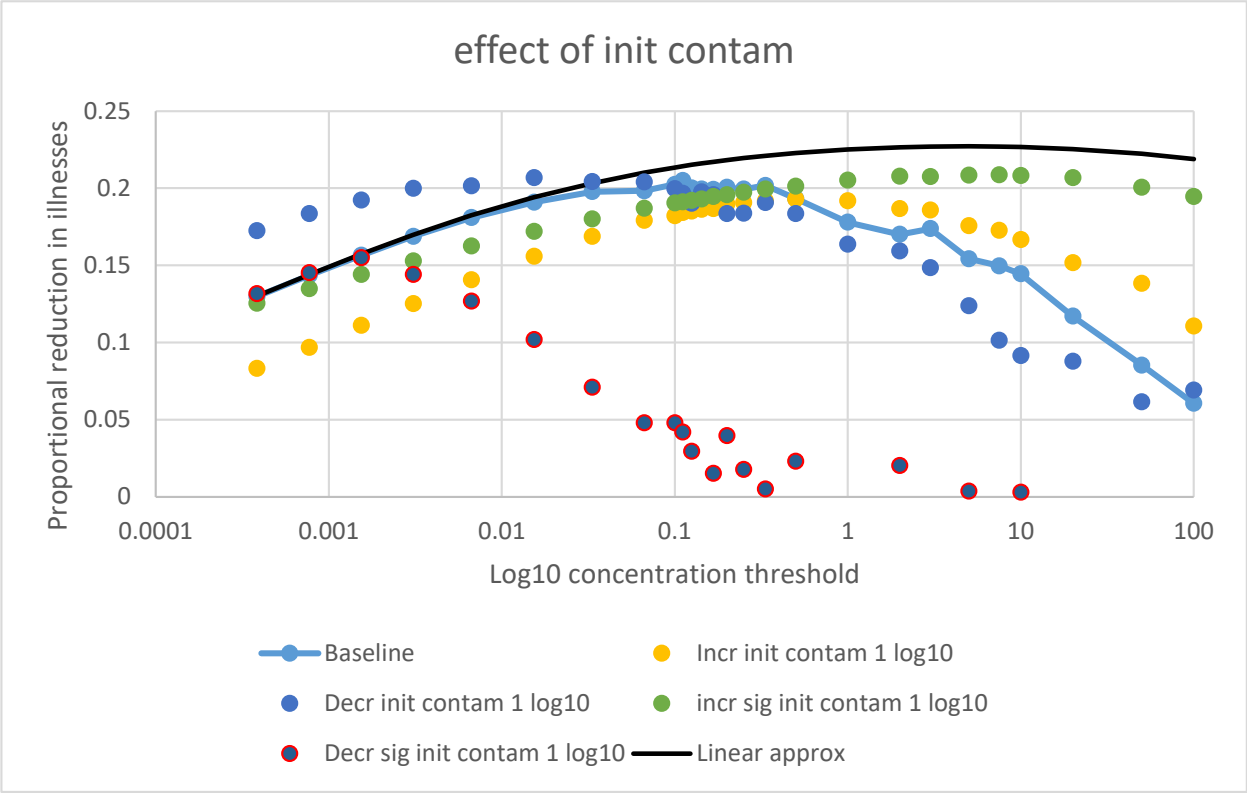


Figure 27: Effect of initial contamination in comminuted chicken

Changing the mean of the initial contamination distribution shifts this distribution to higher or lower concentrations in log10 values. Increasing this mean by 1 log10 (from -3.7 to -2.7) results in a right-shifting of the effectiveness curve – relative to the baseline predictions – so that the proportional reduction in illnesses is larger for higher concentration thresholds and smaller for lower concentration thresholds. The opposite effect is noted when we decrease the mean initial contamination from -3.7 log10 to -4.7 log10; the effectiveness curve is shifted to the left relative to the baseline predictions. Nevertheless, the amplitude of the effectiveness curve (i.e., maximum effectiveness) for either increasing or decrease the mean of the initial contamination is similar to the baseline. It should be noted, however, that increasing or decreasing the mean of the initial contamination distribution substantially generates indefensible probability of illness per serving estimates that are too high or low.

Increasing the standard deviation (sig) of the initial contamination distribution from 1.949 (baseline) to 2.949 creates higher and lower contamination levels in the distribution's tails. Consequently, the effectiveness of higher concentration thresholds is greater than the baseline model and smaller than the baseline for lower concentration thresholds.

Decreasing the standard deviation of the initial contamination distribution to 0.949 reduces the frequencies of larger (and smaller) contamination levels. Consequently, the public health effectiveness is reduced across the full range of concentration thresholds, although at very low concentration thresholds there is very little difference from the baseline. This pattern is similar to what is observed for the chicken parts baseline predictions; in that model, the default initial contamination has a low standard deviation such that the effectiveness of higher concentration thresholds is diminished because there are few lots with high enough concentrations to be detected and diverted.

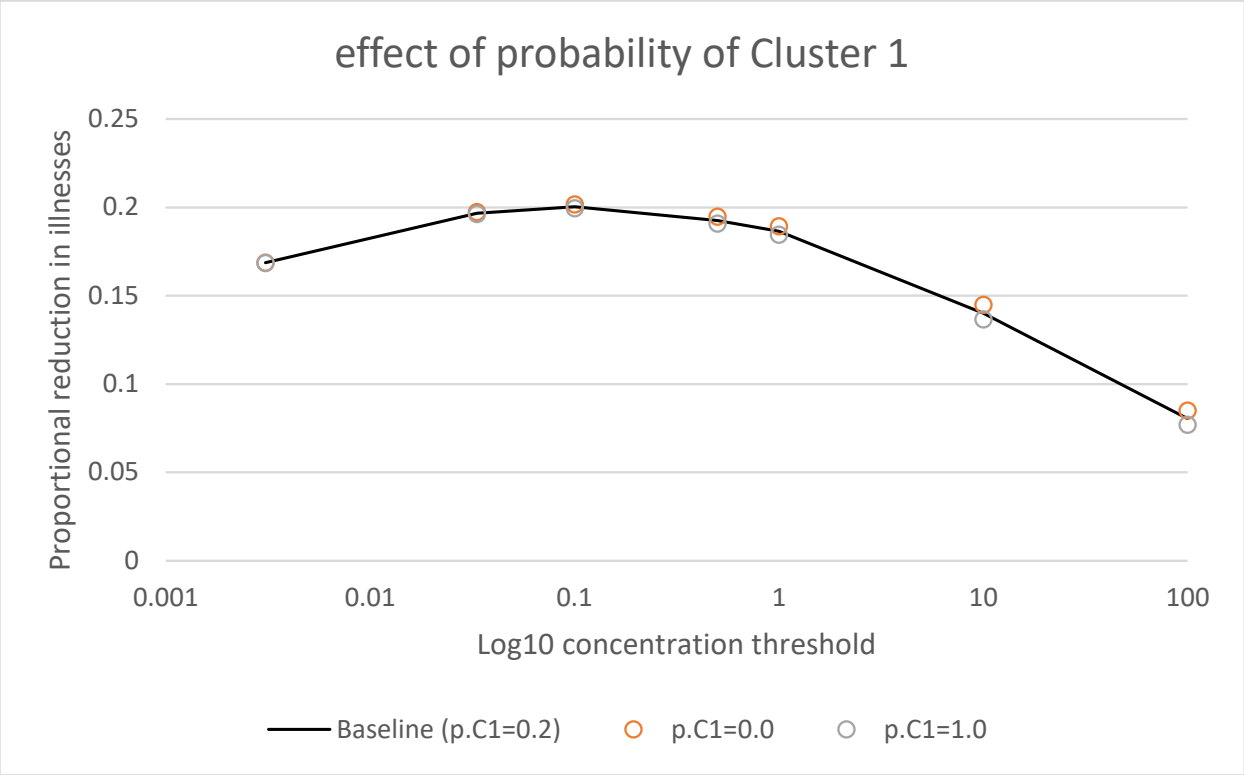


Figure 28: Effect of probability of cluster 1

Increasing or decreasing the probability of cluster 1 organisms relative to the baseline assumption of 0.20 results in very little change in the proportional reduction in illnesses across the range of concentration thresholds. Although the probability of illness per serving is somewhat larger when 100% of organisms are assumed to be C1, the public health effectiveness is slightly lower for all concentration thresholds because the dose-response function for C1 is more non-linear than that for C2. Consequently, the opposite behavior is evident when we assume 0% of organisms are C1; the probability of illness per serving is somewhat smaller but the public health effect is slightly greater across the range of concentration thresholds.

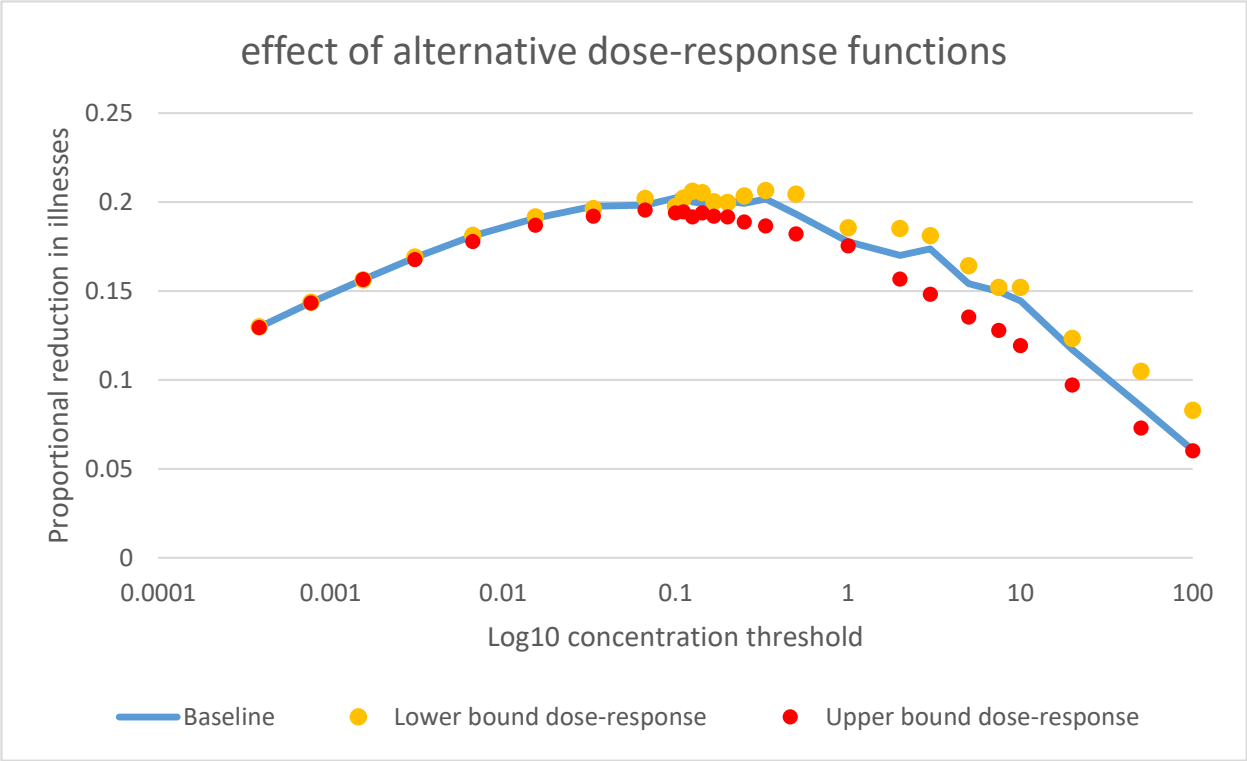


Figure 29: Effect of alternate dose-response functions on comminuted chicken model.

Using the lower or upper bound dose-response relationships for C1 and C2 results in very little change in the proportional reduction in illnesses across the range of concentration thresholds. Because the upper bound dose-response relationships demonstrate departures from linearity at lower doses than the baseline or lower bound relationships, the public health effect estimated for the upper-bound dose-response scenario is slightly lower than the baseline scenario at the higher concentration thresholds. Nevertheless, the probability of illness per serving for the lower and upper bound scenarios are smaller and larger than the baseline scenario's probability of illness per serving.

Uncertainty analysis

We want to examine the effects of uncertainty on the estimated annual number of illnesses prevented by different concentration thresholds. To motivate this analysis, we use the following equation and propagate uncertainty about each component to estimate an uncertainty distribution about illnesses prevented.

$$I_{avoided} \sim \left(1 - \frac{P_{new}(ill)}{P_{baseline}(ill)}\right)\lambda_{ill} \sim Prop.reduct \times \lambda_{ill}$$

Uncertainty about the proportional reduction in illnesses ($Prop.reduct$) is modeled as a

$Pert(min, mode, max)$ distribution. For a given concentration threshold, the mode is the effectiveness estimated using our full model. As illustrated in our sensitivity analysis (**Figure 26)**, boundary effectiveness estimates can be defined by assuming a minimal or maximal attenuation effect on the consumed dose.

For a particular concentration threshold, the minimum effectiveness value is estimated from the full model but with the mean of the attenuation distribution set equal to +2 log10 (instead of -5 log10 in the modal scenario). In a sense, this change eliminates the 7 log10 reduction expected between production and consumption while still adjusting the initial concentration to account for a 2 log10 serving size. This change results in larger doses for all exposures and, correspondingly, a more substantial effect from the non-linear dose-response relationship. This nonlinearity, at higher concentration thresholds, lessens the differences in the probability of illness between failing and passing lots and, consequently, lessens the effect from diverting failing lots.

For a particular concentration threshold, the maximum effectiveness value is the approximation we get when we assume the dose-response relationships are linear. This approach is essentially the same as a scenario in the full model where the mean attenuation reduces the initial contamination by more than 9 log10 (see ***Figure 26***) and, consequently, all exposures are very small doses. Such a change ensures that differences in average concentration between failing and passing lots are not modified by a non-linear dose-response relationship and the predicted effect of diverting failing lots is maximized.

Uncertainty about the annual number of illnesses associated with chicken before the effects of a policy (λ_{ill}) is modeled as $\lambda_{ill} \sim \frac{F \times B \times D \times u \times a}{C}$, where F=7600 is a typical annual number of FoodNet *Salmonella* cases reported (CDC, 2022), B=0.66 is the fraction of foodborne cases among all *Salmonella* cases (Beshearse, 2021), D=0.89 is the fraction of cases acquired domestically (Scallan, 2011) and C=0.15 is the FoodNet catchment area fraction. The variable *u* is the under-diagnosis multiplier for *Salmonella* and is modeled as a $gamma(32.83, 1/0.74)$ distribution (Ebel, 2012c). The variable *a* is the attribution fraction of *Salmonella* cases associated with consumption of chicken and is modeled as a $Pert(9.4, 17.3, 25.3)$ distribution (IFSAC, 2022)[7]. For the carcass final product standard, the total *Salmonella* illnesses associated with chicken are used for λ_{ill} . . For the chicken parts and comminuted chicken standards, the values for λ_{ill} are further adjusted to be 0.83 and 0.06 of the total *Salmonella* illnesses associated with ccken (see section **3.2**).

The summarized components of our uncertainty analysis for concentration thresholds of interest to FSIS risk managers demonstrate that the credible range for the proportional reduction in illnesses spans at least an order of magnitude (**Table 39**). Nevertheless, the boundaries about the proportional reduction in illnesses are most observable for concentration threshold values that are below the range of interest

[7] The reference provides a mode and 90% confidence intervals directly. Minimum and maximum values are 99.9 percentiles estimated as $x_{0.999} = \tilde{x} \pm Z_{0.999}\left(\frac{x_{0.95} - \tilde{x}}{Z_{0.95}}\right)$, where Z_k is the k^{th} quantile of a $Normal(0,1)$ distribution and $\tilde{x}$ is the reported mode.

to FSIS (e.g., between 1 cfu/65mL and 1 cfu/2mL for chicken parts) (**Figure 30**). The 95 percent credible interval for annual chicken (carcass) illnesses spans approximately one log10, but this range is less for chicken parts and comminuted chicken.

Table 38: Descriptions of the uncertainty distributions for the parameters used to estimate annual illnesses prevented are shown.

Concentration threshold*	Variable	Chicken carcasses	Chicken parts	Comminuted chicken
	λ_{ill} , mean (95% CI)	125,000 (73,000 – 193,000)	104,000 (60,000 – 160,000)	8,000 (4,000 – 12,000)
0.03 cfu/mL	percent reduction	Pert(min=0.0095, mode=0.0369, max=0.0444)	Pert(min=0.0192, mode=0.0757, max=0.1005)	Pert(min=0.096, mode=0.1964, max=0.2033)
1 cfu/mL	percent reduction	Pert(0.0016, 0.0194, 0.0392)	Pert(0.0011, 0.0138, 0.0383)	Pert(0.04, 0.1852, 0.2252)
10 cfu/mL	percent reduction	Pert(0.0003, 0.0083, 0.03)	Pert(0, 0.0021, 0.0083)	Pert(0.0148, 0.1389, 0.2267)
100 cfu/mL	percent reduction	Pert(0.0001, 0.0017, 0.0186)	Pert(0, 0.0002, 0.0008)	Pert(0.0042, 0.0817, 0.2189)

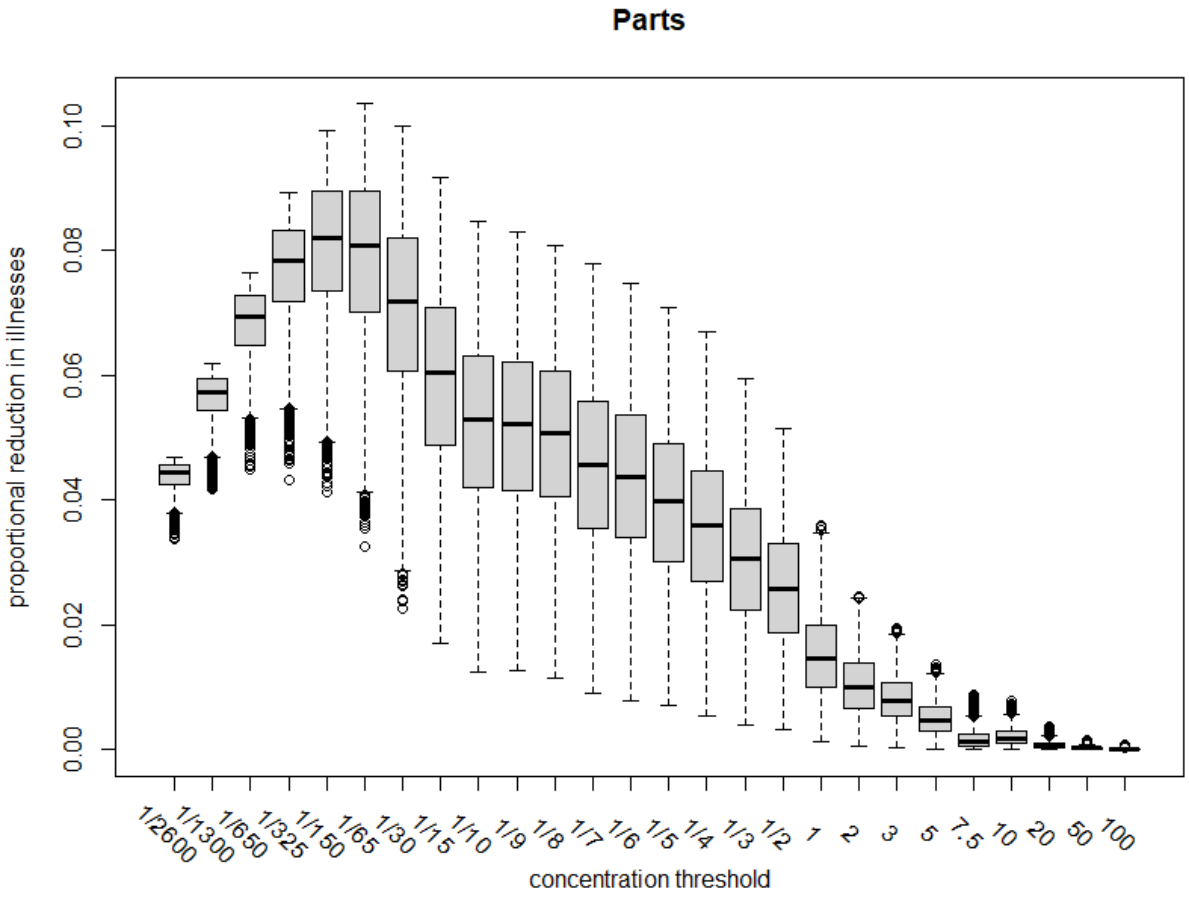


Figure 30: Box and whiskers plot of the proportional reduction in illnesses for chicken parts across a range of concentration thresholds is shown. The mode of the pert distribution for each concentration threshold is that predicted by the full (default) model while the lower bound is estimated by assuming the attenuation distribution mean is +2 log10 and the upper bound is based on an approximation that assumes linearity of the dose-response functions.

Following Monte Carlo simulation (1 million iterations) of the product of the proportional reduction and the starting annual number of illnesses, we estimate distributions for annual illnesses prevented at the concentration thresholds of interest. These results suggest substantial overlap in the 95 percent credibility intervals across progressively higher concentration thresholds. For example, the comminuted chicken credible intervals are very similar across all the concentration thresholds while, for chicken carcasses, these intervals overlap for the 0.03 cfu/mL to 10 cfu/mL concentration thresholds. Overlapping credible intervals suggest that differences in the most likely effectiveness between different concentration thresholds may not be meaningful.

Table 39: Estimated annual illnesses prevented by final product concentration standards are shown for *Salmonella* threshold levels of interest. Values are rounded to nearest 100 illnesses.

Annual illnesses prevented, most likely (95% credible interval)

Threshold level	Chicken carcasses	Chicken parts	Comminuted chicken
0.03 cfu/mL or /g	4600 (2000, 7100)	7900 (3300, 12700)	1500 (800, 2200)
1 cfu/mL or /g	2400 (700, 5000)	1400 (400, 3600)	1400 (600, 2100)
10 cfu/mL or /g	1000 (200, 3100)	200 (40, 700)	1000 (400, 1900)
100 cfu/mL or /g	200 (0, 1500)	20 (0, 100)	600 (200, 1500)

5.4 Discussion

Predictions made in this risk assessment are average effects. Across repeated events, we would expect the average effect to be similar to our predictions. However, this model does not demonstrate the variability between events (e.g., diversion of units) or across establishments within an industry, or across time. This variability might contribute to substantial fluctuations in actual illnesses prevented, if our foodborne illness surveillance was sensitive enough to detect changes in actual illnesses attributed to individual products (Ebel, 2017).

As expected, our analysis demonstrates that higher levels of *Salmonella* on raw products are associated with higher risk of illness, on average, compared with lower levels. Nevertheless, the level and its frequency are inversely related; lower levels occur much more frequently than higher levels.

With respect to level-based final product standards, this analysis suggests that increasing the level threshold above the current LOD results generally in a lower fraction of lots being identified as failing the standard and, correspondingly, a smaller number of illnesses prevented. The comminuted chicken analysis illustrates a counter-intuitive effect, where increasing the level threshold above the current LOD results in a lower fraction of lots being diverted, but the number of illnesses prevented actually increases. This effect only occurs when the level threshold is slightly increased and can be explained as the result of not diverting lower risk lots and replacing them with higher, but still average, risk lots.

With respect to serocluster final product standards, this analysis suggests that removing the most virulent *Salmonella* serotype cluster should decrease human illnesses. Nevertheless, any risk management option must first detect a lot as *Salmonella*-positive before the serotypes of concern can be identified. Furthermore, similar to final product levels, the most virulent serogroups are generally the least frequently detected. Therefore, the requirement to detect and identify virulent serotypes necessarily limits the effectiveness of a test and diversion option based on serogroups relative to an option that diverts only *Salmonella*-positive lots.

Some of the modest reductions in illnesses are the result of treating FSIS's current sample collection and laboratory resources as an application of a lot-by-lot testing scheme used to set a microbiological criterion for foods (FAO/WHO, 2016; Whiting, 2006). This can be illustrated by considering that the maximum possible illness reduction for a theoretical level threshold and a replacement of diverted lots by passing lots (**Table 40**) does not achieve the HP2030 25% reduction in salmonellosis illness goal. While the necessary samples per lot or flock is application dependent, the sample sizes for these applications typically range from between 4 and 14 (Skov, 1999; Whiting, 2006), which would require the collection of more than 1.6 million samples for carcass testing alone.

Table 40:The maximum benefit of replacing diverted lots by passing lots.

Chicken Product	**Carcasses**	**Parts**	**Comminuted**
Maximum Illnesses Reduction Concentration	1 cfu/150mL	1 cfu/2600mL	1 cfu/2600g
Illnesses Prevented	5,000 (4%)	14,000 (13.4%)	1,740 (23%)

Direct and Indirect Effects

This analysis has focused only on the direct effects of final product standards. Following implementation of such standards, it is reasonable to expect that industry risk-managers might respond by taking actions to reduce the likelihood that their products would fail the standards. Nevertheless, this risk assessment cannot predict the potentially substantial effects of these actions. Therefore, this analysis can only attempt to measure the direct effects of identifying and removing tested lots that fail a standard.

In contrast to the direct effects predicted here, prior FSIS risk assessments for *Salmonella* performance standards have focused entirely on indirect effects of testing programs. These performance standards only serve to identify establishments (not lots) as passing or failing those standards. Furthermore, there is no mandatory enforcement action once an establishment fails a performance standard. Risk assessments of performance standards attribute improvements in public health to the actions taken by failing establishments to become passing establishments (FSIS, 2015). Those assessments assume that the motivation for improvement is via market forces that penalize failing establishments and/or reward passing establishments. Therefore, failing establishments and public health improve as an indirect effect of the performance standards.

Chapter 6 Receiving Guidelines

The first risk management option states:

> What is the public health impact (change in illnesses, hospitalizations, and deaths) achieved by eliminating at receiving a proportion of chicken contaminated with specific levels of *Salmonella* and/or specific *Salmonella* subtypes?

As stated previously, FSIS does not have regulatory discretion in the pre-harvest environment, nor does it routinely collect data on the nature of flocks that are presented for slaughter—where FSIS jurisdiction begins. Specifically, FSIS does not have data on the *Salmonella* serotypes present on live birds, nor the *Salmonella* contamination levels. Further, FSIS does not have robust, generalizable data on the types of pre-harvest interventions, such as vaccination, employed by the live bird industry. FSIS worked collaboratively with the UMD-JIFSAN through a Cooperative Agreement to obtain industry-shared data. At the time of writing, these efforts, while effective in laying the groundwork for future data sharing endeavors and have not been fruitful in producing new pre-harvest data for use in this risk assessment.

As a result, receiving guidelines are modeled using the rehang sample data.

Given that there is no reliable information to parameterize a level distribution, risk management options that address *Salmonella* levels will not be assessed. While no specific information is available for explaining the occurrence of serotypes in flocks, some simplifying assumptions allow for an approximation of the effect of removing flocks where a larger fraction of highly pathogenic serotypes are identified.

6.1 Maximum Potential Benefit of Receiving Guidelines

Our model for final product standards provides a crude description of the effect of incrementally reducing the share of all *Salmonella* that are C1 (**Figure 31**). If the model parameter describing the fraction of *Salmonella* that are C1 (c) is progressively reduced from its current value of 20 percent to zero percent, we calculate a linear reduction in chicken illnesses that extends from 0 to about 55,000 illnesses prevented per year. At the extreme, these results demonstrate that C1 *Salmonella* constitute about 44 percent of the 125,515 annual human illnesses attributed to chicken consumption.

Alternatively, the final product standards model for serogroups introduces the concept of two schemes for contaminated lots (Scheme 1 has a high proportion of C1 while Scheme 2 has a low proportion). The baseline model assumes that nearly 15 percent of lots are in Scheme 1 ($P(S1)$ in our analysis). Reducing this fraction mimics the broad effect of eliminating lots or flocks with a high proportion of C1 and replacing these with lots or flocks having a low proportion of C1 via some pre-harvest strategy (e.g., vaccination). In this case, however, the maximum effect of this reduction is about 27,000 illnesses prevented (22 percent of all chicken illnesses). This effect is more limited, relative to the prior scenario, because the frequency of C1 in Scheme 2 is still almost 10 percent.

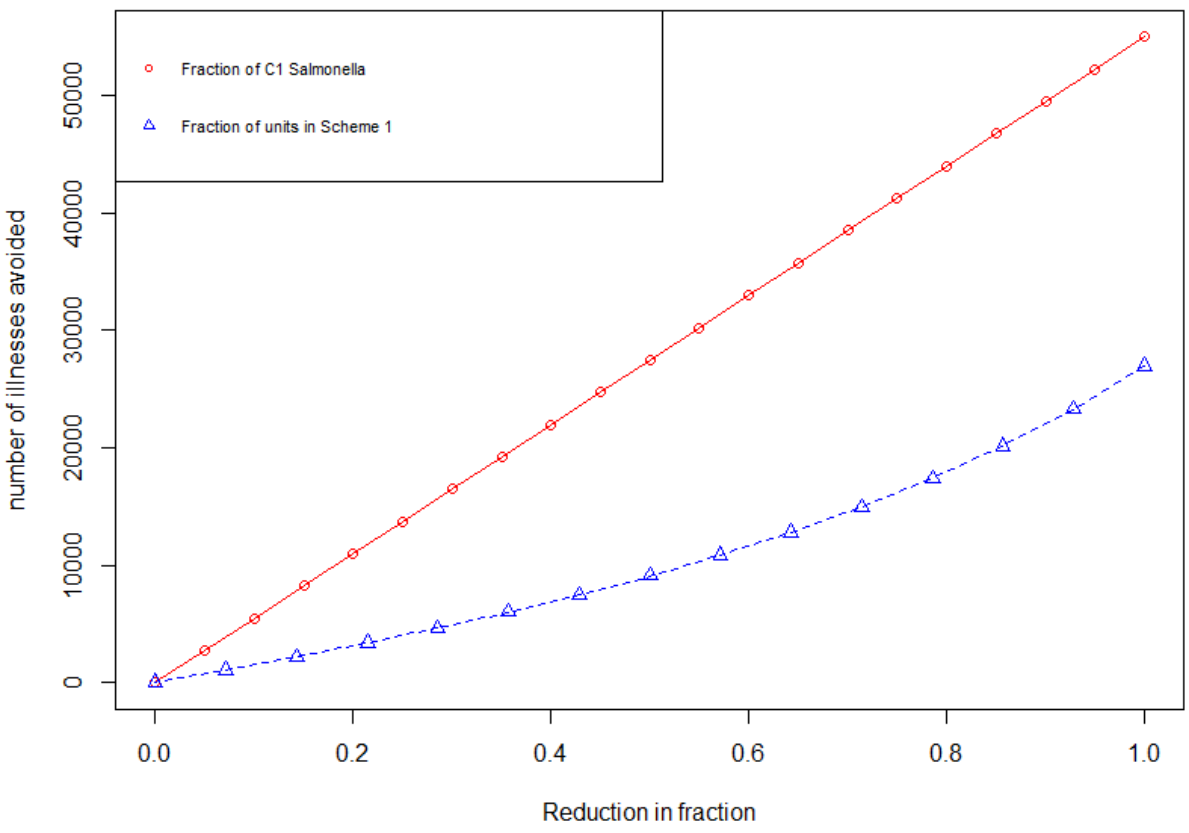


Figure 31: An illustration of the broad effects of reducing the total fraction of C1 *Salmonella*, or the fraction of flocks in Scheme 1, is shown.

Although the actual share of C1 *Salmonella* across live chicken flocks is unknown, it is likely to vary beyond the two schemes assumed for the final product standard. Therefore, the remainder of this section assesses how flock sampling could be used to inform receiving guidelines.

6.2 Flock Sampling and Receiving Guidelines

The first assumption of this analysis is that all flocks contain a mixture of serotypes, some of which belong to the higher virulence cluster denoted by $C1$and with proportion F_{C1}. The second simplifying assumption is that the serotype observed for each tested carcass represents the only serotype on the carcass (or at least that the observed serotype is sufficiently dominant as to explain the majority of the probability of illness given exposure). While the first assumption is reasonable, the validity of the second assumption is questionable (Thompson, 2018) and will lead to estimates of the potential reductions in illness that are likely larger than what would occur in practice.

Addressing the effect of removing lots that contain a higher fraction of the more virulent serotypes

requires defining how the distribution of the fraction of chicken carcasses whose dominant serocluster, F_{C1}, varies across the roughly 204,000 flocks that constitute the annual chicken production of the roughly 200 slaughter establishments that are subject to FSIS performance standards in the U.S. This distribution is constructed by assuming that the serotypes identified in the 168 samples that were positive at both rehang and post-chill are indicative of the underlying distribution of F_{C1}across the population of flocks. **Table 42** describes the agreement between the serotypes in the high- and low-virulence clusters such that the fraction of cluster 1 serotypes is not significantly different at either location (i.e., 0.202 and 0.208 at rehang and post-chill, respectively). Noting this similarity, it is assumed that the average ofF_{C1}across the population of flocks is 0.205.

Establishment Production Volume

A three-way (2x2x2) contingency table was constructed by discretizing the paired dataset (**Table 26**) based on production volume. That is, the data was classified by serotype cluster at rehang, cluster at post-chill, and establishment production volume. Using a threshold of 10 million birds slaughtered, the three-way table can be displayed as a set of two-way tables (**Table 41**) for high volume and low-volume establishments.

Let n_{ijk} be the number of samples for which rehang is identified by *i* (1=cluster 1, 2=cluster 2), post-chill is *j* (1=cluster 1, 2=C luster 2), and establishment size is *k* (1=high, 2=low). Additionally, let "+" indicate the summation over one variable such that $n_{ij+} = \sum_k n_{ijk}$, for example. Thus, there are $n_{++1} = 93$. samples from high volume establishments and $n_{++2} = 75$. samples from low-volume establishments. The proportion of cluster 1 serotypes observed at rehang and post-chill are roughly similar (averaging around 0.2) for high volume and low-volume establishments. Furthermore, the concordance rates (i.e., t rate at which the same cluster was detected at both rehang and post-chill) were comparable by establishment size. Paired samples in low-volume establishments detected the same cluster at rehang and post-chill at a rate of 0.853 whereas paired samples from high volume establishments observed a concordance rate of 0.806.

The relative agreement in these descriptive metrics suggests that the underlying cluster mixtures at rehang and post-chill are reasonably equivalent regardless of production volume. First, consider a trivial model that all cells in the three-way table are equivalent to their expected values. In two-way tables, the expected value for each cell is estimated by $E_{ij} = \frac{n_{i+} n_{+j}}{n}$, whereas in three-way tables, the expected vues are estimated by $E_{ijk} = \frac{n_{i++} n_{+j+} n_{++k}}{n^2}$ . . The chi-square test can then be used to measure the divergence of the observed from the expected. Here, the null is rejected ($\chi^2 = 38.58, p = 8.519e-08$.)Noting that for the chi-squared approximation to work well, the expected values are required to be sficiently large (e.g., most values greater than 5). In this case, 80% of the expected values are above this threshold (**Table 41**), and thus, rehang cluster, post-chill cluster, and establishment size are not mutually independent. Secondly, a joint independence model testing whether rehang and post-chill cluster results are jointly independent of establishment size resulted in the failure to reject the hypothesis ($\chi^2 = 2.2445, p = 0.5232$). That is, there was no evidence that differences between the clustering results by establishment size exist.

Furthermore, the conditional log odds ratio which allows one to test for conditional independence of two variables, given a third, was considered. Controlling for establishment size, the conditional log odds ratios for high volume and low-volume establishments are 2.213 (1.070, 3.356) and 2.872 (1.456, 4.289), respectively, indicating that rehang and post-chill are conditionally associated and the odds for detecting matching clusters on positive samples at rehang and post-chill is significantly higher than positive samples at rehang and post-chill not matching on the cluster.

Table 41: Contingency tables for establishments by volume. Expected counts are indicated in parentheses.

Small Establishments					**Large Establishments**				
		Post-chill					**Post-chill**		
		Cluster 1	Cluster 2	Total			Cluster 1	Cluster 2	Total
Rehang	Cluster 1	9 (3.16)	7 (12.02)	16	**Rehang**	Cluster 1	11 (3.92)	7 (14.90)	18
	Cluster 2	4 (12.46)	55 (47.36)	59		Cluster 2	11 (15.45)	64 (58.72)	75
	Total	13	62	75		Total	22	71	93

Describing the Fraction of a Flock Contaminated by the Higher Virulence Serocluster

Given the assumed similarity in the distribution of F_{C1} across the industry, the next step is to model a distribution for this parameter. Assume that $F_{C1} \sim Beta(a,b)$, where the parameters are such that $a/(a+b) \approx 0.205$, and the agreement between single samples selected at rehang and post-chill would match the values observed in the contingency table for testing across all establishment sizes (**Table 42**).

Table 42: Contingency table for all establishments.

		Post-chill		
		Cluster 1	Cluster 2	Total
Rehang	Cluster 1	20	14	34
	Cluster 2	15	119	134
	Total	35	133	168

The parameters of the distribution were estimated using a Bayesian Monte Carlo model (Smith, 1992) with a weighting function. The function for the resampling weights used the product of the χ^2statistic (describing the agreement between the contingency table derived from the Monte Carlo draw and the observed data) and a *Z*-statistic (testing the agreement between the Monte Carlo estimate of F_{C1} and the observed value).

Prior distributions for the parameters of the beta distribution were $a \sim Exponential(1)$ and $b \sim Exponential(0.25)$. These weakly informative priors were chosen to explore the range of plausible parameter values. The posterior distribution of the a and b parameters represent the most likely values for the distribution explaining the fraction of more virulent serotypes within each flock (**Figure 32**).

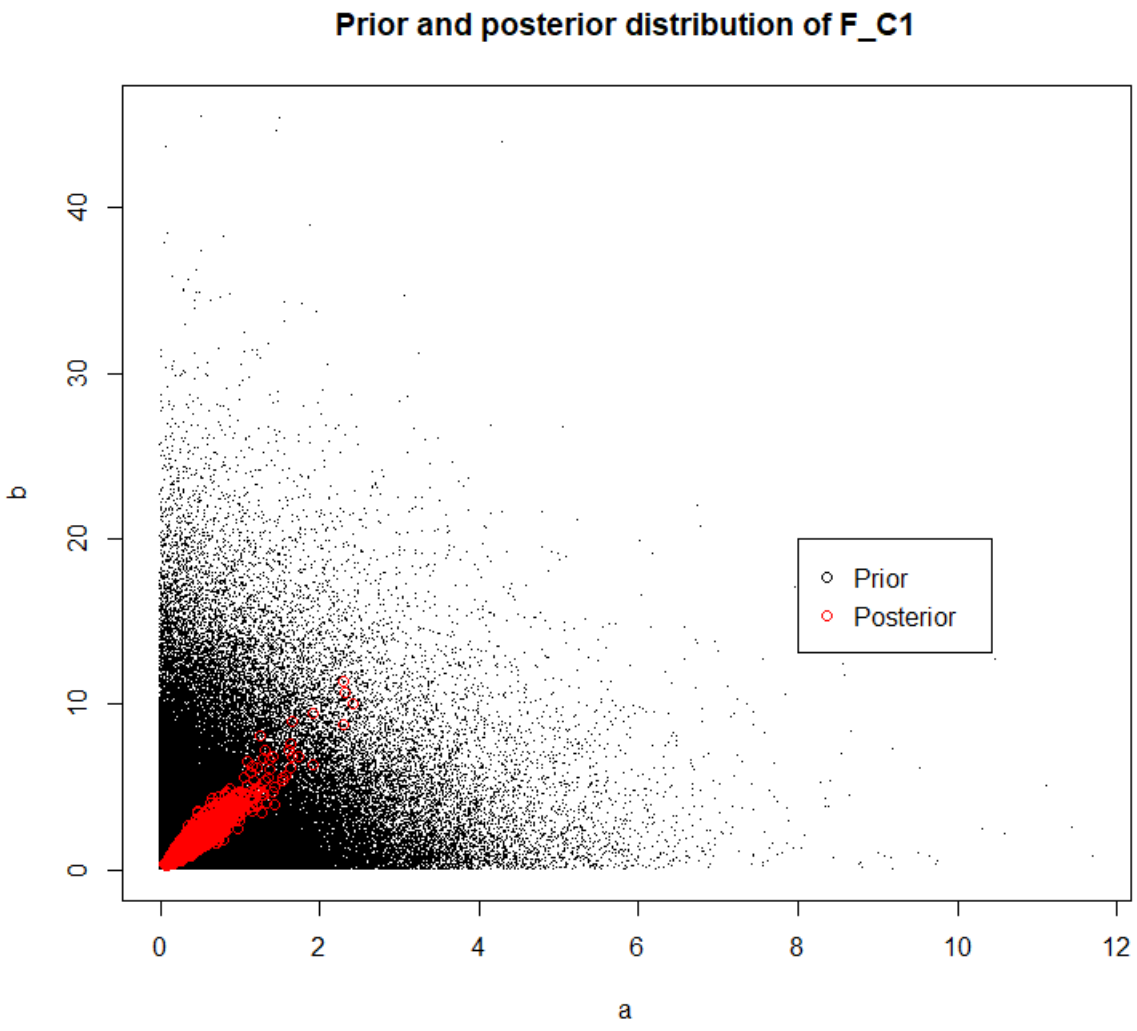


Figure 32: Comparison of the prior and posterior distributions for the proportion of higher virulence serotypes (cluster 1) within each flock.

The parameters of the most likely distribution are $a = 0.37$and $b = 1.53$, with the cumulative distribution given in **Figure 33**. The mean of the posterior is $\bar{F}_{C1} = 0.201$with corresponding contingency

table values in **Table 43**.

Table 43: Contingency table simulated from Monte Carlo sampling of the posterior distributions for the proportion of higher virulence serotypes (cluster 1) within each flock.

		Post-chill		
		Cluster 1	Cluster 2	Total
Rehang	Cluster 1	18.6	14.9	33.4
	Cluster 2	15.6	119.0	134.6
	Total	34.2	133.9	

The cumulative distribution function finds that most flocks have an estimated fraction of higher virulence serotypes that is less than 20%. Specifically, about 51% of flocks have less than 10% of carcasses whose dominant serotypes is a member of the higher virulence cluster. On the upper end of the distribution, 96% of flocks have less than 75% of carcasses testing positive for a higher virulence serotype (i.e., only ~4% of flocks have greater than 75% of carcasses predicted to test positive for higher virulence serotypes).

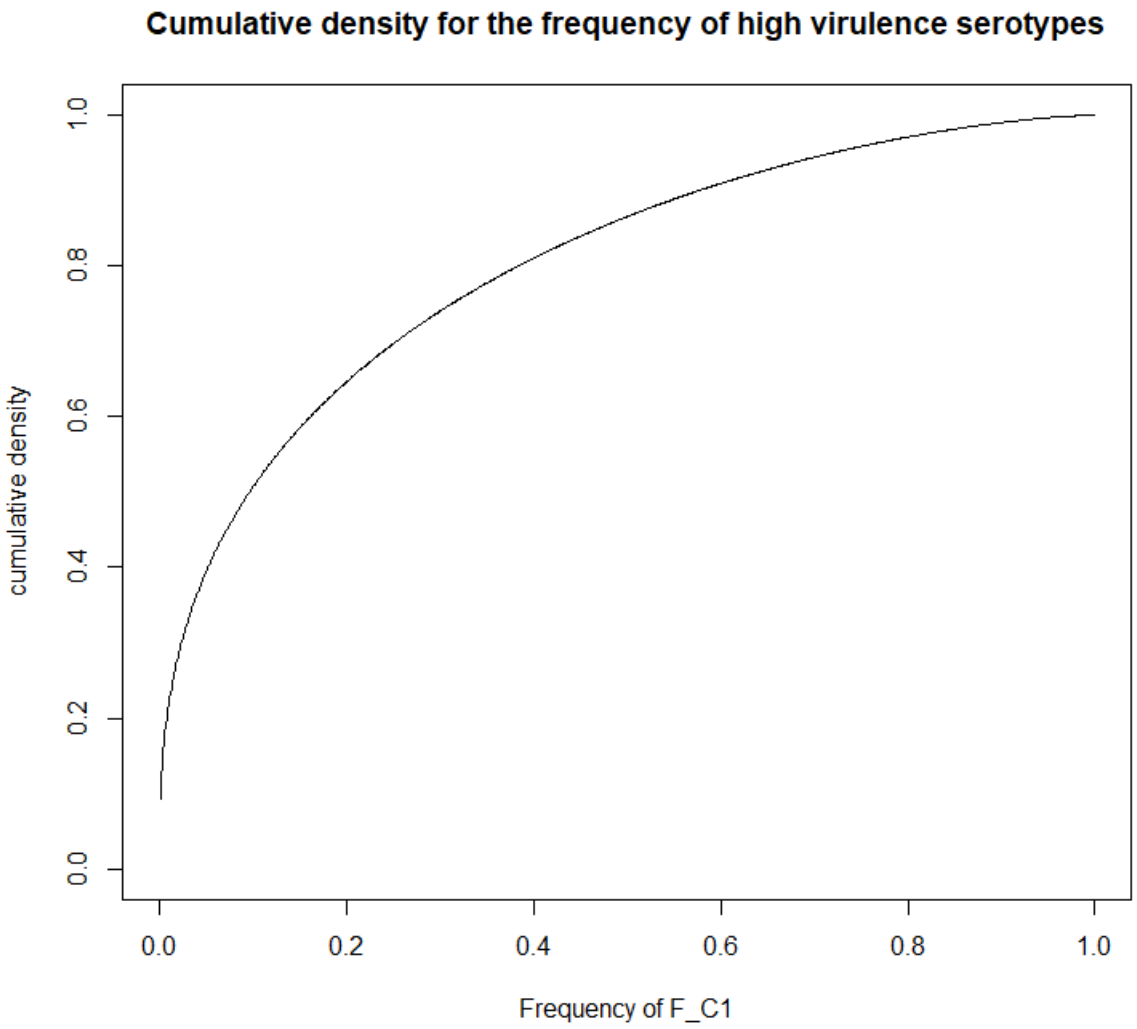


Figure 33: Cumulative density of the distribution describing the fraction of each flock that is contaminated with higher virulence serotypes.

6.3 Proposed approach for identifying higher risk flocks

This receiving guideline risk management question specifies that flocks would be eliminated at receiving, but we assume that diversion of a flock at receiving would not be possible because of the stress and elevated mortality associated with both transporting birds and delaying entry to slaughter would likely be contradictory to FSIS' humane handling requirements (FSIS, 2018). Therefore, it is assumed that all testing would need to occur prior to a flock's preparation for transport (i.e., prior to feed withdrawal, etc.). It is also assumed that testing would be done on individual live birds, rather than the typical boot sampling. This assumption is necessary so that the linkage between individual carcasses and serotype occurrence can be determined. An example of such a scheme is the testing of dead-on-arrival birds described in a recent study (De Villena, 2022).

The model will assume that each flock will be tested for the presence of cluster 1 serotypes using

samples size of 1 and 4. The first sample sizes were chosen because it is customary in the U.S. to base decisions on a single sample. The sample size of 4 birds was chosen because this sample size would be similar to other attributes sampling plans for assessing the microbiological criteria in foods (Anderson, 2016).

This model assumes a rule for accepting a flock is the occurrence of cluster 1 serotypes being either $s = 0$ samples in cluster 1 out of $n = 1$ or $s \leq 1,2,3$ cluster 1 samples out of $n = 4$ samples for each flock. The average flock rejection rate for each of these sampling plans is illustrated in **Figure 34**. If at least one sample is identified as a cluster 1 serotype out of 4 samples, then the average rejection rate is approximately 0.416 (95% CI 0.346, 0.487). Requiring stricter rules yields lower flock rejection rates; namely, if cluster 1 occurs in at least 2, 3, or 4 of 4 samples, then the rejection rate averages 0.223 (95% CI 0.165, 0.285), 0.110 (95% CI 0.069, 0.157), and 0.039 (95% CI 0.019, 0.065), respectively. In the single sample design (i.e., 1 of 1 found to be a cluster 1 serotype), flocks are rejected at a rate of 0.197 (95% CI 0.153, 0.244).

Performance metrics are investigated for each sampling scheme rule. In particular, interest lies in the misclassification rate of high-risk flocks containing more than a particular threshold of cluster 1 prevalence (e.g., 0.05, 0.10). **Figure 34** through **Figure 36** depict the change in sensitivity (Se) and specificity (Sp) based on different cluster 1 risk classification thresholds. For example, if the fraction of cluster 1 in a sampled flock is greater than 5%, the sensitivity of the rejection rule is high (>0.95) for each sampling program while the specificity varies between 0.44 and 0.64. Raising the threshold to define high risk flocks, the sensitivity for each program starts to drop while observing increases in specificity. The performance is maximized in the 1 of 1 sampling program when using a 0.20 cluster 1 risk threshold (Se = 0.82, Sp = 0.76). In the multiple sample programs, employing stricter rejection rules (i.e., increasing minimum cluster 1 detections) generally exhibits better performance characteristics.

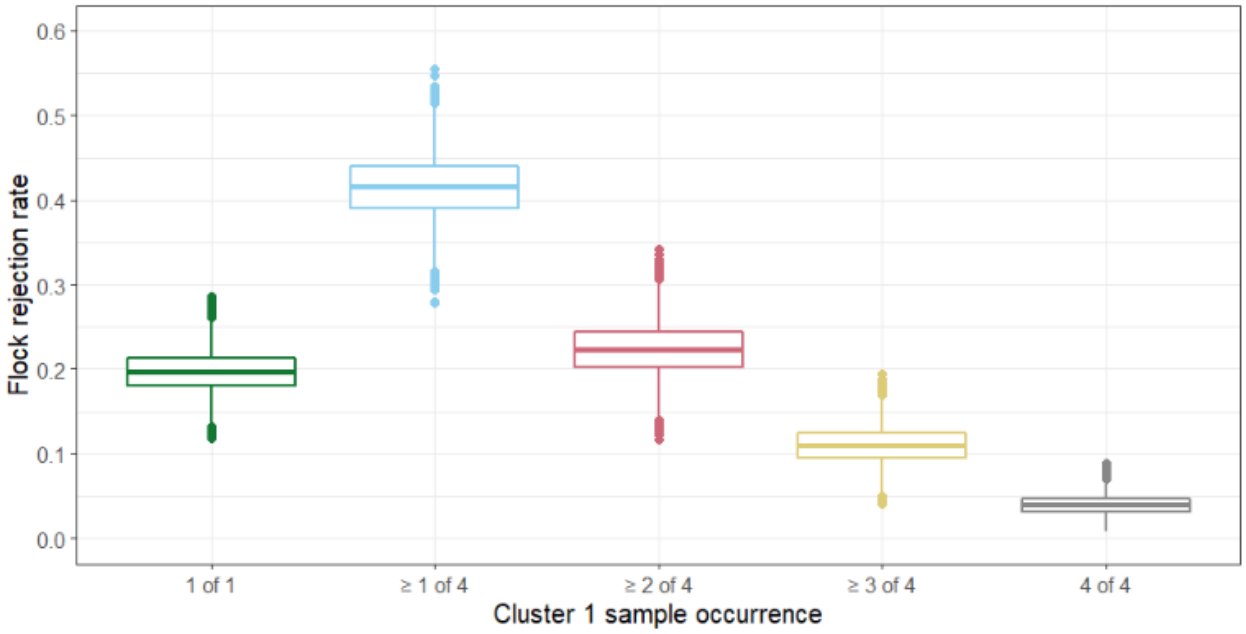


Figure 34: Simulation results regarding flock rejection rate based on cluster 1 occurrence in the

implemented sampling plan.

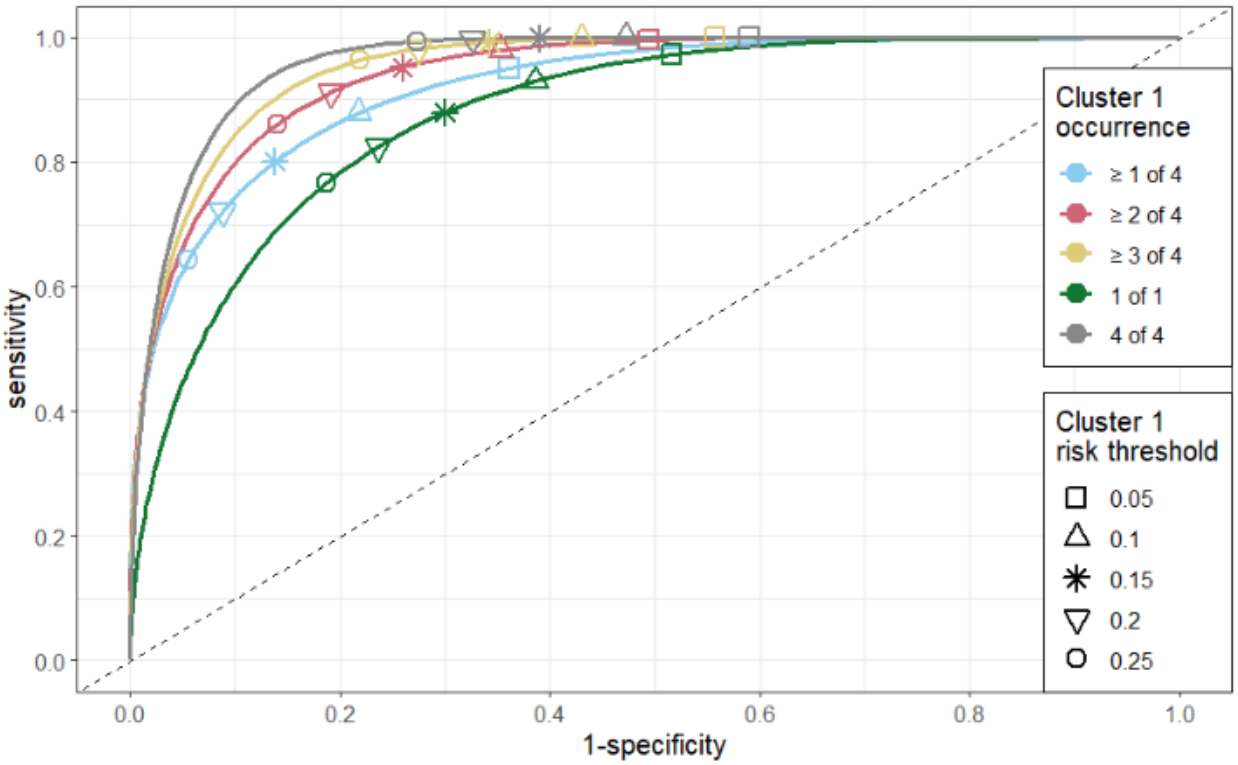


Figure 35: Performance characteristics of each sampling program and rejection rule.

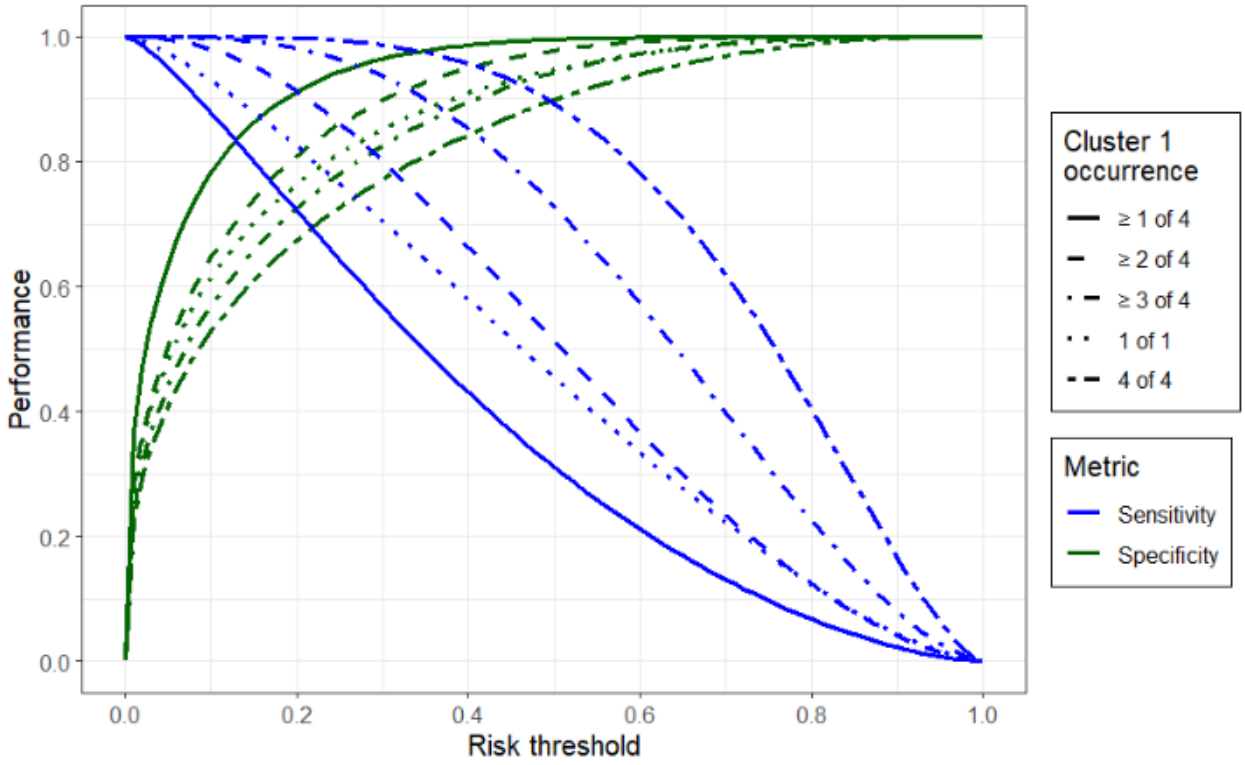


Figure 36: Performance metrics by risk threshold for each sampling program and rejection rule.

6.4 Modeling the impact of risk management options

Assessing the total impact of different risk management options is complicated by an inability to predict the fate of rejected flocks. If the number of flocks rejected is small, they may be diverted to rendering. Conversely, if the number of flocks is large, the industry may divert all flocks to cooked product or designate establishments that pasteurize all product. Given these unknowns, the simplest metric to assess is the reduction in the overall probability of illness per serving as a consequence of diverting flocks that are classified as failing under the proposed testing strategies.

The other metric of interest is the number of flocks diverted, denoted by Div, as a consequence of the option. This value is determined using Monte Carlo simulation.

The baseline probability of illness prior to the implementation of a new standard is estimated by

$$P_{baseline}\left(ill\right) = \left(F_{C1.baseline} \times P(ill \mid Cluster1) + \left(1 - F_{C1,baseline}\right) \times P(ill \mid Cluster2)\right),$$

where

$$P(ill \mid Cluster1) = \int R_{C_1}\left(d\right) f\left(d\right) \partial d \text{ and } P(ill \mid Cluster2) \text{ is defined similarly.}$$

It is assumed that the effect of the reduction in the occurrence of highly virulent serotype propagates

through all poultry products, so $f(d)$ is the lognormal distribution with parameters $\hat{\mu}_{all} = -3.037$ and $\hat{\sigma}_{all} = 1.28$ that describes the overall level per gram of *Salmonella* at the end of production.

For each flock rejection rule, the resulting fraction of carcasses in the higher virulence cluster estimates $F_{C1,new}$ and the new probability of illness is estimated as

$$P_{new}(ill) = (F_{C1,new} \times P(ill \mid Cluster1) + (1 - F_{C1,new}) \times P(ill \mid Cluster2).$$

The proportional reduction in the probability of illness is given as

$$Prop.reduction = \left(1 - \frac{P_{new}(ill)}{P_{baseline}(ill)}\right).$$

6.5 Results

Table 44: Results of pre-harvest risk management options.

Options	n (number of samples taken)	s (number of samples C1 positive)	$F_{C1,new}$(the posterior dist)	Annual flocks diverted (rejection rate)	Proportional reduction in probability of illness
Option 1	1	1	0.105	41,000 (0.20)	0.23
Option 2	4	1	0.037	85,000 (0.42)	0.40
Option 3	4	2	0.082	46,000 (0.23)	0.29
Option 4	4	3	0.128	23,000 (0.11)	0.17
Option 5	4	4	0.170	8,000 (0.04)	0.07

As displayed in **Table 30** and **Figure 34**, there is relatively no difference between the first and third options, because they both allow for roughly 10 percent of the flocks to be rejected. As the required number of samples classified as cluster 1 increases, a smaller fraction of flocks are rejected, with these flocks being more highly contaminated with the cluster 1 serotypes. Nevertheless, a substantial fraction of all carcasses that would be classified as cluster 1 remain amongst the large fraction of flocks that were classified as containing predominantly cluster 2 serotypes.

6.6 Summary

This analysis sought to explore the potential public health impacts of new guidelines at the receiving step at chicken slaughter establishments. Given the lack of pre-harvest data, FSIS developed scenarios using data the Agency had access to; namely chicken rehang data.

Further, due to limited serotype data at the post-chill step, among other technical issues, this analysis does not estimate the public health impact of eliminating a portion of chicken carcasses that are

contaminated with a specific level of *Salmonella*. Rather, this analysis focuses on the public health impact of eliminating specific *Salmonella* serotypes at rehang, with a specific focus on the serotype clusters discussed in the Final Product Standards chapter above.

However, estimates of the public health impact of new receiving guidelines could be modeled using the serotype cluster-based approach described in the Final Product Standard chapter of this risk assessment. Guidelines were modeled as a 'flock rejection' standard; such that any incoming flock tested at rehang would be diverted if higher virulence serotypes—cluster 1—were identified in the sample.

A flock rejection standard that requires diverting every flock that tests positive for cluster 1 at rehang has a potentially substantial public health impact. Notably, direct effects of a flock rejection standard are sufficient to meet the HP2030 goals of a 25% reduction in *Salmonella* illnesses from chicken when the standard is sufficiently stringent (i.e., flocks are rejected when any cluster 1 serotypes are found). These results should serve as proof-of-concept approach to illustrate the strengths and potential weaknesses of a flock diversion risk management option.

Chapter 7 Process Control

The third risk management question is:

> What is the public health impact of monitoring/enforcing process control from rehang to post-chill? Monitoring could include analytes such as Enterobacteriaceae, Aerobic Plate Count, or other indicator organisms, analysis could include presence/absence or levels and the monitoring could also include variability of actual result versus expected result, log reduction, absolute sample result, or other individual establishment specific criteria.

The utility of measuring levels of indicator organisms at a single location in the slaughter process has been studied previously and found to be of limited effectiveness for controlling pathogenic bacteria (Altekruse, 2009a), so multipoint sampling (e.g., rehang and post-chill) is required in the evaluation of process control, and thus, not assessed for parts and comminuted product. We address this question by analyzing the log reduction and presence/absence of two indicator organisms that can be effectively measured on chicken carcasses at different stages of the slaughter process, namely, AC and Enterobacteriaceae (EB).

Essentially, we analyze the scenario of underperforming establishments that adjust practices toward meeting a level of control according to indicator organism metrics such as AC log reduction and AC elimination from rehang to post-chill. That is, by setting a log reduction or presence fraction target/guideline, we consider the overall prevalence that results.

7.1 Data Description

For each sample in FSIS' Young Chicken Carcass Exploratory Sampling Program (2022), the level of AC and EB was estimated. The limits of both detection (LOD) and quantification (LOQ) for the quantitative PCR assay employed is 10 cfu/mL. FSIS has previously assessed the utility of using either AC and EB using data from FSIS microbiological baseline studies (FSIS, 2009b; FSIS, 2012) and concluded AC is a more practical indicator organism because levels are more likely to be above the LOD (Williams, 2015; Williams, 2017). **Table 45** compares the proportions of samples that are above the LOD at both sampling locations for the 2022 FSIS Exploratory Sampling ((FSIS, 2022d)) and the data collected in the previous FSIS chicken carcass baseline study (FSIS, 2009b). The percentage of samples above the LOD at both time periods is similar at rehang. The percentages of samples above the LOD at post-chill is much lower for both classes of indicator organisms in the current data, suggesting that the industry has increased the degree of removal and inactivation of all bacteria. The lack of enumerated EB samples at post-chill makes this organism inappropriate for assessing the degree of reduction in this organism between different locations in the slaughter process. For this reason, there was no further analysis of the data collected for this organism.

Table 45: Percentage of samples whose indicator organism level was above the LOD.

Indicator	Rehang 2007-2008	Rehang 2022	Post-chill 2007-2008	Post-chill 2022
Aerobic Count	99.9	99.8	97.1	70.0
Enterobacteriaceae	99.7	98.7	57.4	16.1

Figure 37 summarizes the AC levels for all samples at rehang and post-chill, where the orange vertical lines depicting the mean AC levels at both sampling locations. These values are 4.40 at rehang and 1.39 at post-chill, for an average log reduction of 3.01. It is insightful to compare these values to those of the 2007-2008 FSIS microbiological baseline study, where the average AC levels were 4.50 at rehang and 2.46 at post-chill, for an average log reduction of 2.04. This demonstrates that while there have been essentially no changes in the incoming AC loads on chicken carcasses, the additional processing interventions implemented in the last 15 years are achieving about an additional 1-log reduction (i.e., on average only 1 aerobic bacterium out of every 1000 is surviving between rehang and post-chill, as compared to 1 out of 100 in the 2007-2008 period). In a previous study (Williams, 2015) log reductions in AC, generic E. *coli (GEC)* and *Salmonella* were similar in magnitude (2.04, 2.3 and 2.08, respectively), so it is reasonable to expect reductions of roughly 3 logs in *Salmonella.*

In previous studies of indicator organisms, a correlation was observed between an establishment's reduction in the average levels of AC between rehang and post-chill and the occurrence of GEC and pathogenic bacteria, with establishments that had larger reductions tending to have both lower levels and occurrence of pathogens (Williams, 2015; Williams, 2017). When indicator organisms are present in nearly all samples, the average reduction is simply calculated as

$$\Delta_{APC,j} = \frac{1}{n_j}\sum_{i=1}^{n_j} y_{rh,i} - \frac{1}{n_j}\sum_{i=1}^{n_j} y_{pc,i},$$

Where $y_{*,i}$ is the level of sample *i* from establishment *j* at either rehang (rh) or post-chill(pc). When the fraction of samples that are below the LOD is small, as was the case in the 2007-2008 microbiological baseline study, simple ad hoc adjustments such as substituting ½ of the LOD for samples where AC was not found, are reasonable solutions. Similarly, some samples can have levels that exceed the assay's ability to enumerate the sample. When this occurs infrequently, the ad hoc solution of substituting twice the upper limit of quantification is reasonable. As demonstrated in **Table 45**, almost all rehang samples are above the LOD, but a much larger fraction of post-chill samples had AC levels below the LOD (**Table 45**). Simple ad hoc adjustments in this case are inappropriate and create large biases in the level estimates, unless used sparingly (Helsel, 2009; Helsel, 2010).

To account for the large fraction of samples below the LOD, a maximum likelihood routine was used to fit a lognormal distribution to the data for each establishment, while accounting for the censored observations (Williams, 2014). The $\hat{\mu}_j$ derived from the fitted distribution is used to estimate the average AC level for every establishment. The average log reduction, accounting for the censored data, is estimated as

$$\Delta_{APC,j} = \frac{1}{n_j}\sum_{i=1}^{n_j} y_{rh,i} - \hat{\mu}_j .$$

Figure 38 compares the $\hat{\mu}_j$ values from each establishment to the estimated mean when a value of ½ the LOD (i.e., 5) is substituted for all samples where the estimated AC value was <10. The largest discrepancies occur for establishment where the average post-chill AC levels are low. There is no difference between the estimated means when all post-chill samples are above the LOD. Note that the high-volume establishments that also belong to large corporations tend to have significantly lower average AC levels at post-chill.

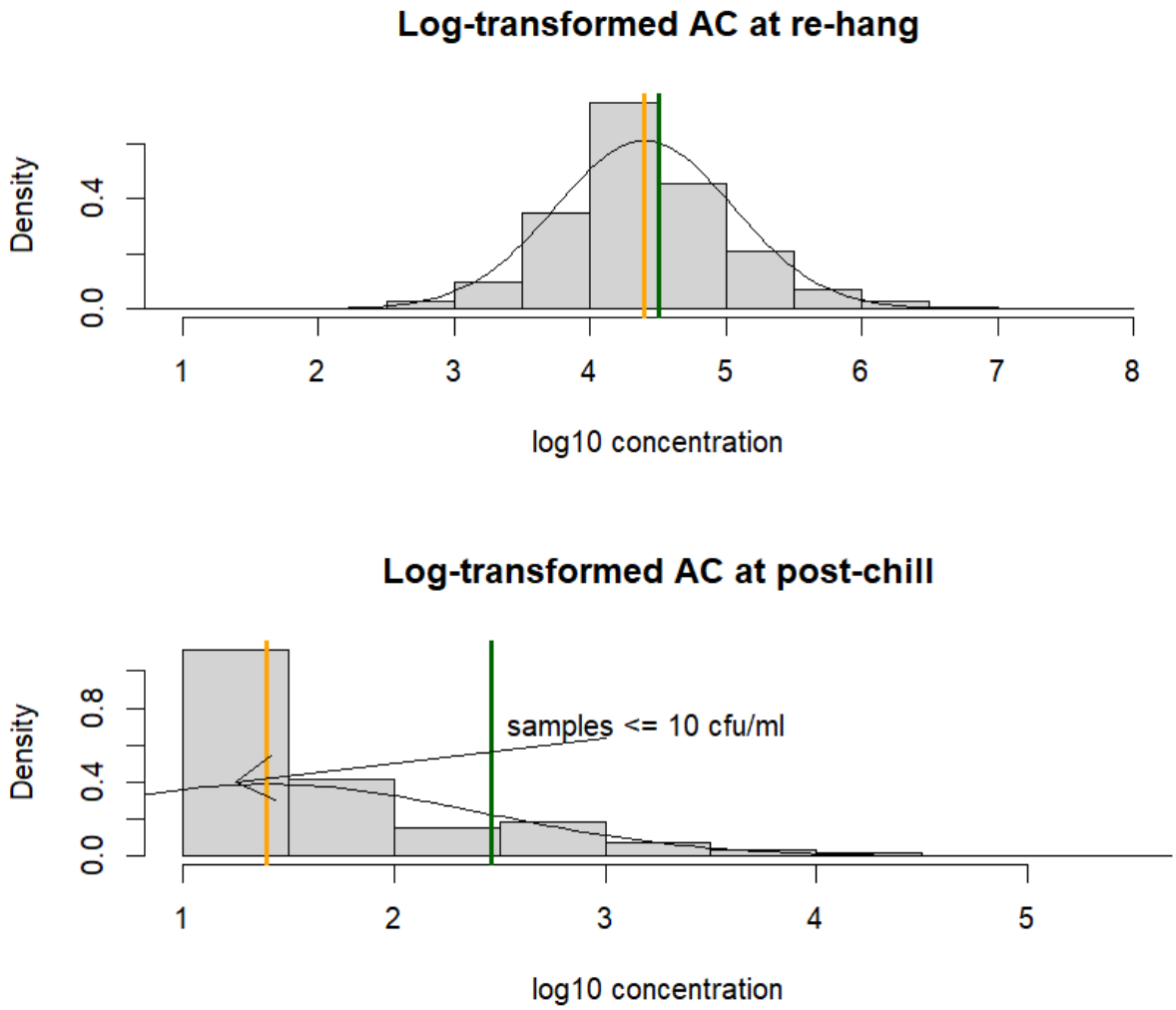


Figure 37: Distribution of AC level at both rehang and post-chill. The orange vertical lines represent the current mean level, while the green lines represent the mean level from the previous microbiological baseline study.

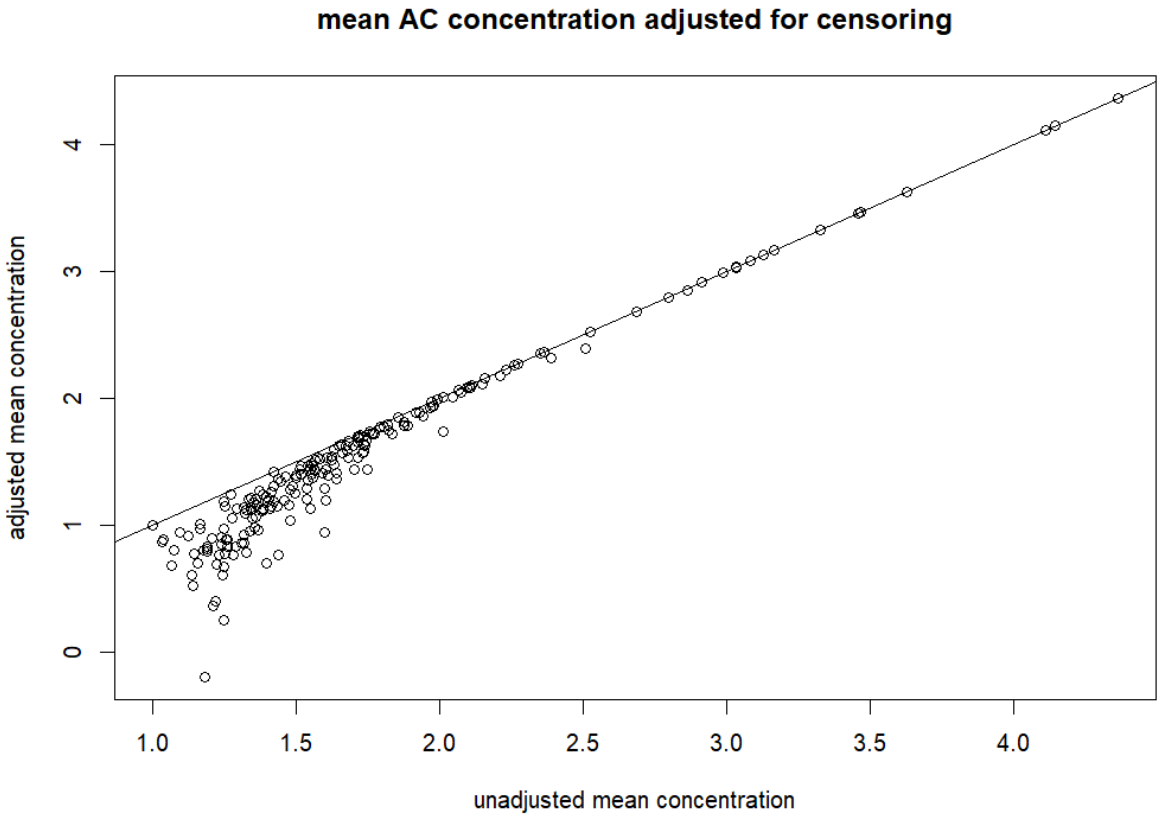


Figure 38: Estimated mean levels for each establishment, after adjustment for the LOD censoring, compared to the mean level derived by substituting 5 cfu/mL (1/2 the LOD) for all sample where no AC were detected.

7.2 AC-Based Performance Standards

Interest lies in determining if performance standards that assess an establishment's capacity to remove and/or inactivate both pathogenic and non-pathogenic bacteria would be an effective alternative to FSIS' current prevalence-based performance standards, which focus on pathogen occurrence on final product. Given that pathogenic bacteria are only consistently isolated from samples collected early in the slaughter process (De Villena, 2022; FSIS, 2009b; Williams, 2015), performance standards based on characteristics of more common and easily measured indicator organisms are appealing. An analysis comparing the occurrence of *Salmonella* at post-chill with the level of AC at both rehang and post-chill identified two statistically significant relationships that could be utilized to propose indicator-organism based performance standards.

The first option is to use changes in the average level of AC between rehang and post-chill to assess an establishment's ability to reduce microbial contamination during slaughter and processing. Evidence shows that *Salmonella* occurrence is lower in establishments with higher average log reduction (**Figure 39**). As shown in **Figure 39**, the overlayed lines represent a logistic regression model ($a < 2 \times 10^{-16}$) predicting the occurrence of *Salmonella* on post-chill samples as a function of average log10 AC reductions between rehang and post-chill. Establishment production volume and corporate

ownership are also highlighted, indicating that lower-volume, independently operated establishments generally have the highest rates of contamination and lower average log10 reductions in AC between rehang and post-chill. While the covariates related to ownership and production volume are interesting and have significant explanatory value, they are not risk management factors that can be adjusted to reduce risk and will not be considered in the risk assessment.

In the U.S., multiple studies have attempted to link levels of indicator organisms at a single point in the production process to either levels or the occurrence of pathogenic bacteria (Altekruse, 2009b; Williams, 2017). Nevertheless, a second indicator-based performance standard option is supported by the observation that a large fraction of post-chill samples have levels of AC below the 10 cfu/mL LOD of the assay. Although roughly 30 percent of all AC samples had no detectable AC, this varies from 0 to nearly 80 percent of samples across the population of establishments. **Figure 40** shows the relationship between the fraction of samples with no detectable AC compared to the proportion of *Salmonella*-positive post-chill samples, with the overlayed lines representing a logistic regression model fitted to the data. As was the case with the average log10 reductions in AC, the independently operated, low-volume establishments rarely have post-chill samples that are free of aerobic bacteria.

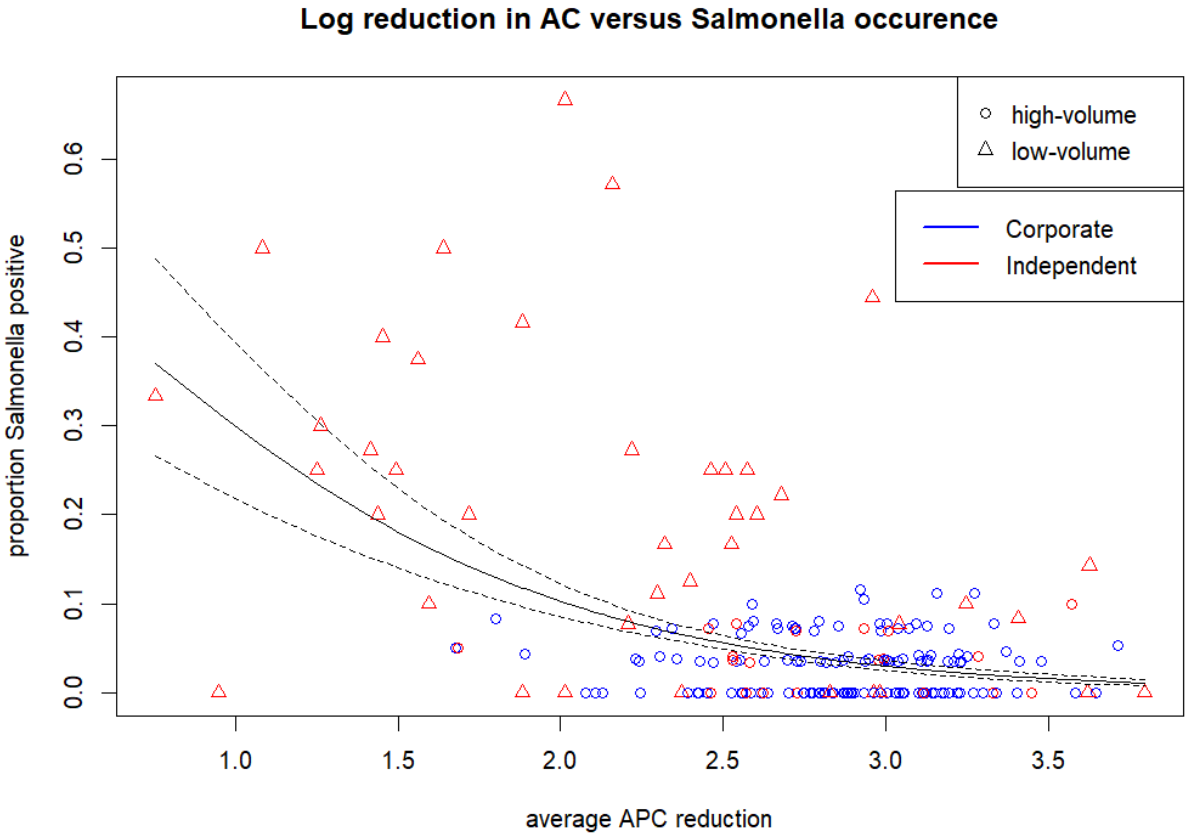


Figure 39: The relationship between log reduction in AC between rehang and post-chill as it relates to *Salmonella* on post-chill samples. Both establishment size and whether establishments belong to a corporation that operates multiple establishments are depicted.

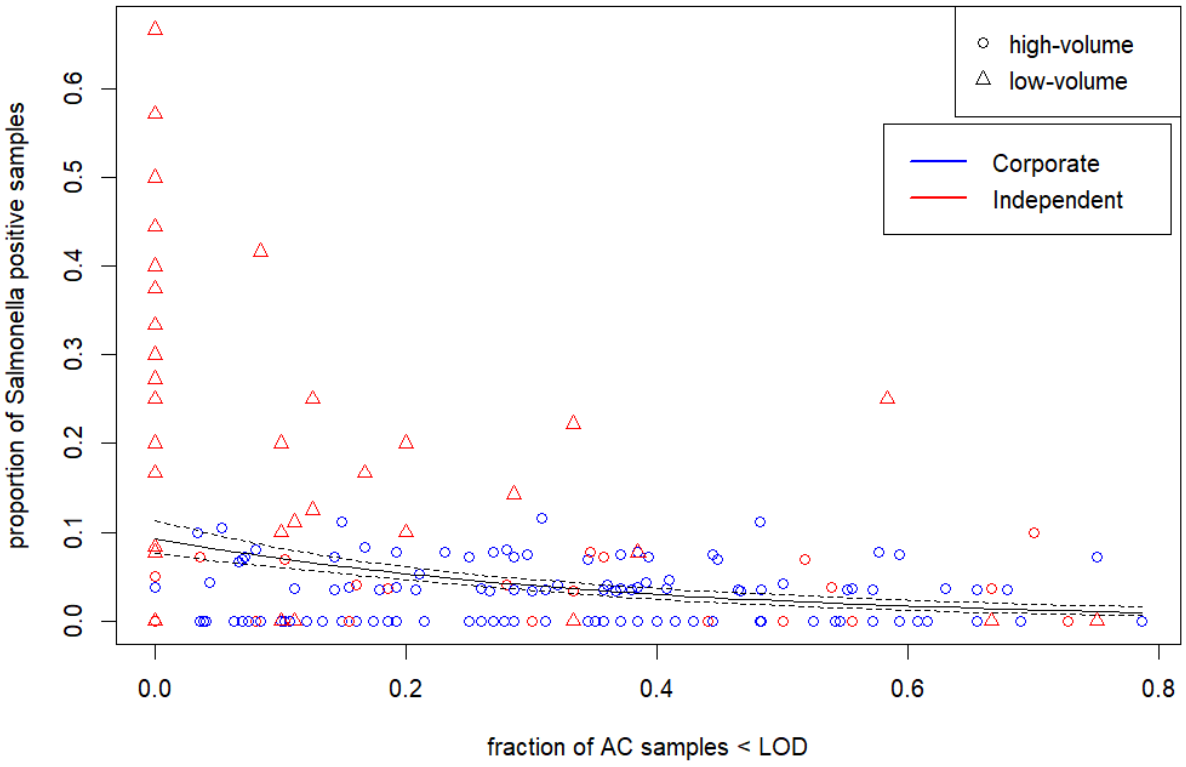


Figure 40: Relationship between *Salmonella* occurrence and the fraction of samples where no AC was isolated.

While establishments with the highest occurrence of *Salmonella* tend to be independent, low-volume producers, there is more diversity in the results for large establishments, with some having only 5 to 10% of samples with no detected AC. Both establishment size, and whether establishments belong to a corporation that operates multiple establishments, is also depicted.

The results in **Figure 39** and **Figure 40** provide evidence that could support two proposed AC-based performance standards:

1. the AC-reduction standard sets a minimum value for the difference in average log10 AC levels between rehang and post-chill, and

2. the AC-elimination standard sets a minimum fraction of post-chill samples where no AC are observed with the current assay (i.e., samples with <10 cfu/mL).

7.3 Modeling Approach

The goal of this analysis is to assess the viability of replacing the existing FSIS performance standards with an alternative framework, using the same basic modeling structure. This risk assessment model predicts the effect of imposing a new performance standard on all slaughter establishments. Once the

performance standard is implemented, establishments will be subjected to the collection and testing of samples, and their results will be used to classify the establishment as either compliant or non-compliant with the standard. Assuming that prevalence is a stable characteristic of establishments, this classification creates two strata. On average, compliant establishments will have a lower prevalence of *Salmonella* compared to non-compliant establishments.

Given that some fraction of establishments would initially fail the performance standard, and some fraction of these establishments would now be either voluntarily compelled or required to lower their carcass contamination frequency, some or all establishments would change their processing to become compliant. This change, from before and after implementation of the performance standard, is how the human health effect of the proposed performance standard is measured. Because it is assumed that there are two basic types of slaughter establishments (i.e., compliant and non-compliant), this approach is called the "two-strata model" (Ebel, 2012b).

For performance standards based on a set of samples collected from the establishment, this analytic approach requires a fraction of production volume associated with establishments that initially pass (ω .). Before the performance standard is implemented,

$$exp) = \omega P_{compliant}\left(exp\right) + \left(1-\omega\right) P_{noncompliant}\left(exp\right). ,$$

e $P_{compliant}\left(exp\right)$ and $P_{noncompliant}\left(exp\right)$. are the prevalence of contaminated carcasses among all slaughter establishments that would pass or fail the performance standard, respectively. $P_{baseline}\left(exp\right)$ is the current prevalence of *Salmonella*-positive samples. These values are estimated using a two-stage cluster sampling approach (Cochran, 1977). The weighting constant ω is the production volume produced by compliant establishments (i.e., establishments whose estimated average log10 AC reduction exceeds the value chosen for the AC-reduction performance standard).

Once the performance standard is implemented, and noncompliant establishments are identified, some fraction,α , of those establishments would change their production practices in order to pass the performance standard. We present values for two different compliance fractions, with the first one assuming that compliance with the performance standard is mandatory and all failing establishments will improve their interventions sufficiently to meet the standard $(\alpha = 1)$. The second choice of compliance fraction assumes that the performance standards are not mandatory, but that half of failing establishments will add additional interventions to meet the standard $\left(\alpha = \frac{1}{2}\right)$.

It is assumed that the additional reductions in all bacteria in failing establishments is such that the failing establishments would ultimately attain a prevalence of *Salmonella*-contaminated carcasses equal to those that pass the performance standard (i.e., $P_{compliant}\left(exp\right)$). Given this expected change, the estimated overall prevalence following implementation of the performance standard is given by;

$$P_{new}\left(exp\right) = \left(\omega + \alpha - \omega\alpha\right) P_{compliant}\left(exp\right) + \left(1-\omega\right)\left(1-\alpha\right) P_{noncompliant}\left(exp\right).$$

The choice of the appropriate cut-off values for both performance standards is determined by setting a specific reduction in the occurrence of pathogens. For this example, we will assume that the reduction target is informed by the HP2030 goal of a 25% reduction in human illnesses (HHS, 2020) and further assume that reductions in the occurrence of pathogen-contaminated samples are proportional to

reductions in cases of salmonellosis.

The new weighted average of the prevalence among compliant and noncompliant establishments is calculated after some fraction of those establishments initially non-compliant have changed their practices so that they are compliant with the performance standard. For a mandatory standard, $P_{new}\left(exp\right) = P_{compliant}\left(exp\right)$ because all noncompliant establishments are expected to change their production processes to match those of the compliant establishments.

7.4 Comparison of AC-Reduction and AC-Elimination Standards

Evaluation of the AC-Reduction Standard

Figure 41 shows the results of applying an AC log reduction standard to the chicken carcass industry over a range of reductions from 1.1 to 3.6 log10. The patterns for $P_{compliant}\left(exp\right)$ and $P_{noncompliant}\left(exp\right)$ demonstrate that changes in log reduction do not necessarily equate to consistent changes in the prevalence value because the establishments that enter and exit compliant and noncompliant subpopulations do not necessarily have a higher or lower prevalence (i.e., log reductions in AC counts are only correlated with the occurrence of *Salmonella* rather than being a perfect predictor).

The vertical lines on **Figure 41:** denote the log reductions that achieve the intended goal. The first vertical line at a 2.9 log10 reduction in AC levels corresponds to an enforceable standard that requires all establishment to achieve that level of reduction. If adoption of the standard is voluntary, the overall log reduction must be set higher at a 3.3 log10 reduction in AC levels to offset the additional *Salmonella* that enters the food supply from establishments that do not adopt the standard.

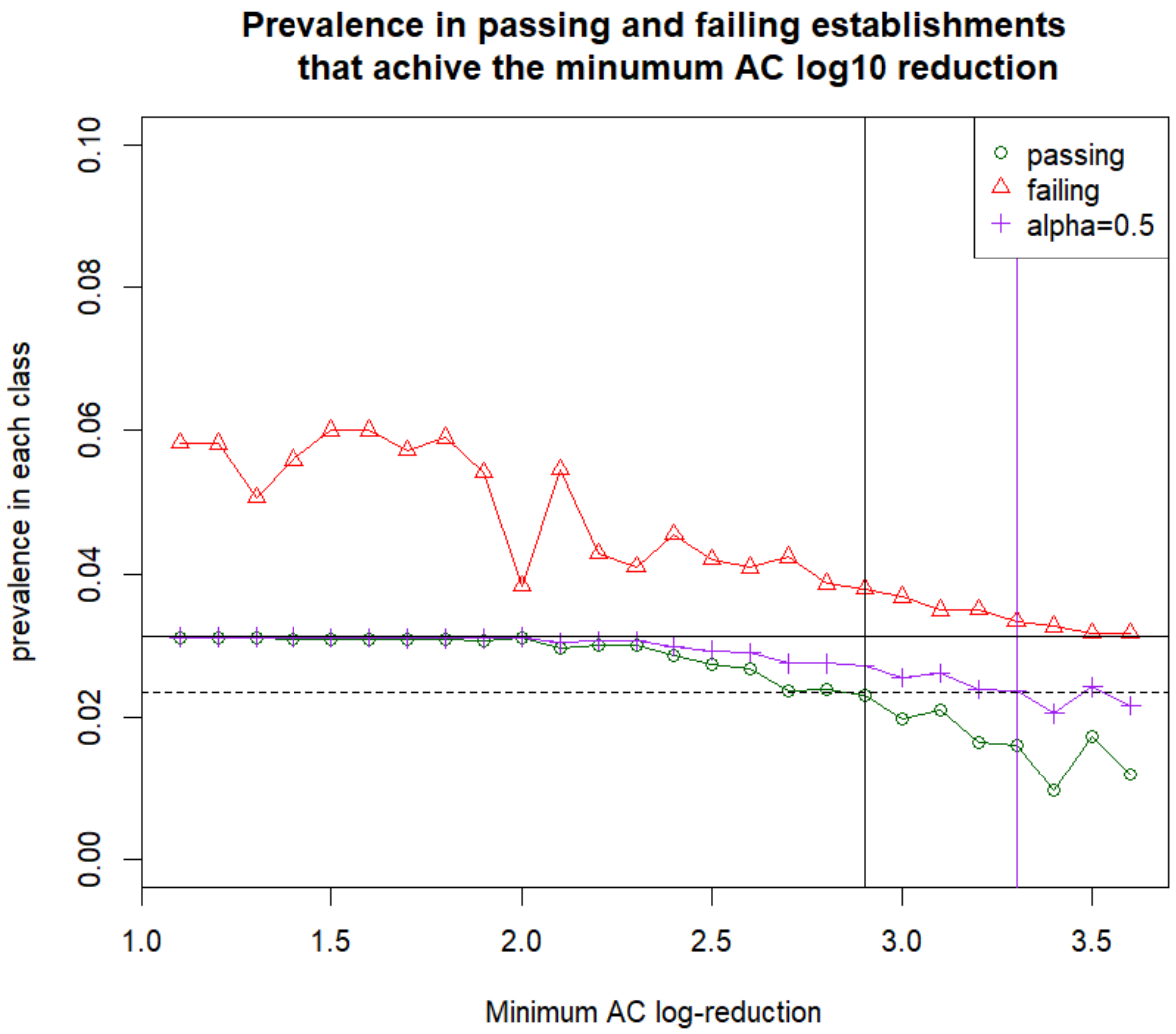


Figure 41: Prevalence in the failing and passing establishments as a function of a range of log10 AC reduction performance standards. The solid horizontal line represents the current prevalence for the industry and the dashed horizontal line represents a 25 percent reduction. A mandatory standard would result in the lower green line while a lower level of compliance (50%) would result in a new *Salmonella* prevalence that represents a mixture of both compliant and noncompliant establishments.

The cost of implementing the standard relates to both the number of establishments that would fail the standard at the time of implementation and the volume of product produced in those establishments. **Figure 42** provides the proportion of the industry that would be passing the standard across the range of average log reduction values. Across the entire range, the proportion of production volume that is initially passing is always greater than the proportion of establishments. This phenomenon occurs because, as demonstrated in **Figure 39**, pathogen contamination tends to be inversely related to production volume in chicken establishments, so low-volume establishments have a higher occurrence of pathogen contamination.

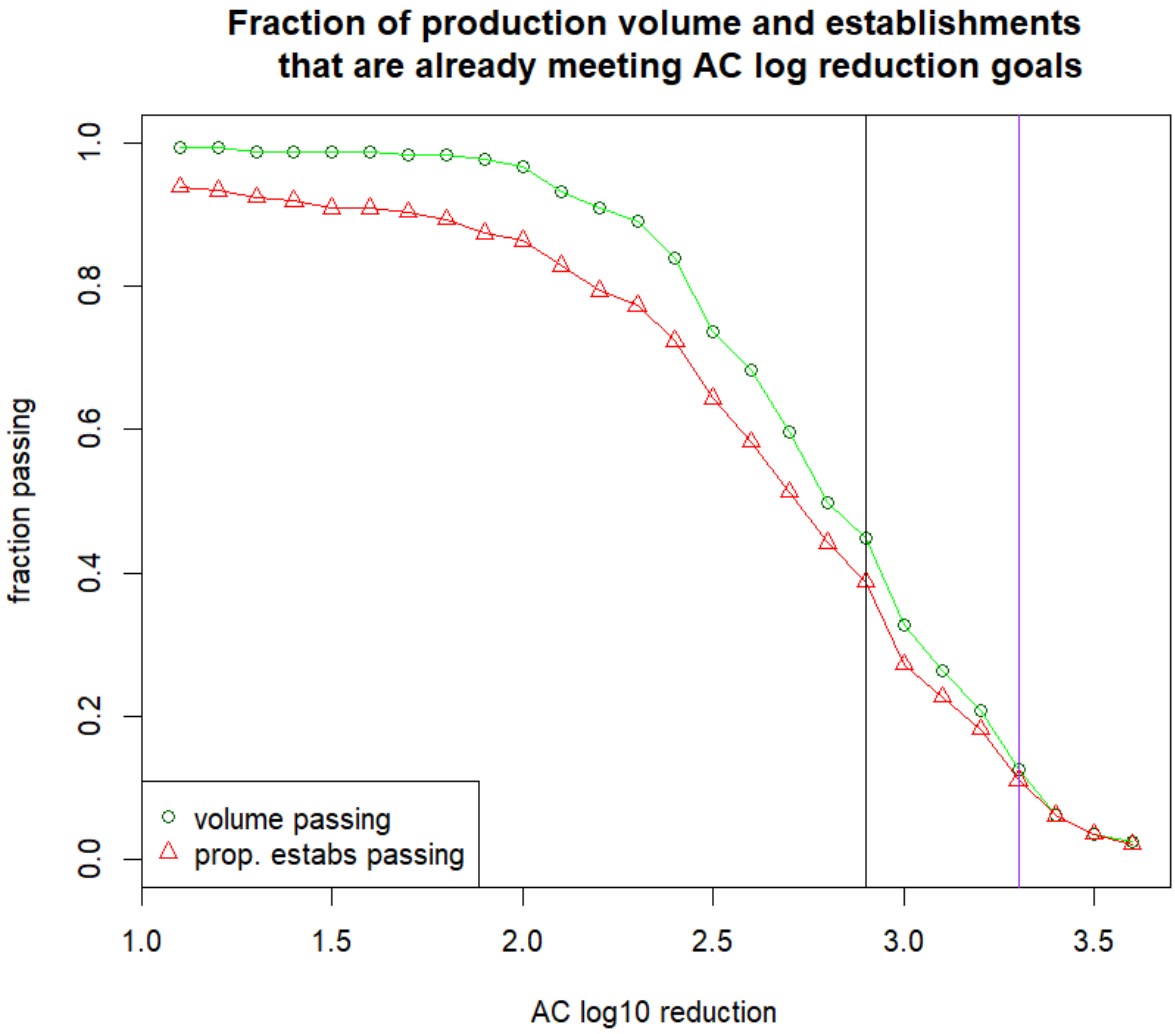


Figure 42: The proportion of production volume and the proportion of establishment for establishments currently passing a range of standards (1.1 to 3.6 log10 reduction).

Since the inception of FSIS performance standards in the mid-1990s, FSIS has always considered a) whether the performance standards were technically feasible and b) the chance that a passing establishment was misclassified as failing. These considerations ensured that a performance standard did not place an unreasonable burden on the industry (FSIS, 1997b). Technical feasibility has been ensured by never setting performance standards lower than what some reasonable fraction of the industry was already achieving. **Figure 42** demonstrates that the possible AC reduction standards are already technically feasible, though less than 50% of establishments are current exceeding about a 2.7 log10 reduction in AC.

The issue of misclassification of establishments under an AC reduction standard is less straight forward because there can be multiple measures of this concept. Given that interest lies in replacing the existing prevalence-based standards with an AC-based standard, a natural measure of misclassification is determining if establishments that are currently passing the prevalence-based performance standard

would also pass the AC reduction standard (i.e., the Se=sensitivity of the new standard in the sense that a reduction in *Salmonella* occurs conditional on a given AC reduction) and similarly an establishment failing the AC reduction standard would also be currently failing the prevalence-based standard (i.e., Sp=specificity). **Figure 43** provides the estimated sensitivities and specificities of an AC reduction standard across the range of possible log10 AC reductions. In general, a good performance standard has both a high Se and Sp value. The sum of the two performance characteristics (Se+Sp) is maximized for average log reductions of 2.2, 2.3 and 2.4. Unfortunately, the log reductions in this range would result in almost no perceptible reduction in the occurrence of *Salmonella*-positive samples (**Figure 41:**). For the log reductions that are predicted to achieve a desired HP2030 reduction of 25% (i.e., 2.9 and 3.3), the sensitivity is less than 50 percent, while specificity exceeds 80 percent.

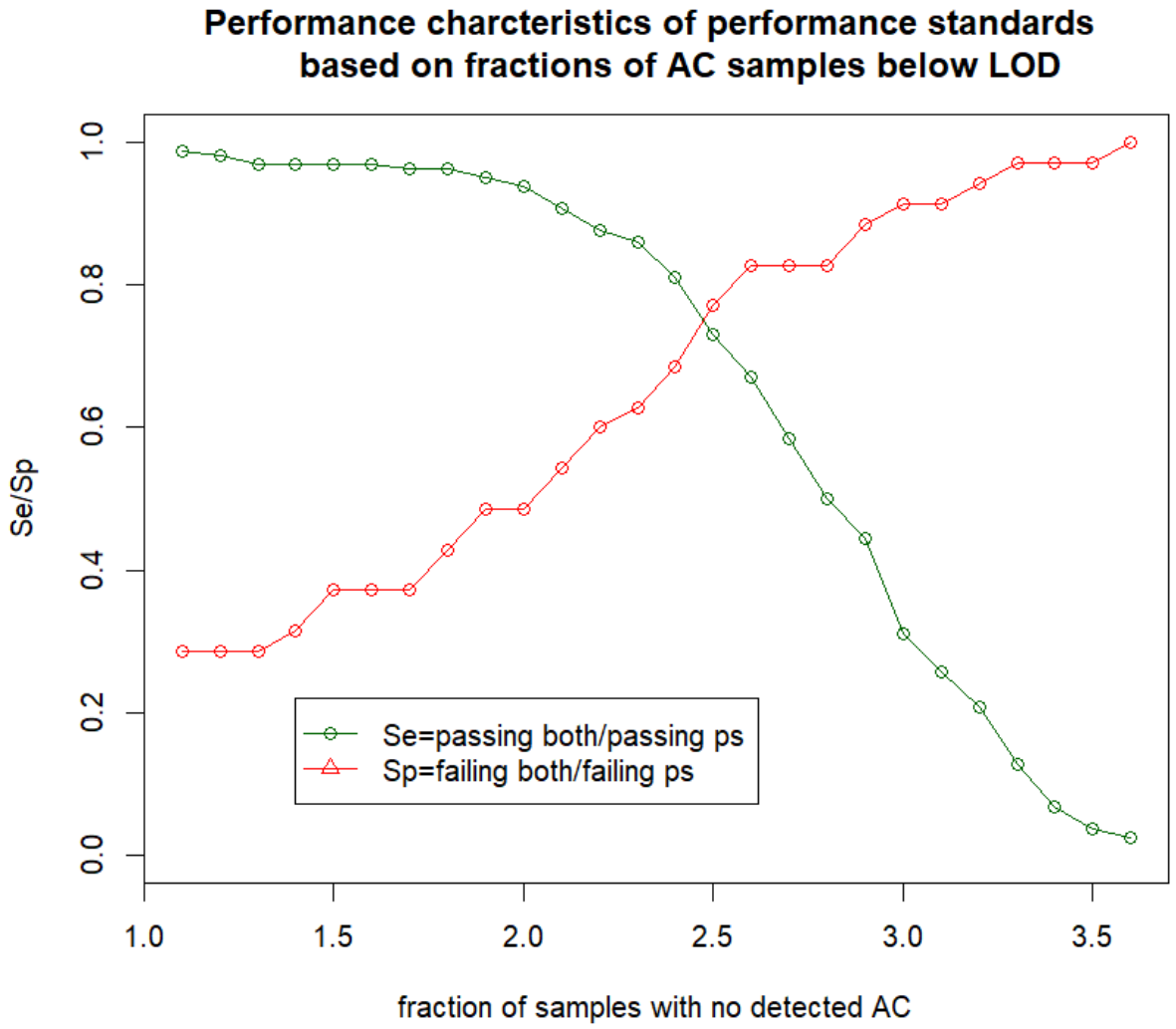


Figure 43: Approximate sensitivity and specificity curves for the AC-based performance standards. The denominators are based on passing or failing the existing FSIS prevalence-based performance standard of 5 positives out of 51. The black vertical line corresponds to a mandatory standard and the purple vertical line corresponds to a voluntary standard.

Evaluation of the AC-elimination standard

Figure 44 shows the results of applying an AC-elimination standard to the chicken carcass industry over a range of sample where no AC were isolated from 0 to 60 percent. Similar to the log-reduction standard, the patterns for $P_{compliant}\left(exp\right)$ and $P_{noncompliant}\left(exp\right)$ demonstrate that changes in the standard do not necessarily equate to consistent changes in the prevalence value because the establishments that enter and exit compliant and noncompliant subpopulations do not necessarily have a higher or lower prevalence.

The vertical lines on **Figure 44** denote the fraction of AC-negative samples that achieve the intended goal, with the vertical line at 35% corresponding to an enforceable standard that requires all establishments to achieve the performance standard. If adoption of the standard is voluntary, the overall goal of a 25% reduction in *Salmonella*-positive samples cannot be achieved, suggesting that additional reductions would need to occur in passing establishments to offset the *Salmonella* contamination that remains in the food supply from failing establishments.

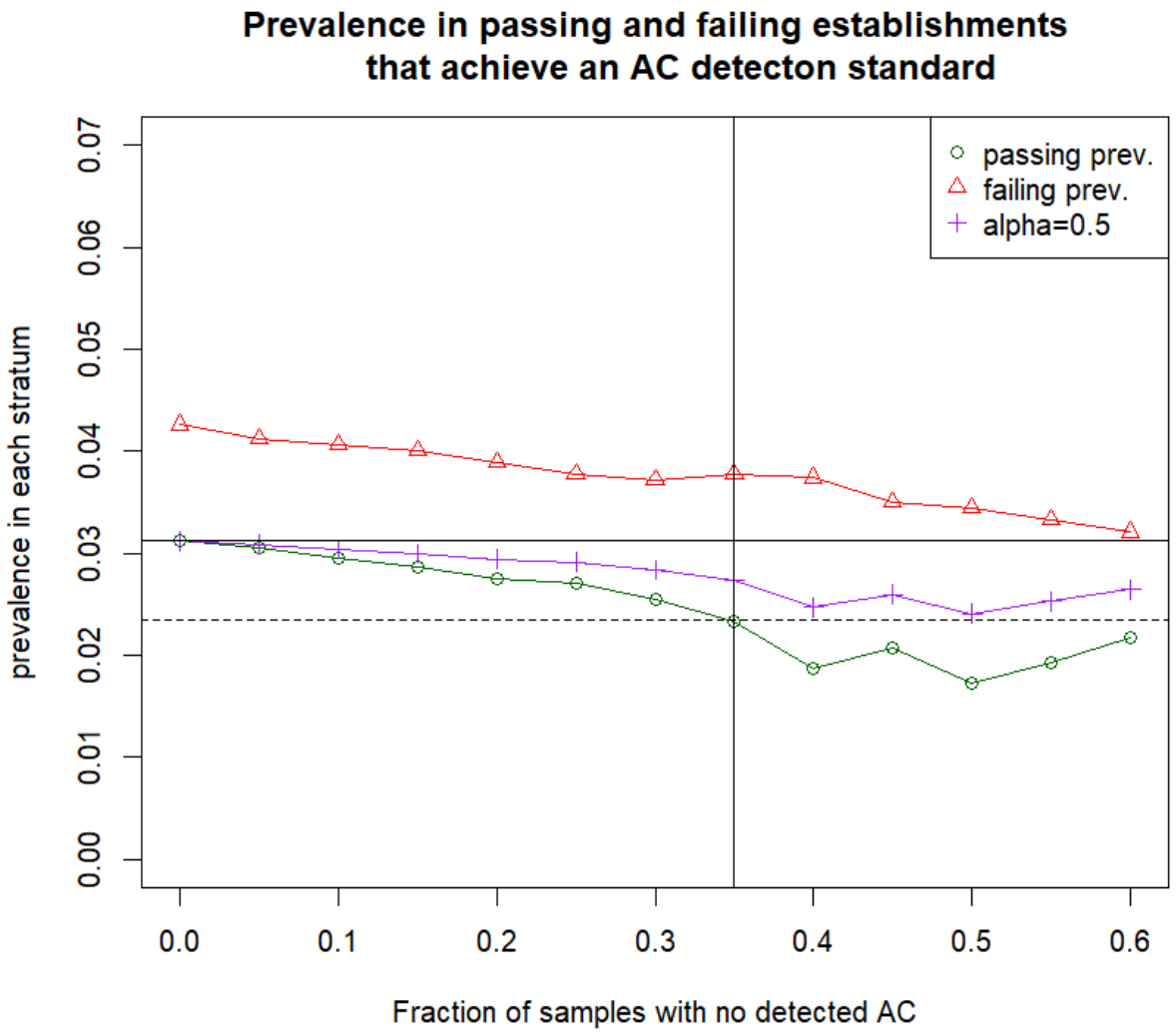


Figure 44: Prevalence in the compliant and noncompliant establishments as a function of AC-elimination performance standards. The horizontal line represents the current prevalence for the industry and the dashed horizontal line represents a 25 percent reduction. A mandatory standard would result in the lower green line while a lower level of compliance (50%) would result in new prevalence that represents a mixture of both compliant and noncompliant establishments.

Figure 45 provides the proportion of the industry that would pass the AC-elimination standard across the range of values. For the enforceable standard, 45% of production volume and about 37% of establishments are already passing the standard that achieves the 25% reduction goal, which is almost identical to the 44 and 38% values for the AC-reduction standard.

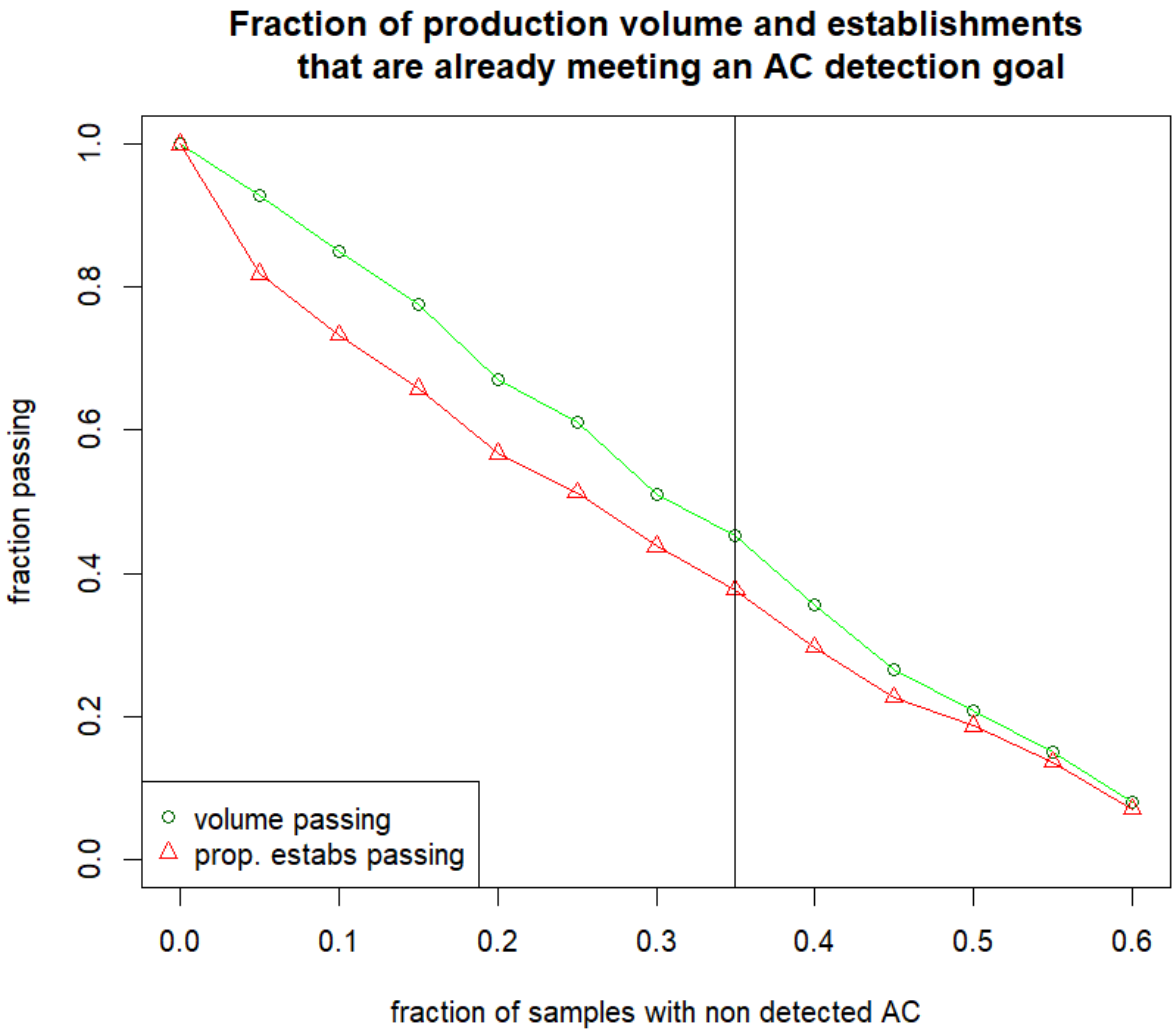


Figure 45: Proportion of production volume and establishments that would already be passing the standard at the time of implementation.

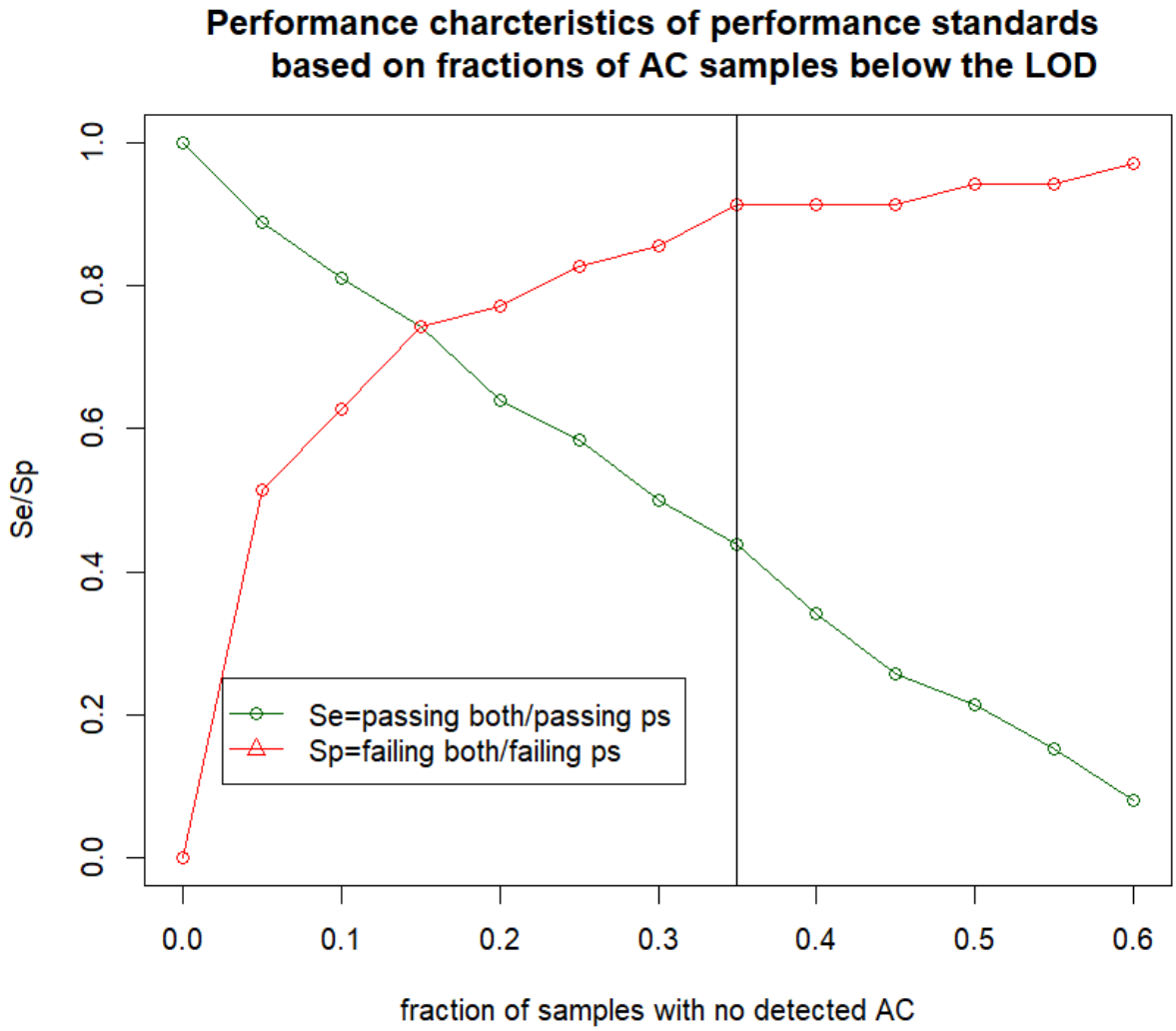


Figure 46: Approximate sensitivity and specificity curves for the AC-based performance standards. The denominators are based on passing of failing the existing FSIS prevalence-based performance standard of 5 positives out of 51.

Figure 46: provides the estimated sensitivities and specificities of an AC-elimination standard across the range of possible samples where no AC was detected. The sum of the two performance characteristics (Se+Sp) for the mandatory standard that achieves the intended 25% reduction is very similar to the value for the mandatory AC-reduction standard; in this case, sensitivity is 40% while specificity is 90%

7.5 Summary

Two mandatory process control standards were considered in this analysis:

1. the AC-reduction standard sets a minimum value for the difference in average log10 AC levels between rehang and post-chill, and

2. the AC-elimination standard sets a minimum fraction of post-chill samples where no AC are observed with the current assay (i.e., samples with <10 cfu/mL).

For each of these approaches, the results were approximately equal in terms of impact on the industry and the performance characteristics. The advantage of the AC-elimination standard is that samples only need to be collected at one location, which has benefits from the associated cost savings in time, materials, and laboratory resources. However, an AC-elimination standard can only be effective if FSIS requires mandatory compliance with the standard. If adoption of the standard is voluntary, the overall goal of a 25% reduction in *Salmonella*-positive samples cannot be achieved,

It is not surprising that both AC-based standards would perform less well than a comparable prevalence-based performance standard because AC reduction are only moderately correlated, rather than highly predictive of the occurrence of *Salmonella.* Probably the best example of this phenomenon is to note that, for the AC-reduction standard, there are 5 large corporate establishments that are predicted to be failing the prevalence-based performance standard and 4 of these have average log10 AC reductions of greater than 2.9 **(Figure 39**). Situations where AC-based standards may be most effective are standards intended to achieve reductions in multiple pathogens (e.g., both *Salmonella* and *Campylobacter)* because the removal and inactivation of all pathogens (bacterial and viral) should be correlated with reductions in indicators organisms.

Chapter 8 Discussion

The risk assessment provides answers to three risk management questions.

> ***Risk Management Question #1****: What is the public health impact (change in illnesses, hospitalizations, and deaths) achieved by eliminating at receiving a proportion of chicken contaminated with specific levels of Salmonella and/or specific Salmonella subtypes?*

Developing a response to this question was significantly hampered by the biggest data gap identified in the course of this risk assessment: the lack of live receiving *Salmonella* data. As a result, rehang sampling was used a proxy for live receiving data, but this is of limited practicality as numerous interventions occur between slaughter and rehang. As live receiving only occurs in slaughter establishments, scenarios for this risk management question could only be assessed for chicken carcasses.

Due to limited paired *Salmonella* level data at the post-chill step (only 14 of the 216 *Salmonella* positive post-chill samples could be enumerated for *Salmonella* in the 2022 FSIS Exploratory Sampling data), among other technical issues, this analysis does not estimate the public health impact of eliminating a portion of chicken carcasses at rehang that are contaminated with a specific level of *Salmonella*. Rather, this analysis focuses on the public health impact of eliminating specific *Salmonella* serotypes at rehang, with a specific focus on the serotype clusters discussed in the Final Product Standards chapter above.

Given the available data, two interpretations of this risk management question were considered:

1. Rehang sampling as a *verification* of *Salmonella* control strategies that were undertaken before (i.e., vaccination, defeathering, etc.), and

2. Rehang sampling as a *location* of potential FSIS action (which requires the consideration of sampling frequency and other logistics).

For the first interpretation of this risk management question—rehang sampling as a verification of *Salmonella* control strategies that were undertaken before—up to 55,000 (a 44% illness reduction) annual salmonellosis cases could potentially be prevented if higher virulence serotypes could be completely removed from flocks, and up to 27,000 (a 22% illness reduction) annual cases could be prevented if the proportion of higher virulence serotypes in flocks could be decreased.

For the second interpretation of this risk management question—rehang sampling as a location of potential FSIS action—up to 50,000 (a 40% illness reduction) annual salmonellosis cases could potentially be prevented.

Scenarios were run that consider the logistics of sampling at rehang for the second question interpretation. Two sampling programs were considered: 1 sample collected per flock and 4 samples collected per flock, with 0 to 3 allowable test positive higher virulence serotype samples. With this approach, if FSIS were to increase sampling at rehang, approximately 5,000 more illnesses could potentially be prevented if flocks with higher virulence serotypes could be removed.

> ***Risk Management Question #2****: What is the public health impact (change in illnesses, hospitalizations, and deaths) achieved by eliminating final product contaminated with specific*

levels of Salmonella and/or specific Salmonella subtypes?

This risk management question was of primary interest for FSIS risk managers and stakeholders. As such, the bulk of model development was geared toward answering this question. The model assesses only changes in overall illnesses, including those that result in hospitalization and deaths, but does not explore direct changes to hospitalizations and deaths.

Level-Based Final Product Standards

The model was run to predict effects for diverting each type of chicken product across a range of level thresholds from 1 cfu/2600mL (-3.41 log10) to 100 cfu/mL (2 log10) for parts and carcasses, and 1 cfu/2600g (-3.41 log10) to 100 cfu/g (2 log10) for comminuted products. Uncertainty analysis was conducted for the major threshold scenarios under consideration by the risk managers. The illnesses prevented estimates with the 95% credible intervals are summarized in **Table 46** below.

Table 46: Estimated annual illnesses prevented by final product concentration standards are shown for concentration thresholds of interest. Values are rounded to nearest 100 illnesses.

Annual illnesses prevented, most likely (95% credible interval)

Threshold level	Chicken carcasses	Chicken parts	Comminuted chicken
0.03 cfu/mL or /g	4600 (2000, 7100)	7900 (3300, 12700)	1500 (800, 2200)
1 cfu/mL or /g	2400 (700, 5000)	1400 (400, 3600)	1400 (600, 2100)
10 cfu/mL or /g	1000 (200, 3100)	200 (40, 700)	1000 (400, 1900)
100 cfu/mL or /g	200 (0, 1500)	20 (0, 100)	600 (200, 1500)

A chicken carcass performance standard that diverts test-positive product lots based on a threshold of 0.033 cfu/mL is the most effective risk management option to reduce foodborne *Salmonella* from chicken carcasses, with 4,700 illnesses prevented annually, which equates to 3.8 percent of the approximately 125,000 overall chicken illnesses that occur each year. The public health impact (in terms of illnesses prevented) of the chicken carcass final product standards encompasses the illnesses estimates for all secondary chicken products, as the majority of those secondary products are fabricated from carcasses.

A chicken parts performance standard that diverts test-positive lots based on a threshold of 0.033 cfu/mL is the most effective risk management option to reduce foodborne *Salmonella* from chicken parts, with 7,850 illnesses prevented annually, which equates to 7.5 percent of the approximately 104,000 chicken parts illnesses that occur each year. The second most effective risk management option to reduce foodborne *Salmonella* from chicken parts is a level threshold of 1cfu/mL, which prevents about 1,400 illnesses annually (a 1.4 percent reduction).

A comminuted chicken performance standard that diverts test-positive lots based on a threshold of

1cfu/g is the most effective risk management option to reduce foodborne *Salmonella* from comminuted chicken, with 1,400 illnesses prevented annually, which equates to approximately 7,500 comminuted chicken illnesses that occur annually (a 19 percent reduction).

A comminuted chicken performance standard that diverts test-positive lots based on a threshold of 0.0307 cfu/g is the second most effective risk management option, with about 1,250 illnesses prevented annually, which equates to approximately 7,500 comminuted chicken illnesses annually (a 17 percent reduction), due in part to a high average risk per lot comminuted chicken.

These results suggest substantial overlap in the 95 percent credible intervals across progressively higher concentration thresholds, which indicates that differences in the most likely effectiveness between different concentration thresholds may not be meaningful.

Serotype-Based Final Product Standards

Serotype-based final product standards were developed solely for chicken carcasses due to data limitations in assessing the mixture of serotypes in chicken parts and comminuted chicken product lots. A chicken carcass performance standard that diverts lots that test positive for the higher virulence *Salmonella* serotypes (any of those in cluster 1) should decrease human illnesses, with 1,800 illnesses per year prevented annually. It is important to note that any final product standard that targets serotypes is only a subset of the most effective risk management options above: a level threshold of 0.0333 cfu/mL. That is, a lot must first test *Salmonella*-positive before the serotype can be identified. Consequently, a serotype-based risk management policy option will not be as effective as a level-based risk management approach alone.

Achieving the public health benefits identified in these final product standards assumes that the *Salmonella* contamination decreases that have been achieved by the chicken industry since 2015 are maintained.

> ***Risk Management Question #3****: What is the public health impact of monitoring/enforcing process control from rehang to post-chill? Monitoring could include analytes such as Enterobacteriaceae, Aerobic Count, or other indicator organisms, analysis could include presence/absence or levels and the monitoring could also include variability of actual result versus expected result, log reduction, absolute sample result, or other individual establishment specific criteria.*

Process control scenarios were assessed for the chicken slaughter industry. Monitoring was interpreted as the effect of a log reduction in AC between the rehang and post-chill. While other indicator organisms and metrics are outlined in the risk management question, there was insufficient data to consider those other metrics across the poultry industry and, furthermore, AC is the only analyte that has been shown to correlate with end point *Salmonella* contamination, though even this correlation has shown to be weak, as discussed below.

Analyses of current (2022 FSIS exploratory paired carcass sampling) and past (2007 FSIS carcass microbiological baseline) data indicates that the chicken industry is consistently achieving a large reduction in AC. Over the past 14 years, the chicken industry as a whole has achieved a 1-2 log reduction in AC. This finding, however, demonstrates that any new performance standard that rely on changes in process control would be limited in its ability to reduce the overall burden of *Salmonella* illnesses from

chicken.

Further complicating efforts to achieve significant decreases in *Salmonella* illnesses from chicken is the fact that, as has been the case historically, indicator organisms are not strongly correlated with the presence of *Salmonella* at post-chill. Specifically, a weak correlation between post-chill *Salmonella* prevalence and AC on two fronts:

1. AC-reduction between rehang and post-chill, and

2. the fraction of post-chill samples where no AC is observed.

The latter of these two correlations is new and did not exist when *Salmonella* rates were higher (Williams, 2015), it will be referred to as AC-elimination.

As a result of these weak relationships between AC and *Salmonella* prevalence, it follows that the correlation between AC 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.

Given these two pieces of information, this analysis instead focused on assessing the potential of two process control performance standards to achieve the HP2030 illness reduction targets for *Salmonella* (U.S. Department of Health and Human Services 2023). These approaches are:

1. an AC-reduction standard that sets a minimum value for the difference in average log10 AC levels between rehang and post-chill,

2. an AC-elimination standard sets a minimum fraction of post-chill samples where no AC are observed with the current assay (i.e., samples with <10 cfu/mL).

As a result of the aforementioned overall reductions in *Salmonella* contamination in the chicken carcass industry, the first standard described above would only be effective at achieving HP2030 illness reduction targets if there was a 2.5log to 3log industry-wide AC reduction standard in AC between rehang and post-chill. Further, to achieve the HP2030 illness reduction targets, the AC standard must be enforceable; voluntary (and thus partial) compliance will not achieve the HP2030 target.

One advantage of second proposed approach, the AC-elimination standard, is that samples only need to be collected at one location (post-chill), which has benefits from the associated cost savings in time, materials, and laboratory resources. However, an AC-elimination standard can only be effective if FSIS requires mandatory compliance with the standard. If adoption of the standard is voluntary, the overall goal of a 25% reduction in *Salmonella*-positive samples cannot be achieved.

It is not surprising that both AC-based standards would perform less well than a comparable prevalence-based performance standard because AC reductions are only moderately correlated, rather than highly predictive, of the occurrence of *Salmonella*. Situations where AC-based standards may be most effective are standards intended to achieve reductions in multiple pathogens (e.g., both *Salmonella* and Campylobacter) because the removal and inactivation of all pathogens (bacterial and viral) should be correlated with reductions in indicators organisms.

Risk Management Question #4*: What is the public health impact of implementing combinations of the risk management options listed above?*

The fourth risk management question could not be answered by the analyses summarized in this document. Analytical challenges and data gaps prevented a full treatment of combination scenarios. Cursory explanations for how scenarios could be combined are outlined below, along with some of the limitations researchers face in pursuing this work.

Conceptually, final product standards can be combined as level and serotype-focused approaches, as well as across each of the three product types (i.e., carcasses, parts and comminuted). For example, a carcass final product standard that combines a level-based threshold of 0.033 cfu/mL and detection of higher virulence *Salmonella* is already estimated in this report. Other combinations of this type could be explored (e.g., a level of 1 cfu/mL and detection of a higher virulence serotype) for carcasses, but serotype data for parts and comminuted will not support the approach used in this analysis. We lack data to impute different schemes for higher virulence cluster frequency for parts or comminuted. Therefore, the effect of detecting higher or lower virulence serotype clusters given that a parts or comminuted lot fails a level-based threshold is the same.

For any combination, there is also the option to define the standard as the union of the individual components (e.g., a lot fails based on concentration or detection of higher virulence *Salmonella*). In this case, we could sum the predicted effects of each component and subtract the prevented illnesses that the individual effects share. For example, if a carcass standard failed a lot when its concentration was >0.0.33 cfu/g or a higher virulence *Salmonella* was detected, then we could add the predictions from **Table 35** (4700 illnesses prevented) and **Table 36** (1800 illnesses prevented), but subtract the illnesses prevented predicted when both criteria apply (i.e., the same 1800 illnesses prevented in **Table 36**). Alternatively, the standard could be defined as the intersection of the individual components (e.g., a lot fails based on concentration and detection of higher virulence serotype). As illustrated in our example, this intersection is the 1800 illnesses prevented by the scenario in **Table 36**. Despite its relative simplicity, this example also illustrates the fact that a union combination policy will generally predict prevented illnesses that are greater than an intersection combination policy.

Combining level-based standards across the product types would essentially be a union of the individual effects. Production lots of carcasses, parts and comminuted are unrelated. Therefore, testing would result in diversion when any lot failed. Because it is reasonable to assume that any combination will include a carcass standard, our suggested approach to estimating the prevented illnesses for a combination of standards across products is to first determine the illnesses prevented by the carcass standard, then remove these illnesses from the population before determining for parts and/or comminuted. This approach has the effect of removing the prevented illnesses shared by the multiple standards. Alternatively, a conservative approach that avoids the potential of overestimating the effect of a combination of standards would be to only count the illnesses prevented by the carcass standard or only count the maximum of the illnesses prevented across the combination of standards.

Combining receiving, processing and final product standards is also possible from a risk management perspective. The analysis of potential effects for a receiving standard, however, were limited to serotype considerations. That analysis assumed that all incoming lots would be subjected to testing with one of four testing regimens. For a chosen testing regimen, it is possible to adjust the expected frequency of

higher virulence *Salmonella* that would be applicable to final product testing. For example, the receiving standard analysis suggests that rejecting a lot if a single sample is found positive for a higher virulence *Salmonella* would generate a new value of 0.105 (**Table 44**) for the parameter in the final product standards model. The final product model could be adjusted to account for this change to determine the effect of a combination of the receiving and final product standards. Nevertheless, the substantial cost and consequence associated with testing all incoming lots suggests that such a combination of standards may be unwieldly.

The process control analysis is focused on predicting changes in *Salmonella* prevalence following the imposition of AC log reduction standards. As such, it is not amenable directly to combining with receiving or final product standards. Nevertheless, it is feasible that the initial contamination concentration distribution used in the final product standards could be adjusted to account for predicted changes in *Salmonella* prevalence from a process control standard. As explained in Ebel and Williams (2015), such an adjustment could be scalar (only influencing the lognormal parameter) or non-scalar (influencing both and lognormal parameters). Such adjustments are only warranted if the process control standards are mandatory and based on sufficient testing that limits establishment misclassification.

8.1 Research Needs

The following research needs were identified during the development of the risk assessment in raw chicken products. The following list is not prioritized:

1. Within Flock *Salmonella* Variability

   Serotype Mixtures: Analysis of the FSIS 2022 Exploratory Sampling data with two data points (rehang and post-chill) per flock indicates that multiple *Salmonella* serotypes occur in flocks. While it is possible that some flocks do not contain a single, dominant serotype, no data exist that describes what other per flock mixtures of serotypes may be present or how prevalent they are in the U.S. poultry population.

   Per Unit Salmonella Population: It is plausible that multiple serotypes are present on a given carcass, part, or comminuted unit where units are flocks and lots of product or single samples (Berrang, 2009; Cox, 2020; Obe, 2023; Rasamsetti, 2023; Siceloff, 2022). The model developed in this document does not require resolution down to the level of single chicken carcass to answer the risk management questions, and therefore this data gap was surmountable. However, a microbial profile at the carcass level could shed light on population *Salmonella* variability.

2. Industry Behavior Data
   While it can be postulated that any regulation that establishes enforceable standards will have an indirect effect on the poultry industry's pathogen control measures, no data is available at this time describing the magnitude of that effect. Therefore, only the direct of effect of such a risk management approach was assessed in this risk assessment.

More clarity and insight could be gained from data describing interventions that target levels of *Salmonella* and their efficacy. While some serotype-specific interventions are known (e.g., vaccination), their current usage and effectiveness is not well understood and no data on industry-wide usage are available at this time.

The declaration of *Salmonella* as an adulterant may lead to an industry-wide shift of control measures on the same scale as the STEC O157 policy (FSIS, 2002); however, no after-action analyses of the STEC shift are available at this time. One feature that is well understood is the industry change in lotting practices on the basis of STEC microbial independence in response to the introduction of the STEC O157 adulterant policy change. This risk assessment used the average industry lot sizes for chicken products (flocks and days production), but future research on this topic may refine lot size on the basis of *Salmonella* survival capability throughout slaughter and processing.

3. Efficacy of Preharvest Interventions

   NACMCF 2023 response (NACMCF, 2023) outlined the microbiological criteria that could be established to encourage control of *Salmonella* at preharvest, but there remains little concrete data on the effectiveness of existing preharvest interventions and how widespread their usage is. Therefore, future risk assessments would benefit from both data collection and a systematic literature review and data extraction in the style of (Wang, 2023).

4. *Salmonella* virulence capacity

   The genetic basis of *Salmonella* virulence has not been fully elucidated and is likely to be complex. Virulence genes in *Salmonella* are heavily influenced by gene acquisition facilitated by horizontal gene transfer and gene loss through pseudo-gene formation. The clustering approach undertaken in this risk assessment relied on the presence/absence of *Enterobacteriaceae* virulence gene markers without directly accounting for their biological function. As research into *Salmonella* virulence factors and their gene functions continues to develop, clustering should be revisited to ensure the reliability/consistency, and potentially, the resolution. Additionally, outbreak data was used to validate the constructed seroclusters to estimate relative risk to public health. Further exploration into the individual strains within broad seroclusters would continue to improve future risk assessment analyses.

References

Aitchison, J. (1982). The statistical analysis of compositional data. *J R Stat Soc: Series B Stat Methodol*, *44*(2), 139-160.

Altekruse, S. F., Berrang, M. E., Marks, H., Patel, B., Shaw Jr, W. K., Saini, P., Bennett, P. A., & Bailey, J. S. (2009a). Enumeration of *Escherichia coli* cells on chicken carcasses as a potential measure of microbial process control in a random selection of slaughter establishments in the United States. *Appl Environ Microbiol*, *75*(11), 3522-3527. https://doi.org/10.1128/aem.02685-08

Altekruse, S. F., Berrang, M. E., Marks, H., Patel, B., Shaw, W. K., Saini, P., Bennett, P. A., & Bailey, J. S. (2009b). Enumeration of *Escherichia coli* cells on chicken carcasses as a potential measure of microbial process control in a random selection of slaughter establishments in the United States. *Applied and Environmental Microbiology*, *75*(11), 3522-3527.

Altman, D. G., & Bland, J. M. (1994). Diagnostic tests 2: Predictive values. *Bmj*, *309*(6947), 102. https://doi.org/10.1136/bmj.309.6947.102

Anderson, W., Gathura Gichia, M., Buchanan, R. L., Donahue, D., Hielm, S., Kiermeier, A., Kumar Jain, A., Paoli, G., Fernando Perez, F., Toyofuku, H., Williams, M. S., & Zwietering, M. H. (2016). *Statistical aspects of a microbiological criteria related to foods: A risk managers guide* (Vol. 24). FAO/WHO. http://www.who.int/foodsafety/publications/mra_24/en/

Armstrong, B. G. (1998). Effect of measurement error on epidemiological studies of environmental and occupational exposures. *Occup Environ Med*, *55*(10), 651-656. https://doi.org/10.1136/oem.55.10.651

Bassett, J., Jackson, T., Jewell, K., Jongenburger, I., & Zwietering, M. (2010). *Impact of microbial distributions on food safety* (9789078637202). www.ilsi.org/Europe/Publications/Microbial%20Distribution%202010.pdf

Batz, M. B., Richardson, L. C., Bazaco, M. C., Parker, C. C., Chirtel, S. J., Cole, D., Golden, N. J., Griffin, P. M., Gu, W., Schmitt, S. K., Wolpert, B. J., Kufel, J. S. Z., & Hoekstra, R. M. (2021). Recency-weighted statistical modeling approach to attribute illnesses caused by 4 pathogens to food sources using outbreak data, United States. *Emerg Infect Dis*, *27*(1), 214-222. https://doi.org/10.3201/eid2701.203832

Bauer, N., Evans, P., Leopold, B., Levine, J., & White, P. (2014). *White Paper: Current and Future Development and Use of Molecular Subtyping by USDA-FSIS*. https://www.fsis.usda.gov/sites/default/files/media_file/2020-06/Molecular-Subtyping-White-Paper.pdf

Belgorodski, N., Greiner, M., Tolksdorf, K., Mueller-Graf, C., Schueller, K., Yang, Y., & Junker, M. (2016). rrisk [R package].

Berrang, M. E., Bailey, J. S., Altekruse, S. F., & Shaw, W. K., Jr. (2008). Presence and numbers of *Campylobacter*, *Escherichia coli*, and *Salmonella* determined in broiler carcass rinses from United States processing plants in the hazard analysis and critical control point-based inspection models project. *J Appl Poult Res*, *17*(3), 354-360. https://search.ebscohost.com/login.aspx?direct=true&db=lah&AN=20093236803&site=ehost-live

Berrang, M. E., Bailey, J. S., Altekruse, S. F., Shaw, W. K., Patel, B. L., Meinersmann, R. J., & Fedorka-Cray, P. J. (2009). Prevalence, serotype, and antimicrobial resistance of *Salmonella* on broiler carcasses postpick and postchill in 20 US processing plants. *J Food Prot*, *72*(8), 1610-1615.

Beshearse, E., Bruce, B. B., Nane, G. F., Cooke, R. M., Aspinall, W., Hald, T., Crim, S. M., Griffin, P. M., Fullerton, K. E., Collier, S. A., Benedict, K. M., Beach, M. J., Hall, A. J., & Havelaar, A. H. (2021). Attribution of Illnesses Transmitted by Food and Water to Comprehensive Transmission Pathways Using Structured Expert Judgment, United States. *Emerg Infect Dis*, *27*(1), 182-195. https://doi.org/10.3201/eid2701.200316

Bulmer, M. G. (1974). On fitting the Poisson lognormal distribution to species abundance data. *Biometrics*, *30*, 651-660.

Centers for Disease Control and Prevention. (2021). National Outbreak Reporting System (NORS) Dashboard. *http://www.cdc.gov/norsdashboard*, *Accessed December 26, 2022*. www.cdc.gov/norsdashboard

Centers for Disease Control and Prevention. (2022a). FoodNet Fast: Pathogen Surveillance Tool. *https://www.cdc.gov/foodnet/foodnet-fast.html*, *Accessed December 26, 2022*. http://wwwn.cdc.gov/foodnetfast.

Centers for Disease Control and Prevention. (2022b). National Health and Nutrition Examination Survey (NHANES): 2017-March 2020 Data Documentation, Codebook, and Frequencies. *https://www.cdc.gov/nchs/nhanes/index.htm*(December 2022).

Centers for Disease Control and Prevention. (2022). Preliminary Incidence and Trends of Infections Caused by Pathogens Transmitted Commonly Through Food— Foodborne Diseases Active Surveillance Network, 10 U.S. Sites, 2016–2021

Chaney, W. E., Englishbey, A. K., Stephens, T. P., Applegate, S. F., & Sanchez-Plata, M. X. (2022). Application of a Commercial *Salmonella* Real-Time PCR Assay for the Detection and Quantitation of *Salmonella* enterica in Poultry Ceca. *J Food Prot*, *85*(3), 527-533. https://doi.org/10.4315/JFP-21-285

Chen, R. X., Cheng, R. A., Wiedmann, M., & Orsi, R. H. (2022). Development of a Genomics-Based Approach To Identify Putative Hypervirulent Nontyphoidal *Salmonella* Isolates: *Salmonella enterica* Serovar Saintpaul as a Model. *mSphere*, *7*(1), e0073021. https://doi.org/10.1128/msphere.00730-21

Chen, Y., Jackson, K. M., Chea, F. P., & Schaffner, D. W. (2001). Quantification and variability analysis of bacterial cross-contamination rates in common food service tasks. *J Food Prot*, *64*(1), 72-80. https://doi.org/10.4315/0362-028x-64.1.72

Cochran, W. G. (1950). Estimation of bacterial densities by means of the most probable number *Biometrics*, *6*, 105-116. https://doi.org/10.2307/3001491

Cochran, W. G. (1977). *Sampling Techniques* (3 ed.). John Wiley and Sons.

Commeau, N., Parent, E., Delignette-Muller, M. L., & Cornu, M. (2012). Fitting a lognormal distribution to enumeration and absence/presence data. *Int J Food Microbiol*, *155*(3), 146-152. https://doi.org/10.1016/j.ijfoodmicro.2012.01.023

Cowling, D. W., Gardner, I. A., & Johnson, W. O. (1999). Comparison of methods for estimation of individual-level prevalence based on pooled samples. *Prev Vet Med*, *39*(3), 211-225. https://doi.org/10.1016/s0167-5877(98)00131-7

Cox, N. A., Berrang, M. E., House, S. L., Hinton Jr, A., Eric Line, J., & Wiggins, L. T. (2020). Detection of multiple naturally occurring *Salmonella* serotypes from commercial broiler carcasses with conventional methods. *J Food Safety*, *40*(2), e12761. https://doi.org/https://doi.org/10.1111/jfs.12761

De Villena, J. F., Vargas, D. A., Bueno López, R., Chávez-Velado, D. R., Casas, D. E., Jiménez, R. L., & Sanchez-Plata, M. X. (2022). Bio-Mapping indicators and pathogen loads in a commercial broiler processing facility operating with high and low antimicrobial intervention levels. *Foods*, *11*(6), 775. https://doi.org/10.3390/foods11060775

Ebel, E., Williams, M., Golden, N., & Marks, H. (2012a). Simplified framework for predicting changes in public health from performance standards applied in slaughter establishments [Article]. *Food Control*, *28*(2), 250-257. https://doi.org/10.1016/j.foodcont.2012.05.016

Ebel, E. D., & Williams, M. S. (2015). When are qualitative testing results sufficient to predict a reduction in illnesses in a microbiological food safety risk assessment? *J Food Prot*, *78*(8), 1451-1460. https://doi.org/10.4315/0362-028X.JFP-15-042

Ebel, E. D., Williams, M. S., & Amann, D. M. (2020). Quantifying the effects of reducing sample size on 2-class attributes sampling plans: Implications for United States poultry performance standards. *Food Control*, *111*, 107068. https://doi.org/10.1016/j.foodcont.2019.107068

Ebel, E. D., Williams, M. S., Golden, N. J., & Marks, H. M. (2012b). Simplified framework for predicting changes in public health from performance standards applied in slaughter establishments. *Food Control*, *28*(2), 250-257. https://doi.org/10.1016/j.foodcont.2012.05.016

Ebel, E. D., Williams, M. S., Hoekstra, R. M., Golden, N. J., Cole, D., Travis, C. C., & Klontz, K. C. (2016). Comparing disease characteristics of sporadic and outbreak foodborne illnesses. *Emerg Infect Dis*, *22*, 1193-1200.

Ebel, E. D., Williams, M. S., & Schlosser, W. D. (2012c). Parametric distributions of underdiagnosis parameters used to estimate annual burden of illness for five foodborne pathogens. *J Food Prot*, *75*, 775–778. https://doi.org/10.4315/0362-028X.JFP-11-345

Ebel, E. D., Williams, M. S., & Schlosser, W. D. (2017). Estimating the Type II error of detecting changes in foodborne illnesses via public health surveillance. *Microbial Risk Analysis*, *7*, 1-7. https://doi.org/https://doi.org/10.1016/j.mran.2017.10.001

Engen, S., Lande, R., Walla, T., & DeVries, P. J. (2002). Analyzing spatial structure of communities using the two-dimensional Poisson lognormal species abundance model. *Am Nat*, *160*, 60-73.

Faes, C., Molenberghs, G., Hens, N., Muller, A., Goossens, H., & Coenen, S. (2011). Analysing the composition of outpatient antibiotic use: a tutorial on compositional data analysis. *J Antimicrob Chemother*, *66*(suppl_6), vi89-vi94. https://doi.org/10.1093/jac/dkr461

FAO/WHO. (1999). Codex Commission, Principles and Guidelines for the Conduct of Microbiological Risk Assessment. (CAC/GL 30-1999). https://www.fao.org/fao-who-codexalimentarius/sh-proxy/en/?lnk=1&url=https%253A%252F%252Fworkspace.fao.org%252Fsites%252Fcodex%252FStandards%252FCXG%2B30-1999%252FCXG_030e_2014.pdf

FAO/WHO. (2000). *Ad Hoc Expert Consultations on Risk Assessment of Microbiological Hazards in Foods* (Report of the Joint FAO/WHO Expert Consultation on Risk Assessment of Microbiological Hazards in Foods, Issue. https://www.fao.org/fileadmin/templates/agns/pdf/jemra/SL00_en.pdf

FAO/WHO. (2002). *Risk Assessments of Salmonella in Eggs and Broiler Chickens* (Microbiological Risk Assessment Series, Issue. https://www.fao.org/3/y4392e/y4392e00.htm

FAO/WHO. (2016). *Statistical aspects of microbiological criteria related to foods: a risk manager's guide* (9241565314). (Microbiological Risk Assessment Series, no 24, Issue. https://apps.who.int/iris/handle/10665/249531

FAO/WHO. (2021). Microbial risk assessment – Guidance for food. https://www.who.int/publications/i/item/9789240024892 (Microbiological Risk Assessment Series)

Fenske, G. J., Pouzou, J. G., Pouillot, R., Taylor, D. D., Costard, S., & Zagmutt, F. J. (2023). The genomic and epidemiological virulence patterns of Salmonella enterica serovars in the United States. *PLoS One*, *18*(12), e0294624. https://doi.org/10.1371/journal.pone.0294624

Fuller, W. A. (2009). *Sampling Statistics*. John Wiley & Sons.

Gasparro, A. (2018, January 4, 2018). Rotisserie Chickens: The '90s Gift to Supermarkets That Keeps on Giving. *Wall Street Journal*.

Gonzales-Barron, U., & Butler, F. (2011). A comparison between the discrete Poisson-gamma and Poisson-lognormal distributions to characterise microbial counts in foods. *Food Control*, *22*(8), 1279-1286. https://doi.org/10.1016/j.foodcont.2011.01.029

Gould, L. H., Walsh, K. A., Vieira, A. R., Herman, K., Williams, I. T., Hall, A. J., Cole, D., & (CDC), C. f. D. C. a. P. (2013). Surveillance for foodborne disease outbreaks - United States, 1998-2008. *MMWR Surveill Summ*, *62*(2), 1-34. https://www.ncbi.nlm.nih.gov/pubmed/23804024

Grimont, P. A., & Weill, F.-X. (2007). *Antigenic Formulae of the Salmonella Serovars* (9th ed.). WHO Collaborating Center for Reference and Research on *Salmonella*, Institut Pasteur. http://www.scacm.org/free/Antigenic%20Formulae%20of%20the%20Salmonella%20Serovars%202007%209th%20edition.pdf

Haas, C. N. (1989). Estimation of microbial densities from dilution count experiments. *Appl Environ Microbiol*, *55*(8), 1934-1942. https://doi.org/10.1128/aem.55.8.1934-1942.1989

Hazra, A., & Gogtay, N. (2017). Biostatistics Series Module 7: The Statistics of Diagnostic Tests. *Indian J Dermatol*, *62*(1), 18-24. https://doi.org/10.4103/0019-5154.198047

Helsel, D. (2009). Much ado about next to nothing: incorporating nondetects in science. *Ann Occup Hyg*, *54*(3), 257-262. https://doi.org/10.1093/annhyg/mep092

Helsel, D. R. (2005). *Nondetects and Data Analysis: Statistics for Censored Environmental Data*. Wiley Inter-science.

Helsel, D. R. (2010). Summing nondetects: Incorporating low-level contaminants in risk assessment. *Integr Environ Assess Manag*, *6*(3), 361-366. https://doi.org/10.1002/ieam.31

Hoffmann, S., & Ahn, J.-W. (2021). Updating Economic Burden of Foodborne Diseases Estimates for Inflation and Income Growth. (ERR-297).

Hsi, D. J., Ebel, E. D., Williams, M. S., Golden, N. J., & Schlosser, W. D. (2015). Comparing foodborne illness risks among meat commodities in the United States [Article]. *Food Control*, *54*, 353-359. https://doi.org/10.1016/j.foodcont.2015.02.018

Interagency Food Safety Analytics Collaboration. (2019). *Foodborne illness source attribution estimates for 2017 for Salmonella, Escherichia coli O157, Listeria monocytogenes, and Campylobacter using multi-year outbreak surveillance data, United States* https://www.cdc.gov/foodsafety/ifsac/pdf/P19-2017-report-TriAgency-508.pdf

Interagency Food Safety Analytics Collaboration (IFSAC). (2022). Foodborne illness source attribution estimates for 2020 for *Salmonella*, *Escherichia coli* O157, and *Listeria monocytogenes* using multi-year outbreak surveillance data, United States. *https://www.cdc.gov/foodsafety/ifsac/pdf/P19-2020-report-TriAgency-508.pdf*.

Izsak, R. (2008). Maximum likelihood fitting of the Poisson lognormal distribution. *Environ Ecol Stat*, *15*, 43–156. https://doi.org/10.1007/s10651-007-0044-x

Joelsson, A. C., Barkhouse, D., Kahle, K., & Siciliano, N. A. (2022). *GENE-UP® enviroPRO™, formerly known as enviroPRO™*. https://members.aoac.org/AOAC_Docs/RI/22PTM/22C_061801_ISenviropro.pdf

Karanth, S., Tanui, C. K., Meng, J., & Pradhan, A. K. (2022). Exploring the predictive capability of advanced machine learning in identifying severe disease phenotype in Salmonella enterica. *Food Res Int*, *151*, 110817. https://doi.org/10.1016/j.foodres.2021.110817

Kralik, P., & Ricchi, M. (2017). A Basic Guide to Real Time PCR in Microbial Diagnostics: Definitions, Parameters, and Everything [Review]. *Front Microbiol*, *8*. https://doi.org/10.3389/fmicb.2017.00108

Lambertini, E., Ruzante, J. M., Chew, R., Apodaca, V. L., & Kowalcyk, B. B. (2019). The public health impact of different microbiological criteria approaches for *Salmonella* in chicken parts. *Microb Risk Anal 12*, 44-59. https://doi.org/10.1016/j.mran.2019.06.002

Lambertini, E., Ruzante, J. M., & Kowalcyk, B. B. (2021). The Public Health Impact of Implementing a Concentration-Based Microbiological Criterion for Controlling *Salmonella* in Ground Turkey [Article]. *Risk Analysis*, *41*(8), 1376-1395. https://doi.org/10.1111/risa.13635

Lillard, H. S. (1988). Comparison of sampling methods and implications for bacterial decontamination of poultry carcasses by rinsing. *J Food Prot*, *51*, 405-408. https://doi.org/10.4315/0362-028X-51.5.405

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)*, *22*(3), 276-282.

McMillan, E. A., Weinroth, M. D., & Frye, J. G. (2022). Increased Prevalence of *Salmonella* Infantis Isolated from Raw Chicken and Turkey Products in the United States Is Due to a Single Clonal Lineage Carrying the pESI Plasmid. *Microorganisms*, *10*(7). https://doi.org/10.3390/microorganisms10071478

Mermin, J., Hutwagner, L., Vugia, D., Shallow, S., Daily, P., Bender, J., Koehler, J., Marcus, R., Angulo, F. J., & Emerging Infections Program FoodNet Working, G. (2004). Reptiles, amphibians, and human *Salmonella* infection: a population-based, case-control study. *Clin Infect Dis*, *38 Suppl 3*, S253-261. https://doi.org/10.1086/381594

Miller, E. A., Elnekave, E., Flores-Figueroa, C., Johnson, A., Kearney, A., Munoz-Aguayo, J., Tagg, K. A., Tschetter, L., Weber, B. P., Nadon, C. A., Boxrud, D., Singer, R. S., Folster, J. P., & Johnson, T. J. (2020). Emergence of a Novel Salmonella enterica Serotype Reading Clonal Group Is Linked to Its Expansion in Commercial Turkey Production, Resulting in Unanticipated Human Illness in North America. *mSphere*, *5*(2). https://doi.org/10.1128/mSphere.00056-20

Mitzenmacher, M. (2003). A brief history of generative models for power law and lognormal distributions. *Internet Mathematics*, *1*(2), 226-251.

National Advisory Committee on Microbiological Criteria for Foods (NACMCF). (2023). *Response to questions posed by the Food Safety and Inspection Service: Enhancing Salmonella control in poultry products*. https://www.fsis.usda.gov/policy/advisory-committees/national-advisory-committee-microbiological-criteria-foods-nacmcf/2021.

National Chicken Council. (2021). *U.S. Broiler Exports Quantity and Share of Production*. Retrieved April 13 from https://www.nationalchickencouncil.org/about-the-industry/statistics/u-s-broiler-exports-quantity-and-share-of-production/

National Chicken Council. (2022). *How Broilers are Marketed*. Retrieved December 26 from https://www.nationalchickencouncil.org/statistic/how-broilers-are-marketed/

Njage, P. M. K., Leekitcharoenphon, P., & Hald, T. (2019). Improving hazard characterization in microbial risk assessment using next generation sequencing data and machine learning: Predicting clinical outcomes in *shigatoxigenic Escherichia coli*. *Int J Food Microbiol*, *292*, 72-82. https://doi.org/10.1016/j.ijfoodmicro.2018.11.016

Obe, T., Siceloff, A. T., Crowe, M. G., Scott, H. M., & Shariat, N. W. (2023). Combined Quantification and Deep Serotyping for *Salmonella* Risk Profiling in Broiler Flocks. *J Applied and Environmental Microbiology*, *89*, e02035-02022. https://doi.org/10.1128/aem.02035-22

Ollinger, M., & Bovay, J. (2020). Producer response to public disclosure of food-safety information. *Am J Agric Econ*, *102*(1), 186-201. https://doi.org/10.1093/ajae/aaz031

Oscar, T. (2021). *Salmonella* Prevalence Alone Is Not a Good Indicator of Poultry Food Safety. *Risk Analysis*, *41*(1), 110-130. https://doi.org/10.1111/risa.13563

Painter, J. A., Ayers, T., Woodruff, R., Blanton, E., Perez, N., Hoekstra, R. M., Griffin, P. M., & Braden, C. R. (2009). Recipes for foodborne outbreaks: A scheme for categorizing and grouping implicated foods. *Foodborne Pathog Dis*, *6*(10), 1259-1264. https://doi.org/10.1089/fpd.2009.0350

Painter, J. A., Hoekstra, R. M., Ayers, T., Tauxe, R. V., Braden, C. R., Angulo, F. J., & Griffin, P. M. (2013a). Attribution of foodborne illnesses, hospitalizations, and deaths to food commodities by using outbreak data, United States, 1998-2008. *Emerging infectious diseases*, *19*(3), 407-415. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84874275084&doi=10.3201%2feid1903.111866&partnerID=40&md5=c36bec185caab9ef08c0c6376f298f6d

Painter, J. A., Hoekstra, R. M., Ayers, T., Tauxe, R. V., Braden, C. R., Angulo, F. J., & Griffin, P. M. (2013b). Attribution of foodborne illnesses, hospitalizations, and deaths to food commodities, United States, 1998–2008. *Emerg Infect Dis*, *19*(3), 407-415. https://doi.org/10.3201/eid1903.111866

Petran, R. L., Grieme, L. E., & Foong-Cunningham, S. (2015). Ch. 6. Culture Methods for Enumeration of Microorganisms. In Y. Salfinger & M. L. Tororello (Eds.), *Compendium of Methods for the Microbiological Examination of Foods* (5th ed.). APHA Press, an imprint of American Public Health Association. https://doi.org/https://doi.org/10.2105/MBEF.0222

Pouillot, R., Hoelzer, K., Chen, Y., & Dennis, S. (2013). Estimating probability distributions of bacterial concentrations in food based on data generated using the most probable number (MPN) method for use in risk assessment. *Food Control*, *29*(2), 350-357. https://doi.org/10.1016/j.foodcont.2012.05.041

Powell, M. R. (2016). Trends in reported foodborne illness in the United States: 1996-2013. *Risk Anal*, *36*, 1589-1598. https://doi.org/10.1111/risa.12530

Rasamsetti, S., & Shariat, N. W. (2023). Biomapping *Salmonella* serovar complexity in broiler carcasses and parts during processing. *Food Microbiol*, *110*, 104149. https://doi.org/10.1016/j.fm.2022.104149

Rasekh, J., Thaler, A. M., Englejohn, D. L., & Pihkala, N. H. (2005). Food Safety and Inspection Service Policy for Control of Poultry Contaminated by Digestive Tract Contents: A Review. *J Appl Poult Res*, *14*(3), 603-611. https://doi.org/10.1093/japr/14.3.603

Research And Markets. (2023). *Ready Chicken Products Market Report 2022: Rapid Urbanization and Adoption of Sedentary Lifestyle Drives Growth*. B. Hathaway. https://www.businesswire.com/news/home/20230116005431/en/Ready-Chicken-Products-Market-Report-2022-Rapid-Urbanization-and-Adoption-of-Sedentary-Lifestyle-Drives-Growth---ResearchAndMarkets.com

Richardson, L. C., Bazaco, M. C., Parker, C. C., Dewey-Mattia, D., Golden, N., Jones, K., Klontz, K., Travis, C., Kufel, J. Z., & Cole, D. (2017). An updated scheme for categorizing foods implicated in

foodborne disease outbreaks: a tri-agency collaboration. *Foodborne Pathog Dis*, *14*(12), 701-710. https://doi.org/10.1089/fpd.2017.2324

Rimet, C. S., Maurer, J. J., Pickler, L., Stabler, L., Johnson, K. K., Berghaus, R. D., Villegas, A. M., Lee, M., & França, M. (2019). *Salmonella* Harborage Sites in Infected Poultry That May Contribute to Contamination of Ground Meat [Original Research]. *Front Sustain Food Syst*, *3*. https://doi.org/10.3389/fsufs.2019.00002

@Risk The Decision Tools Suite. In. (2019). (Version 7.6.1) © Palisade Corporation, All Rights Reserved.

Särndal, C. E., Swensson, B., & Wretmann, J. H. (1992). *Moel Assisted Survey Sampling*. Springer-Verlag.

Scallan, E., Hoekstra, R. M., Angulo, F. J., Tauxe, R. t. V., Widdowson, M. A., Roy, S. L., Jones, J. L., & Griffin, P. M. (2011). Foodborne illness acquired in the United States—major pathogens. *Emerg Infect Dis*, *17*(1), 7. https://doi.org/10.3201/eid1701.P11101

Shah, D. H., Paul, N. C., Sischo, W. C., Crespo, R., & Guard, J. (2017). Population dynamics and antimicrobial resistance of the most prevalent poultry-associated *Salmonella* serotypes. *Poult Sci*, *96*(3), 687-702. https://doi.org/10.3382/ps/pew342

Shorten, P. R., Pleasants, A. B., & Soboleva, T. K. (2006). Estimation of microbial growth using population measurements subject to a detection limit. *Int J Food Microbiol*, *108*(3), 369-375. https://doi.org/10.1016/j.ijfoodmicro.2005.11.024

Siceloff, A. T., Waltman, D., & Shariat, N. W. (2022). Regional *Salmonella* Differences in United States Broiler Production from 2016 to 2020 and the Contribution of Multiserovar Populations to *Salmonella* Surveillance. *Appl Environ Microbiol*, *88*(8), e00204-00222. https://doi.org/10.1128/aem.00204-22

Simmons, M., Fletcher, D. L., Berrang, M. E., & Cason, J. A. (2003a). Comparison of Sampling Methods for the Detection of *Salmonella* on Whole Broiler Carcasses Purchased from Retail Outlets. *J Food Prot*, *66*(10), 1768-1770. https://doi.org/10.4315/0362-028x-66.10.1768

Simmons, M., Fletcher, D. L., Cason, J. A., & Berrang, M. E. (2003b). Recovery of *Salmonella* from Retail Broilers by a Whole-Carcass Enrichment Procedure. *J Food Prot*, *66*(3), 446-450. https://doi.org/10.4315/0362-028x-66.3.446

Skov, M. N., Carstensen, B., Tornøe, N., & Madsen, M. (1999). Evaluation of sampling methods for the detection of *Salmonella* in broiler flocks. *J Appl Microbial*, *86*(4), 695-700. https://doi.org/10.1046/j.1365-2672.1999.00715.x

Smith, A. F. M., & Gelfand, A. E. (1992). Bayesian statistics without tears: A sampling-resampling perspective. *Am Stat*, *46*, 84-88.

Soltys, R. C., Sakomoto, C. K., Oltean, H. N., Guard, J., Haley, B. J., & Shah, D. H. (2021). High-Resolution Comparative Genomics of *Salmonella* Kentucky Aids Source Tracing and Detection of ST198 and ST152 Lineage-Specific Mutations [Original Research]. *Front Sustain Food Syst*, *5*. https://doi.org/10.3389/fsufs.2021.695368

Talukder, S., Lee, K. Y., Atwill, E. R., Pitesky, M., Lavelle, K., Gaa, M. E., Bolkenov, B., Huang, A., Tanaka, M. M., Sebti, J., Okada, Y., Giat, S. L., Yang, W., & Li, X. (2022). Prevalence of non-Typhoidal *Salmonella* in retail chicken and turkey meat in Southern California from 2018 to 2021. Reciprical Meat Conference, Des Moines, IA.

Tate, H., Folster, J. P., Hsu, C. H., Chen, J., Hoffmann, M., Li, C., Morales, C., Tyson, G. H., Mukherjee, S., & Brown, A. C. (2017). Comparative analysis of extended-spectrum-β-lactamase CTX-M-65-producing *Salmonella* enterica serovar Infantis isolates from humans, food animals, and retail chickens in the United States. *Antimicrob Agents Chemother*, *61*(7).

Teunis, P. F. M. (2022). Dose response for *Salmonella* Typhimurium and Enteritidis and other nontyphoid enteric salmonellae. *Epidemics*, *41*, 100653. https://doi.org/10.1016/j.epidem.2022.100653

Teunis, P. F. M., Kasuga, F., Fazil, A., Ogden, I. D., Rotariu, O., & Strachan, N. J. C. (2010). Dose-response modeling of *Salmonella* using outbreak data. *Int J Food Microbiol*, *144*(2), 243-249. https://www.scopus.com/inward/record.uri?eid=2-s2.0-78649503389&doi=10.1016%2fj.ijfoodmicro.2010.09.026&partnerID=40&md5=32391388d81ec0c88f02546cb3cfb629

Teunis, P. F. M., Ogden, I. D., & Strachan, N. J. C. (2008). Hierarchical dose response of E. coli O157:H7 from human outbreaks incorporating heterogeneity in exposure. *Epidemiol Infect*, *136*(6), 761-770. https://doi.org/10.1017/s0950268807008771

Thompson, C. P., Doak, A. N., Amirani, N., Schroeder, E. A., Wright, J., Kariyawasam, S., Lamendella, R., & Shariat, N. W. (2018). High-Resolution Identification of Multiple *Salmonella* Serovars in a Single Sample by Using CRISPR-SeroSeq. *Appl Environ Microbiol*, *84*(21).

Timme, R. E., Pettengill, J. B., Allard, M. W., Strain, E., Barrangou, R., Wehnes, C., Van Kessel, J. S., Karns, J. S., Musser, S. M., & Brown, E. W. (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. https://doi.org/10.1093/gbe/evt159

Tyson, G. H., Li, C., Harrison, L. B., Martin, G., Hsu, C. H., Tate, H., Tran, T. T., Strain, E., & Zhao, S. (2020). A Multidrug-Resistant *Salmonella* Infantis clone Is spreading and recombining in the United States. *Microb Drug Resist*, *27*(6), 792-799. https://doi.org/10.1089/mdr.2020.0389

U.S. Department of Health and Human Services. (2020). Healthy people topics & objectives: Food safety. *Reduce infections caused by Salmonella — FS-04*, *2021*(May 24).

https://health.gov/healthypeople/objectives-and-data/browse-objectives/foodborne-illness/reduce-infections-caused-salmonella-fs-04

U.S. Department of Health and Human Services. (2022). Healthy People Goals 2030. Foodborne Illness. *https://health.gov/healthypeople/objectives-and-data/browse-objectives/foodborne-illness* (Accessed January 28, 2023). https://health.gov/healthypeople/objectives-and-data/browse-objectives/foodborne-illness

USDA Economic Research Service. (2017). Eating out market drove chicken's popularity. https://www.ers.usda.gov/data-products/chart-gallery/gallery/chart-detail/?chartId=81931

USDA Economic Research Service. (2021). Food Availability (Per Capita) Data System. https://www.ers.usda.gov/data-products/food-availability-per-capita-data-system/

USDA Food Safety and Inspection Service. (1996a). Nationwide broiler chicken microbiological baseline data collection program. July 1994–June 1995. *www.fsis.usda.gov/OPHS/baseline/broiler1.pdf*. www.fsis.usda.gov/OPHS/baseline/broiler1.pdf

USDA Food Safety and Inspection Service. (1996b). Pathogen reduction; hazard analysis and critical control point (HACCP) systems; final rule. *Federal Register 61*, 38806-38989. www.fsis.usda.gov/OPPDE/rdad/FRPubs/93-016F.pdf

USDA Food Safety and Inspection Service. (1997a). Nationwide young turkey microbiological baseline data collection program (August 1996–July 1997). https://www.fsis.usda.gov/wps/wcm/connect/f5e01f58-d3a8-4b1a-873f-7d980aec7314/young_turkey_1996-1997.pdf?MOD=AJPERES

USDA Food Safety and Inspection Service. (1997b). Pathogen Reduction; Hazard Analysis and Critical Control Point (HACCP) Systems—Sample Collection—Technical Amendments and Corrections: Direct Final Rule.

USDA Food Safety and Inspection Service. (1998). *Microbiology laboratory guidebook*. Retrieved 09 September 2022 from https://www.fsis.usda.gov/news-events/publications/microbiology-laboratory-guidebook

USDA Food Safety and Inspection Service. (2002). E. coli O157:H7 Contamination of Beef Products. *9 CFR 417*, 62325-62334. https://www.federalregister.gov/d/02-25504

USDA Food Safety and Inspection Service. (2009a). *The nationwide microbiological baseline data collection program: Young chicken survey July 2007 - July 2008*. Washington, D.C.: U. S. Department of Agriculture. Retrieved 13 October, 2022 from https://www.fsis.usda.gov/sites/default/files/media_file/2020-07/Baseline_Data_Young_Chicken_2007-2008.pdf

USDA Food Safety and Inspection Service. (2009b). The nationwide microbiological baseline data collection program: Young chicken survey: July 2007– June 2008. http://www.fsis.usda.gov/PDF/Baseline_Data_Young_Chicken_2007-2008.pdf.

USDA Food Safety and Inspection Service. (2012). The nationwide microbiological baseline data collection program: Market hogs survey. August 2010 – August 2011. http://www.fsis.usda.gov/PDF/Baseline_Data_Market_Hogs_2010-2011.pdf

USDA Food Safety and Inspection Service. (2013). The nationwide microbiological baseline data collection program: Raw chicken parts survey. http://www.fsis.usda.gov/wps/portal/fsis/topics/data-collection-and-reports/microbiology/baseline/baseline

USDA Food Safety and Inspection Service. (2015). Public Health Effects of Raw Chicken Parts and Comminuted Chicken and Turkey Performance Standards. https://www.fsis.usda.gov/news-events/publications/public-health-effects-raw-chicken-parts-and-comminuted-chicken-and-turkey

USDA Food Safety and Inspection Service. (2016). New Performance Standards for *Salmonella* and *Campylobacter* in Not-Ready-to-Eat Comminuted Chicken and Turkey Products and Raw Chicken Parts and Changes to Related Agency Verification Procedures: Response to Comments and Announcement of Implementation Schedule. *Federal Register*, *81*(28), 7285-7300. https://www.federalregister.gov/documents/2016/02/11/2016-02586/new-performance-standards-for-salmonella-and-campylobacter-in-not-ready-to-eat-comminuted-chicken

USDA Food Safety and Inspection Service. (2017). Information on Validation of Labeled Cooking Instructions for Products Containing Raw or Partially Cooked Poultry. https://www.fsis.usda.gov/guidelines/2017-0017

USDA Food Safety and Inspection Service. (2018). Verification of poultry good commercial practices 70 FR 56624. *70 FR 56624* 56624-56626.

USDA Food Safety and Inspection Service. (2021). Sampling Instructions: *Salmonella* and *Campylobacter* verification program for raw poultry products. *10,250.1 Revision 1* https://www.fsis.usda.gov/sites/default/files/media_file/2021-03/10250.1_0.pdf

USDA Food Safety and Inspection Service. (2022a). FY 2022-2026 Food Safety Key Performance Indicator. https://www.fsis.usda.gov/inspection/inspection-programs/inspection-poultry-products/reducing-salmonella-poultry/salmonella-0

USDA Food Safety and Inspection Service. (2022b). Microbiology Laboratory Guidebook. *MLG 41.0*. https://www.fsis.usda.gov/wps/portal/fsis/topics/science/laboratories-and-procedures/guidebooks-and-methods/microbiology-laboratory-guidebook/microbiology-laboratory-guidebook

USDA Food Safety and Inspection Service. (2022c). Proposed Framework for Controlling Salmonella in Poultry. (87 FR 62784), 62784-62786. https://www.federalregister.gov/documents/2022/10/17/2022-22254/proposed-framework-for-controlling-salmonella-in-poultry

USDA Food Safety and Inspection Service. (2022d). Revised young chicken carcass exploratory sampling program. (FSIS Notice 44-22). https://www.fsis.usda.gov/policy/fsis-notice/44-22

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. *J Stat Soft*, *45*(3), 1 - 67. https://doi.org/10.18637/jss.v045.i03

Wang, J., Vaddu, S., Bhumanapalli, S., Mishra, A., Applegate, T., Singh, M., & Thippareddi, H. (2023). A systematic review and meta-analysis of the sources of *Salmonella* in poultry production (pre-harvest) and their relative contributions to the microbial risk of poultry meat. *Poult Sci*, *102*(5), 102566. https://doi.org/10.1016/j.psj.2023.102566

Wheeler, N. E., Gardner, P. P., & Barquist, L. (2018). Machine learning identifies signatures of host adaptation in the bacterial pathogen *Salmonella enterica*. *PLoS Genet*, *14*(5), e1007333. https://doi.org/10.1371/journal.pgen.1007333

Whiting, R. C., Rainosek, A., Buchanan, R. L., Miliotis, M., LaBarre, D., Long, W., Ruple, A., & Schaub, S. (2006). Determining the microbiological criteria for lot rejection from the performance objective or food safety objective. *Int J Food Microbiol*, *110*(3), 263-267. https://doi.org/10.1016/j.ijfoodmicro.2006.04.038

Williams, M. S., Cao, Y., & Ebel, E. D. (2013a). Sample size guidelines for fitting a lognormal probability distribution to censored most probable number data with a Markov chain Monte Carlo method. *Int J Food Microbiol*, *165*(2), 89-96. https://doi.org/10.1016/j.ijfoodmicro.2013.04.026

Williams, M. S., & Ebel, E. D. (2012a). Methods for fitting a parametric 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

Williams, M. S., & Ebel, E. D. (2012b). Methods for fitting the Poisson-lognormal distribution to microbial testing data. *Food Control*, *27*(1), 73-80. https://doi.org/10.1016/j.foodcont.2012.03.007

Williams, M. S., & Ebel, E. D. (2014). Fitting a distribution to censored contamination data using Markov Chain Monte Carlo methods and samples selected with unequal probabilities. *Environ Sci Technol*, *48*(22), 13316-13322.

Williams, M. S., Ebel, E. D., & Allender, H. D. (2015). Industry-level changes in microbial contamination on market hog and broiler chicken carcasses between two locations in the slaughter process. *Food Control*, *51*, 361-370. https://doi.org/10.1016/j.foodcont.2014.11.039

Williams, M. S., Ebel, E. D., & Cao, Y. (2013b). 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., & Golden, N. J. (2017). Using indicator organisms in performance standards for reducing pathogen occurrence on beef carcasses in the United States. *Microb Risk Anal*, *6*, 44-56. https://doi.org/10.1016/j.mran.2017.01.001

Williams, M. S., Ebel, E. D., Golden, N. J., Berrang, M. E., Bailey, J. S., & Hartnett, E. (2010). Estimating removal rates of bacteria from poultry carcasses using two whole-carcass rinse volumes. *Int J Food Microbiol*, *139*(3), 140-146. https://doi.org/10.1016/j.ijfoodmicro.2010.03.022

Williams, M. S., Ebel, E. D., Golden, N. J., Saini, G., Nyirabahizi, E., & Clinch, N. (2022). Assessing the effectiveness of performance standards for Salmonella contamination of chicken parts. *Int J Food Microbiol*, *378*, 109801. https://doi.org/10.1016/j.ijfoodmicro.2022.109801

Williams, M. S., Ebel, E. D., & Nyirabahizi, E. (2021). Comparative history of *Campylobacter* contamination on chicken meat and campylobacteriosis cases in the United States: 1994–2018. *Int J Food Microbiol*, *342*, 109075. https://doi.org/10.1016/j.ijfoodmicro.2021.109075

Williams, M. S., Ebel, E. D., Saini, G., & Nyirabahizi, E. (2020). Changes in Salmonella contamination in meat and poultry since the introduction of the pathogen reduction and hazard analysis and critical control point rule. *J Food Prot*, *83*(10), 1707-1717. https://search.ebscohost.com/login.aspx?direct=true&db=lah&AN=20203438056&site=ehost-live

Williams, M. S., Ebel, E. D., & Vose, D. (2011a). Framework for microbial food-safety risk assessments amenable to Bayesian modeling. *Risk Anal*, *31*(4), 548-565. https://doi.org/10.1111/j.1539-6924.2010.01532.x

Williams, M. S., Ebel, E. D., & Vose, D. (2011b). Framework for microbial food-safety risk assessments amenable to Bayesian modeling. *Risk Anal*, *31*(4), 548-565.

Williams, M. S., Ebel, E. D., & Vose, D. (2011c). Methodology for determining the appropriateness of a linear dose-response function. *Risk Analysis*, *31*(3), 345-350. https://doi.org/10.1111/j.1539-6924.2010.01518.x

Worley, J., Meng, J., Allard, M. W., Brown, E. W., & Timme, R. E. (2018). *Salmonella enterica* Phylogeny Based on Whole-Genome Sequencing Reveals Two New Clades and Novel Patterns of Horizontally Acquired Genetic Elements. *mBio*, *9*(6). https://doi.org/10.1128/mBio.02303-18

Yan, S. S., Pendrak, M. L., Abela-Ridder, B., Punderson, J. W., Fedorko, D. P., & Foley, S. L. (2004). An overview of *Salmonella* typing: Public health perspectives. *Clin Appl Immunol Rev*, *4*(3), 189-204. https://doi.org/10.1016/j.cair.2003.11.002

Appendix A EpiX Analytics' Report "Using genomics to identify nontyphoidal *Salmonella* serovars of concern, and estimating dose-response models amenable to risk assessments in poultry"

Using genomics to identify nontyphoidal *Salmonella* serovars of concern, and estimating dose-response models amenable to risk assessments in poultry

Report prepared for
The United States Department of Agriculture - Food Safety and Inspection Service
By
EpiX Analytics, Fort Collins, Colorado, USA

Original report submitted: December 16, 2022

Updated in response to OMB-required peer review: September 20, 2023

Report authors (in alphabetical order): Dr. Solenne Costard DVM, MSc, PhD; Dr. Régis Pouillot DVM, MPH, PhD; Jane Pouzou MPH, PhD; Dan Taylor DVM, MPH, PhD; Dr. Francisco Zagmutt DVM, MPVM, PhD

For any questions about this document, please contact:
Dr. Francisco Zagmutt
Managing director
info@epixanalytics.com

Contents

Executive summary

Objectives

This report describes the work performed by EpiX Analytics as part of a cooperative agreement with the United States Department of Agriculture's Food Safety and Inspection Service (FSIS) and the University of Maryland. The overarching goal of the agreement was to provide expertise and analysis to be used by FSIS as part of two risk assessments aimed at evaluating public health impact of different microbial criteria applied post-harvest in chicken and turkey (poultry) meat.

The work performed by EpiX Analytics had the following objectives:

1. Use genomics to classify serovars into groups (clusters) based on *virulence*[8] similarities
2. Use dose-response (DR) models for the serovar clusters identified under objective 1.

Methodology

The work was based on patent pending methodology originally developed by our team for application to *Salmonella* risk assessments in beef [1–3], and is summarized in **Figure 1**.

We collated and processed genetic sequences from *S. enterica* isolates from humans, animals, and animal products in the US, and performed genomic analyses to create a catalog of virulence genes for each isolate. We then employed machine learning methods to estimate the closeness of the isolates based on their virulence genes and used statistical classification methods to allocate isolates to two groups (clusters) by their virulence. We then used epidemiological data from the CDC (FoodNet) to externally validate the differences in epidemiological outcomes of virulence between the groups. We also evaluated the robustness of cluster assignments for all strains.

We estimated a DR model (*i.e.,* function), linking the number of ingested bacteria to the probability of illness, for the higher virulence cluster using literature data. Subsequently, we scaled a DR model for the lower virulence cluster so that the observed overrepresentation of strains from the higher virulence cluster in outbreaks in the US was preserved when using the DR models. We achieved this scaling by estimating risk multipliers that adjusted for the relative risks of illness from exposures to serovars belonging to different clusters, resulting from consumption of poultry.

[8] Note that in *Salmonella* microbiology, the term *virulence* is used to describe loci that affect both infectivity and virulence. For consistency with the literature, we use the term *virulence* here but both infectivity and virulence are incorporated in the clustering methods, and then further quantified in the step used to adjust the DR functions.

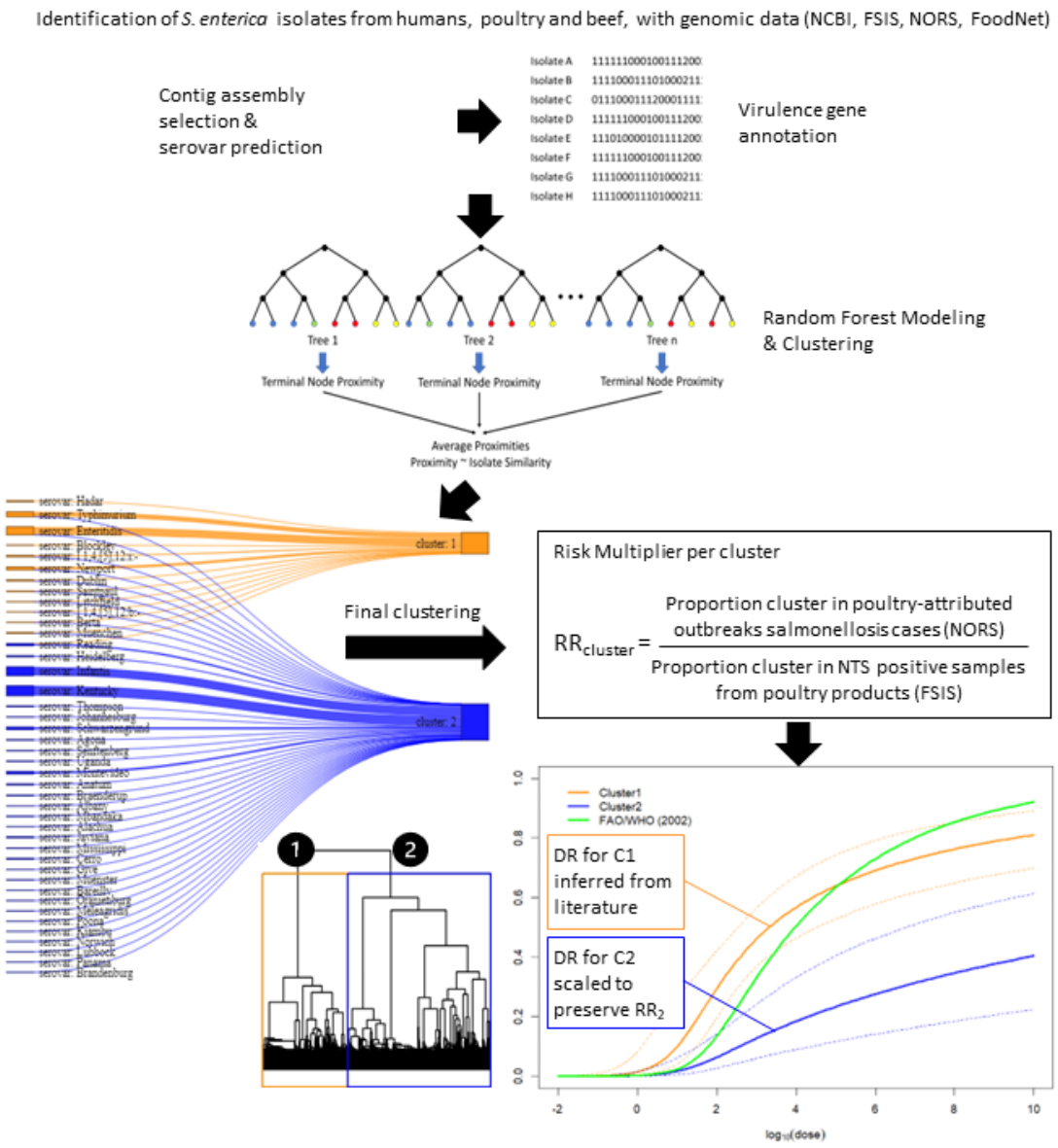


Figure 1: Overview of the analysis performed by EpiX Analytics. First, *S. enterica* isolates were retrieved from NCBI. After serovar prediction and virulence factor gene annotation, isolate assemblies were subjected to unsupervised random forest and hierarchical clustering to determine two virulence groups. Next, poultry and human salmonellosis surveillance data were used to construct risk multipliers used to scale dose-response models describing the two virulence clusters.

Results

We allocated 40,038 *S. enterica* isolates to clusters from the 61,670 isolates initially compiled from human clinical, beef, and poultry isolation sources. The allocation of serovars was stable and robust for two, three, and four clusters. Serovars composing Cluster 1 (the "higher virulence" cluster) remained consistent when allocating isolates to 2-4 clusters and was primarily composed of Enteritidis, Typhimurium, Newport, *S.* I 1,4,[5],12:i:-, and Dublin. Most remaining serovars were assigned to a single "lower virulence" cluster (Cluster 2). When we increased the number of clusters from two to three, the

majority (98%) of Kentucky isolates separated into their own cluster (Cluster 3). Kentucky remained on its own when we increased the number of clusters to four, and most Infantis isolates (88%) formed their own cluster.

Using multiple clustering methods, we allocated 13,106 (99%) of isolates from FSIS poultry *Salmonella* sampling programs to a cluster, allowing us to estimate the risk multipliers with high precision. For example, 33% [95% Confidence Interval: 31-35%] of serovars in poultry belonged to Cluster 1 ("higher virulence" cluster), while we estimated that 71% [58-83%] of human cases attributed to poultry were caused by Cluster 1 serovars. This resulted in a risk multiplier that is 2.1 times higher for Cluster 1 than that without knowing the strain belonged to Cluster 1. The reverse occurred for Cluster 2 ("lower virulence" cluster), where the infection risk was 2.6 times lower than that without knowing the strain belonged to Cluster 2.

The risk multipliers were robust to different modeling choices and type of data used, as established via a sensitivity analysis.

The risk assessment team from FSIS reviewed the clustering results together with the accompanying risk multipliers and decided to proceed with the results for two clusters (k=2, Cluster 1 and Cluster 2). Therefore, we estimated DR models for two clusters, which had remarkably different infection risks. For example, for serovars in Cluster 1, an average of 10,000 *Salmonella* cells had roughly 57% chance of resulting in an infection. In contrast, for serovars in Cluster 2, the maximum evaluated dose of 1.00E+10 cells resulted in approximately a 40% risk of infection. Table 1 summarizes the risk multipliers and top five isolates for both clusters.

Table 1: Summary of the five most frequent serovars by cluster and cluster-specific multipliers

"Higher virulence" Cluster 1 (n=15,788) Risk multiplier: 2.1 [1.7, 2.5]	"Lower virulence" Cluster 2 (n=24,250) Risk multiplier: 0.38 [0.21, 0.58]
Enteritidis, n=5,502	Kentucky, n=6,412
Typhimurium, n=3,403	Infantis, n=5,603
Newport, n=2,724	Montevideo, n=1,531
I 4,5,[5],12:i:-, n=970	Schwarzengrund, n=1,528
Dublin, n=696	Reading, n=1,273

The resulting DR models for the two clusters were provided in the form of functions in the R statistical language amenable for direct integration into FSIS' risk assessment models.

Conclusions

The methodology used in this project provides an objective, science-based framework to estimate heterogeneity in the virulence of serovars and incorporate these differences into quantitative risk assessments. The genomic grouping was validated against epidemiological data, and the model estimates were robust to different analytical and data assumptions.

Introduction

Much of the previous investigations into *Salmonella enterica* subspecies *enterica* virulence mechanisms focus upon Typhimurium as a model organism for *Salmonella* pathogenesis [4–7]. However, given the genomic and phenotypic diversity observed within *Salmonella*, we contend that some virulence models, especially site-specific gene mutations, may not be broadly applicable across the genus. To remediate this, we sought to identify genomic markers which correspond to virulence potential from a curated database of virulence genes identified from Enterobacteriaceae (family containing *Salmonella*, *Escherichia coli*, *Yersinia pestis*, *etc*.). Virulence genes in *Salmonella* are heavily influenced by gene acquisition facilitated by horizontal gene transfer [8,9] and gene loss through pseudo-gene formation [10,11]. Therefore, our goal was to cluster, or group, serovars based upon current virulence gene carriage of isolates commonly implicated in human disease in the U.S.

Accounting for different virulence in serovar groups into a quantitative risk assessment requires having group-specific DR models. The Joint FAO/WHO Expert Meetings on Microbiological Risk Assessment (JEMRA) derived a *Salmonella* DR in 2002 using outbreak data [12]. This DR model didn't consider strain variability, *i.e.,* all *Salmonella* serovar were considered equally virulent. Teunis et al. (2010) used a more sophisticated DR framework using outbreak data to fit a DR model but found no differences between serotypes and susceptibility categories [13]. Re-analyzing those outbreak data with a more flexible approach, Teunis (2022) focused on the major serotypes Enteritidis and Typhimurium, showing that Typhimurium was less infectious and has a lower probability of causing acute illness in infected subjects, but the authors didn't provide a DR model that could be used for the majority of other serotypes of public health concern [14]. Thus, our second goal was to determine DR functions for each cluster that could be used in *Salmonella* risk assessments in poultry.

Materials and Methods

Virulence clusters

Our first objective was to use genomics to group *enterica* isolates based on genetic markers of virulence (virulence factors). First, we collected *enterica* isolate genomes originating from humans, beef, and poultry (*i.e.*, chicken and turkey). Next, we assigned a serovar to each isolate and annotated the isolate assemblies with a custom database of virulence factor. To generate clusters, we fitted an unsupervised random forest model to measure isolate similarity based on the presence of virulence factors and then grouped isolates into clusters. These methods follow our previous (patent pending) work grouping *enterica* isolates from human and beef sources by virulence factors [3]. Below, we briefly outline each step within the isolate clustering protocol and provide additional references, which describe our methods in detail.

Contig assembly selection and quality criteria

S. enterica assemblies from bovine-, chicken-, and turkey-associated isolates came from three primary sources: 1) BioProject PRJNA242847 (FSIS HACCP samples), 2) BioProject PRJNA292666 (FSIS NARMS isolates), and 3) BioProject PRJNA292661 (FDA NARMS isolates). We searched the metadata for the above BioProjects for isolation sources specified as bovine-, chicken-, and turkey-associated or beef, chicken, and turkey origin.

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.

We performed quality control measures on the resultant isolate assembly dataset and applied the following exclusion criteria to generate the final assembly dataset: 1) there was no pre-computed assembly on NCBI, 2) SKESA v. 2.2 assembler did not construct the assembly, 3) > 300 contigs represented the assembly, and 4) the contig n50 < 25,000 base pairs. Finally, any serovar that represented less than 50 isolates was removed from the final assembly dataset.

Serovar prediction

The *Salmonella in silico* Typing Resource (SISTR) assigned a putative serovar to each isolate assembly [15]. 1,077 assemblies failed the subsequent quality control step within the SISTR software, but all 330 genes for the core genome multilocus sequence typing (cgMLST) scheme used to assign a serovar within the software were present within these assemblies. We retained assemblies with all 330 cgMLST loci, even if they failed SISTR software quality control, because they contained all loci necessary for the assignment of a putative serovar to an assembly. The final assembly dataset used as input for the unsupervised random forest model included 36,647 *enterica* assemblies and represented 42 serovars.

Virulence gene annotation

To determine the virulence gene catalogue carried by each *Salmonella* isolate assembly, a custom database of putative virulence factors from *Salmonella*, *Escherichia*, *Shigella*, and *Yersinia* was collated from the virulence factor database (VFDB)[16] and putative virulence factors from *Salmonella*, *Escherichia*, and *Shigella* from Bacterial and Viral Bioinformatics Resource Center (BV-BRC)[17]. We combined amino acid sequences of the open reading frames (ORF) with a reference proteome of *Salmonella* Typhimurium LT2 (https://www.uniprot.org/proteomes/UP000001014) and made the database non-redundant by clustering the open reading frames at 0.90 global identity using cd-hit [18]. We then passed the resultant database to Prokka using the "--proteins" option to specify the database as the primary annotation database in the software pipeline [19]. Additionally, to ensure consistent ORF predictions between assemblies, we trained a model using Prodigal [20] on the chromosome of the reference *Salmonella* Typhimurium LT2 assembly ASM694v2 (https://www.ncbi.nlm.nih.gov/assembly/GCF_000006945.2/) and passed to Prokka using the command "--prodigaltf". We then parsed gene annotations from the resultant Prokka annotation tables to determine the presence/absence of virulence factor genes from the VFDB and BV-BRC non-redundant database in each isolate assembly.

Random Forest model construction

After annotation of the isolate assemblies with the custom virulence factor database, we used the resultant Prokka outputs and constructed a count matrix of virulence genes for each assembly. We excluded putative virulence loci present in more than 95% of assemblies or which were found in fewer than 10 assemblies, which resulted in a final database of 193 loci. Next, we generated row similarity (isolate relatedness) by fitting an unsupervised random forest (10,000 trees, using 60 features loci at each split) to the count matrix of virulence loci (36,647 assemblies x 193 virulence factors) using the randomForest package in R [21].

Grouping isolates and assessing cluster stability

We converted the row-wise proximity matrix (isolate relatedness) output from the random forest model to a distance matrix (1 – similarity) and subjected it to agglomerative clustering using Ward's method [22]. We used the "hclust" functionality from the stats package in R to perform clustering and

bootstrapping via the "Ward.D2" method [23]. Although numerous packages are available in R which can carry out this analysis, due to the computational requirements of clustering a distance matrix for 36,647 isolates, hclust was chosen due to ease of constructing parallel functions. The number of clusters was applied to the resulting trees using the "cutree" function.

The unsupervised random forest algorithm is agnostic to the biological meaning of the virulence factor genes and will cluster observations solely based on similarity. To ensure that the clusters we found were the result of repeatable virulence factor patterns, we conducted scenario analyses to investigate the stability of the clustered results based on varying the number of *k* clusters (*i.e*., k = 2, 3, and 4). Each scenario analysis used 1,000 bootstrap iterations of the distance matrix generated from 36,647 resampled isolates which were then each clustered using the same methods. Multiple measures of stability were applied to assess the consistency of the cluster formation per bootstrap: 1) Jaccard similarity of per-isolate grouping into the same categories over a bootstrap, 2) tendency within a serotype to switch to a different cluster based on majority of isolates swapping, and 3) fraction of serotype isolates assigned to the same cluster. Cluster stability was defined as a Jaccard similarity ≥ 0.75[24].

Following the initial assignment based on unsupervised clustering, isolates which were initially excluded due to low numbers (*i.e.,* < 50) of the total serotype were assigned to clusters using a *supervised* random forest method, where isolates' clusters and their virulence factors were used as a training dataset. The supervised method is also agnostic to serotype, based only on the virulence factors of the clustered training isolates and the non-clustered isolates. This ultimately brought the number of isolates allocated to clusters to 40,038.

DR adjustment by cluster

Use of Risk Multipliers

Following oral exposure to a *Salmonella* strain *s,* the probability of becoming ill[9] given that the strain *s* belongs to cluster C_i may be written, according to Bayes theorem Equation 1) :

Equation 1

$$Pr(ill\ |s \in C_i) = Pr(ill) \times \frac{Pr(s \in C_i|ill)}{Pr(s \in C_i)}$$

We can use epidemiological data to estimate the marginal ratio $\frac{\Pr(s \in C_i|\text{ill})}{\Pr(s \in C_i)}$, over a population by computing the ratio of the proportion of individuals that are sick from a strain from Cluster C_i and the proportion of individuals that ingested a strain of Cluster C_i. We will focus specifically on individuals that acquired salmonellosis from consumption of poultry.

Using NORS and FSIS data, we estimated RR_i, the ratio of the proportion of estimated outbreak cases attributed to poultry linked to Cluster C_i with the proportion of estimated strains of Cluster C_i in poultry, as a proxy of $\frac{\Pr(s \in C_i|\text{ill})}{\Pr(s \in C_i)}$ in Equation 1.

Salmonella in poultry

We used data from the FSIS ground chicken (HC_CH_COM01), chicken parts,

[9] Notice that for simplicity here we assume that the probability of illness given infection is unity. Thus, "ill" and "infected" is used interchangeably but can be addressed separately.

(HC_CPT_LBW01/LO_CPT_LBW01), chicken carcasses (HC_CH_CARC01/LO_CH_CARC01), turkey carcasses (HC_TU_CARC01/LO_TU_CARC01), and ground turkey (HC_TU_COM01/LO_TU_COM01) sampling programs from 2016 to 2021. The isolates were assigned to one of the clusters using the following process:

- If the isolate was used in the previously described clustering process, we assigned the isolate to the cluster as it was allocated at that step;
- For isolates not included in the initial clustering process (i.e., serovars with less than 50 isolates), if it was possible to perform a complete virulence gene annotation for the isolate, we used the cluster predicted by a supervised random forest estimated from the previous classification;
- If it was not possible to perform a complete virulence gene annotation (*e.g.*, the isolate was not sent to NCBI, no assembly, *etc.*), we used a classification based on the isolate's serovar. This classification of cluster per serovar was obtained from the previous supervised random forest. In this case, we tested two modes of assignment.
    - Best cluster: we assigned the isolate to the cluster where the majority of the isolates of its serovar was assigned. For example, assuming two clusters, if the supervised random forest predicted that 20% of the strains of serovar *x* fell in Cluster 1 and 80% in Cluster 2, we set all strains of serovar *x* not yet allocated to a cluster in Cluster 2, or;
    - Proportion cluster: we allocated a value for each isolate, equal to the proportion of strains of its serovar predicted by the supervised random forest. In the previous example, strains of serovar *x* not yet allocated to a cluster would be assigned a value of 0.2 for Cluster 1 and 0.8 for Cluster 2.
- If the isolate's serovar was not one for which any other isolate with a sequence also existed, and therefore, was not sorted by any of the three steps above using genetic information, it was assigned to the lower virulence (e.g., Cluster 2 if using two clusters), assuming that the rarity of this serotype suggests it does not have high infectivity or pose a high probability of exposure.

We estimated the proportion of *Salmonella* in each cluster considering within-program weights (FSIS sampling) based on establishment production volumes and between-program weights based on total consumption rates per product. 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.

To give a lower weight to older data that might be less representative of the current situation, we used a recency weighting as described by Batz et al. (2021)[25]. The weight was 1 for data collected between 01/01/2017 and 12/31/2021 (5 years of collection). The weight for previous data decayed daily using a decay parameter of 5/7 per year (S3).

We used a non-parametric bootstrap to incorporate data uncertainty into our estimates.

Clinical cases attributed to poultry

To determine the proportion of cases attributed to chicken and turkey, we used the 1,616 recorded outbreaks in NORS from 2/4/2009 to 4/9/2021 (local report dates). Of these outbreaks, 792 have an identified food source. Chicken or turkey-attributed outbreaks were categorized as "definitive, "probable" or "possible" depending on the following NORS dataset fields: "CAFC", "FoodName", "CommoditizedFoodOrIngredient", and "IngredientName" (Supplemental figure 1).

Outbreaks classified as “definite” chicken or turkey met at least one of the following criteria:

1. CAFC = Chicken or Turkey;
2. CAFC = NA or Multiple, “FoodName” or “CommoditizedFoodOrIngredient” only contains one food and that one food is chicken or turkey;
3. CAFC = NA or Multiple, “IngredientName” only contains one ingredient, and that one ingredient is chicken/turkey.

Forty-seven of the 1,616 outbreaks included in this analysis are attributed to multiple serotypes based on samples from patients, food, and the environment. Therefore, unique outbreak-serotype combinations, or “sub-outbreaks” were used to group the outbreak-associated illnesses as is shown in the Results section regarding multipliers. The breakdown of these sub-outbreaks and their attribution to poultry is shown in detail in **Figure 3**. The total number of these sub-outbreaks extracted from the NORS databases was 1,690. We assigned a weight to each sub-outbreak equal to the proportion of strains of this serovar isolated within this outbreak.

Attribution to cluster

We assigned the strains associated with the sub-outbreaks to a cluster using a method similar to the one used for the FSIS isolates, that is 1) if isolates were used in the clustering process, using the resulting cluster assignment ; 2) if not, using the supervised random forest if a complete virulence gene annotation was obtained; 3) if not, using a classification based on its serovar (with “Best cluster” or “Proportion cluster” assignments).

Weight per outbreak (considering the number of cases per outbreak)

We tested three methods to weight each outbreak:

- Outbreak counts transformation: Applying a weight equal to 1 (potentially weighted for sub-outbreaks) to each outbreak, we used the number of outbreaks as the outcome. (*i.e.,* we didn’t consider the number of cases per outbreak);
- Estimated primary cases transformation: using the number of cases per outbreak (potentially weighted for sub-outbreaks) as estimated in the NORS database through the “estimated primary cases” field;
- IFSAC transformation: we considered the number of cases using the predicted value of a mixed-effects model, adapted from the method used by the US Interagency Food Safety Analytics Collaboration (IFSAC) for foodborne pathogen attribution based on outbreak data [25] (S3).

We also implemented a recency weighting on the outbreak data, like the one used for FSIS data. Lastly, we considered an adjustment for differential underdiagnosed cases according to severity of illness based on the proportion of bloody diarrhea reported per cluster. The adjustment was similar to the one used by Scallan et al. (2011)[26] (S4).

We used a non-parametric bootstrap to incorporate part of the epistemic uncertainty stemming from data. We did the bootstrap sampling at the level of the sub-outbreak level. The proportion of outbreaks attributed to each cluster was sampled from a Dirichlet distribution using a Bayesian framework, with Jeffrey’s (Dirichlet(α_1 = 0.5, …, α_k = 0.5)) priors. When using the IFSAC transformation, we applied the mixed model to a bootstrap sample of the complete NORS database. The procedure also considered the uncertainty linked to the sub-outbreak allocation. The uncertainty around the underdiagnosing was similar to the one used in Scallan et al. (2011)[26] (S4).

Comparison with FoodNet data

To corroborate the cluster proportion (*i.e.*, proportion of strains from each cluster) using outbreak data linked to poultry, we compared this proportion with the one that is observed for sporadic cases in the U.S. For that purpose, we identified sporadic, domestically acquired enterica isolates from the FoodNet active surveillance network (Specimen Collection Date from 01/01/2000 to 12/31/2019) as described in the previous section. We assigned the cases to the various clusters and applied recency weights using the procedure described previously for FSIS data and NORS data [25]. Recency weights were particularly meaningful for this dataset since it dates back to the year 2000. We also considered the underdiagnosed factor to estimate the proportion of sporadic cases associated to the various clusters in a similar procedure as described for NORS data. Note that sporadic cases in the FoodNet database are not assigned to a given food or food commodities.

Dose-response Models

DR model for Cluster 1 (" higher virulence")

The DR model for Cluster 1 (including Enteritidis and Typhimurium) utilized outbreak data associated to these serovars. We reproduced the Teunis et al.(2010)[13] *Salmonella* DR derivation using Teunis et al.(2008)[27] and Teunis et al.(2010)[13] framework. This framework considers an exact beta-Poisson model of infection for a given dose in a hierarchical model, *i.e*. where α and β parameters follow a variability distribution from outbreak to outbreak.

Contrary to Teunis et al. (2008) framework [27], we used a beta-Poisson model to directly calculate the probability of illness resulting from *Salmonella* exposure and thus, did not consider Teunis et al. (2008)'s model of illness given infection.[10]

We used the data provided in table 1 from Teunis et al.(2010)[13], limited to data from strains belonging to Cluster 1 strains (Enteritidis and Typhimurium), as updated in Teunis (2022)[14] . Using the R nimble package [28], and following Teunis et al.(2008)'s framework [27], we used a Bayesian hierarchical model where the transformed parameters ω and ζ follow a normal distribution from outbreak to outbreak, where $\omega_o = \text{logit}(u_o)$ and $\zeta_o = \log(v_o)$, with $u_o = \alpha_o/(\alpha_o + \beta_o)$ and $v_o = \alpha_o + \beta_o$, α_o and β_o being the parameters of an exact beta-Poisson DR for outbreak o. As in Teunis et al. (2010)[13], we considered heterogeneity in the distribution of the bacteria per meal (negative binomial distribution with parameters dose, the mean dose, and a dispersion parameter r, see Teunis et al. 2008 [27]), and hence used a ${}_2F_1$ hypergeometric confluent function of the second kind.[11] The resulting marginal probability of infection is $1-{}_2F_1(\alpha_o, r_o, \alpha_o+\beta_o, -d_o/r_o)$. See S5 for the derivation.

We obtained the posterior distributions for the hyperparameters of the beta-Poisson models that we can use to derive:

- the variability of αs and βs
- the uncertainty of the DR models

From Cluster 1 to other clusters

If we have a distribution of the exposure in the population, we can derive a DR for the less virulent

[10] Our tests suggested an overparametrized model when the infection and the illness model were considered, since no data were available for the number of infected individuals for the S. Enteritidis and S. Typhimurium outbreaks.

[11] The ${}_2F_1$ hypergeometric confluent function was rewritten using the `nimble` framework from the GNU Scientific Library (gsl) C++ library (https://www.gnu.org/software/gsl/).

strains (*e.g.*, Cluster 2) under the condition:

Equation 2

$$\int DR_{cl2}(d|\alpha',\beta')f(d) = \frac{RR_2}{RR_1}\int DR_{cl1}(d|\alpha,\beta)g(d),$$

Where $DR_{cl2}(d|\alpha',\beta')$ is the beta-Poisson DR for strains of Cluster 2, $DR_{cl1}(d|\alpha,\beta)$ is the DR for strains of Cluster 1, $f(d)$ is the density of the ingested doses of *Salmonella* of cluster 2 from poultry in the US, $g(d)$ is the density of the ingested doses of *Salmonella* of Cluster 1 from poultry in the US, RR_2 is the risk multiplier for strains of Cluster 2 and RR_1 is the risk multiplier for strains of Cluster 1. We'll assume that $f(d) = g(d)$ for all *d*, that is that the density of the ingested dose is the same whatever the cluster (see Assumptions and Discussion).

FSIS provided our team with a lognormal (LN) distribution (base 10) [12] *LN*(-3.037117, 1.279985) representing the distribution of *Salmonella* in raw poultry, and an attenuation distribution *LN*(-5.00, 1.91)[29]. This distribution adjusts the initial dose distribution by the combined effect of cooking, mixing, partitioning, cross-contamination, growth and consumption while considering variability in cooking practices. Under the reasonable assumption of independence between the original distribution and the attenuation, we can assume that $f(d)$ is a lognormal (base 10) distribution $LN\left(-3.037117 - 5.00, \sqrt{1.279985^2 + 1.91^2}\right)$.

As we have one equation (Equation 2) and two parameters (α',β'), an infinite number of solutions are possible. Following Teunis et al. (2010)[13], and Thébault et al.[13] (2013)[30], we assume that the strain variability impacts the mean $u = \alpha/(\alpha+\beta)$ of the underlying beta distribution in the beta-Poisson DR model, but that the parameter $v = \alpha + \beta$ is shared by all strains. Equation 2 can then be solved numerically.

We consider that for each set of (variable) beta-Poisson parameters for cluster 1 correspond a set of beta-Poisson parameters for Cluster 2 that fulfills the relationship stated in Equation 2 , preserving $\alpha + \beta = \alpha' + \beta'$. So, for each pair of (variable) α and β parameters, we have corresponding α' and β' parameters for Cluster 2.

We start from the output of the Bayesian process (empirical posterior distributions for ωm, ζm, ωsd, ζsd and correlation between ω_o *and* ζ_o), and derived 5,001 sets (uncertainty) of 5,000 sets (variability) of αs and βs parameters for Cluster 1. Within each iteration of uncertainty, we find the corresponding set of α' and β' that would fulfill Equation 2 for each of the 5,000 sets, using one iteration of the bootstrap sample for the RR. We repeat the process over the 5,001 iterations of uncertainty. We obtain 5,001 sets (uncertainty) of 5,000 sets (variability) of α's and β's for Cluster 2.

As, ultimately, we are interested in the mean risk (over strains, within a cluster) for a given dose and its confidence interval, we integrated this mean DR numerically using a Monte Carlo simulation. Given the computational complexity of the full DR model, we fitted a polynomial model on the obtained DR so that the model is almost instant to integrate and fully portable.

[12] $x \sim LN(\mu, \sigma)$ if $\log_{10}(x) \sim N(\mu, \sigma)$

[13] For Norovirus. Note however that P. Teunis was a co-author, so we can't consider those assumptions as having been taken independently.

Results

Serovar assignment

In total, there are over 400,000 *Salmonella* isolates housed in the Pathogen Detection Network hosted by NCBI. We extracted 61,670 isolates from the four previously described BioProjects. Based on the exclusion criteria described in the methods section, we further reduced the number of extracted isolates to a final analysis set of 36,647 *enterica* assemblies representing human clinical cases in the US and poultry and beef associated isolates. Within this dataset, which was used as the unsupervised random forest input, 18.4% (6,751) assemblies came from US human clinical infections, 15.2% (5,586) represented isolates from bovine sources, and the remaining 66.3% (24,310) isolates originated from poultry. We assigned a cluster to an additional 3,391 isolates that were initially excluded via supervised random forest, bringing the total number of isolates allocated to a cluster to 40,038.

Serovar assignment for *k*=2, 3, and 4 clusters are provided in Table 2. The serovars composing Cluster 1 remained consistent at the three levels of *k* (**Figure 2**). When *k* was increased from 2 to 3, the majority (98%) of Kentucky isolates separated into their own cluster (Cluster 3, when k=3), while Infantis belonged to Cluster 2. Kentucky remained on its own when *k* was increased to 4 (Cluster 4, when k=4) and most Infantis isolates (88%) formed their own cluster (Cluster 3, when k=4). The remaining serovars comprising Cluster 2 in the *k*=2 designation continued to cluster together as *k* increased to 3 and 4. Isolates (*i.e*., non-serotyped) which were not assigned a serovar due to missing "O" or "H" antigens (*n*=26) may comprise a group of diverse serovars, which split between Cluster 1 and 2 for all levels of *k* based on supervised random forest.

Table 2 includes the serotype names as reassigned using the SISTR methodology, and therefore does not capture all unique partial serotypes that might be found in the FSIS and NORS datasets. If new data is added to this analysis, the serotypes should be characterized using genetic information to assign a cluster to the new isolates, or the cluster should be assigned to the isolate via supervised random forest.

Table 2: Serovar cluster assignments for k= 2, 3 and 4.

		2 Clusters (*k*=2)		3 Clusters (*k*=3)			4 Clusters (*k*=4)			
Serovar	*n*	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 3	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Adelaide	40		100%		100%			100%		
Agbeni	8		100%		100%			100%		
Agona	406		100%		100%			100%		
Alachua	67		100%		100%			100%		
Albany	88		100%		100%			100%		
Anatum	673		100%		100%			100%		
Baildon	6	100%		100%			100%			
Bareilly	143	5%	95%	5%	95%		5%	95%		
Berta	193	98%	2%	98%	2%		98%	2%		
Blockley	163	100%		100%			100%			
Bovismorbificans	56	100%		100%			100%			
Braenderup	525	1%	99%	1%	99%		1%	99%		
Brandenburg	111		100%		100%			100%		
Calabar	1		100%		100%			100%		
Carrau	93		100%		100%			100%		
Cerro	591		100%		100%			100%		
Chailey	7	100%		100%			100%			
Chester	2		100%		100%			100%		
Concord	6		100%		100%			100%		
Cubana	2		100%		100%			100%		
Dublin	697	100%		100%			100%			
Duisburg	1		100%		100%			100%		
Eastbourne	10		100%		100%			100%		
Enteritidis	5510	100%		100%			100%			
Gaminara	1		100%		100%			100%		
Gateshead	1		100%		100%			100%		
Give	164		100%		100%			100%		
Goldcoast	1		100%		100%			100%		
Hadar	558	100%		100%			100%			
Hartford	3		100%		100%			100%		
Heidelberg	728		100%		100%			100%		
Hillingdon	1	100%		100%			100%			
Hvittingfoss	1		100%		100%			100%		
I 1,4,[5],12:b:-	108	95%	5%	95%	5%		95%	5%		
I 1,4,[5],12:i:-	987	98%	2%	98%	2%		98%	2%		
Idikan	1		100%		100%			100%		
Infantis	5604		100%		100%			12%	88%	
Javiana	971		100%		100%			100%		
Johannesburg	158		100%		100%			100%		

Serovar	n	Cluster 1	Cluster 2	Cluster 1	Cluster 2	Cluster 3	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Kentucky	6413		100%		2%	98%		2%		98%
Kiambu	65		100%		100%			100%		
Koessen	3		100%		100%			100%		
Leiden	4		100%		100%			100%		
Litchfield	120	98%	2%	98%	2%		98%	2%		
Livingstone	1		100%		100%			100%		
Lomalinda	4		100%		100%			100%		
Lubbock	103		100%		100%			100%		
Manhattan	4	100%		100%			100%			
Mbandaka	370		100%		100%			100%		
Meleagridis	138		100%		100%			100%		
Miami	47		100%		100%			100%		
Mississippi	263		100%		100%			100%		
Montevideo	1533		100%		100%			100%		
Muenchen	607	94%	6%	94%	6%		94%	6%		
Muenster	258		100%		100%			100%		
Newport	2740	99%	1%	99%	1%		99%	1%		
Norwich	78		100%		100%			100%		
Ohio	11		100%		100%			100%		
Okatie	1		100%		100%			100%		
Oranienburg	203	1%	99%	1%	99%		1%	99%		
Panama	66		100%		100%			100%		
Poona	57		100%		100%			100%		
Potsdam	1		100%		100%			100%		
Reading	1299	2%	98%	2%	98%		2%	98%		
Rissen	6		100%		100%			100%		
Rubislaw	24		100%		100%			100%		
Saintpaul	612	100%		100%			100%			
Sandiego	2		100%		100%			100%		
Schwarzengrund	1528		100%		100%			100%		
Senftenberg	327		100%		100%			100%		
Stanley	47	100%		100%			100%			
Telelkebir	12		100%		100%			100%		
Thompson	549		100%		100%			100%		
Typhimurium	3421	99%	1%	99%	1%		99%	1%		
Uganda	356		100%		100%			100%		
Urbana	18		100%		100%			100%		
Vinohrady	14		100%		100%			100%		
Virchow	7		100%		100%			100%		
Weltevreden	14		100%		100%			100%		
Non-serotyped	26	31%	69%	31%	69%		31%	69%		

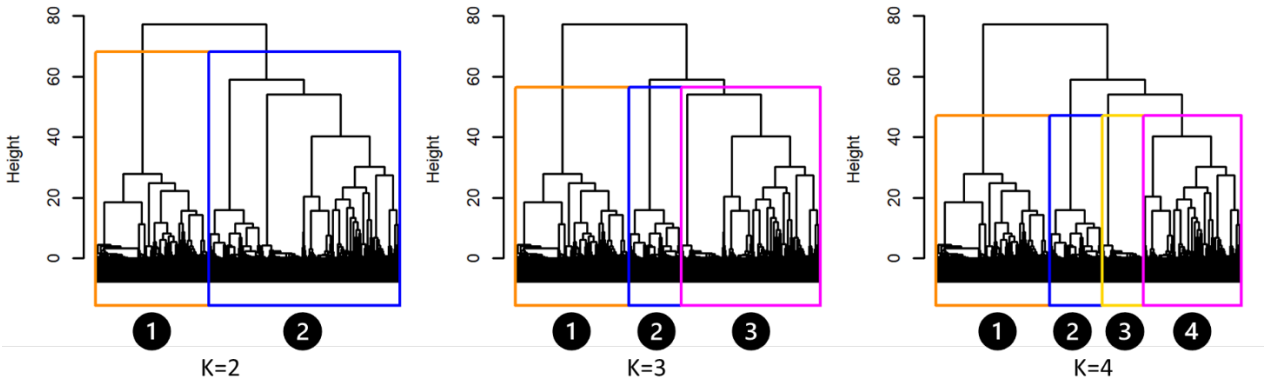


Figure 2: Dendrograms showing isolate groupings for k=2,3, and 4. Numbers correspond to cluster number listed in **Table 2** headings above. Note that Cluster 1 remains consistent as k increases.

Robustness of serovar assignments

Jaccard Similarity

The mean bootstrap Jaccard similarity for all clusters within the *k*=2, 3, and 4 designations was above the 0.75 threshold, indicating cluster stability.

Serotype Switching

Berta (*n*=193) and Saintpaul (*n*=612) isolates switched clusters within the bootstrap samples most often. For *k*=2,3 and 4, the percent of Berta and Saintpaul isolates that switched from Cluster 1 to Cluster 2 in more than 5% of the bootstraps was 100% and 79.4%, respectively. Among the remaining serovars, the level of isolate switching in the bootstrap samples was low, indicating stability of serovar cluster assignments and relatively low variation in virulence factors within serotypes.

Division of serovars between multiple clusters

For *k*=4, a large majority (≥95%) of isolates fall into a single cluster with the remarkable exception of Infantis for which 12% of the isolates (672/5,604) are classified as Cluster 2 while 88% (4,932/5,604) are classified as Cluster 3. However, for *k*=2 and *k*=3, Infantis does not split between clusters and all isolates reside within Cluster 2. The genes most responsible for the split of Infantis isolates into two clusters when *k*=4 are located on the pESI megaplasmid [31]. Most notably, these genes are necessary to produce yersiniabactin, which is a siderophore dependent iron uptake system [32].

Multiplier Estimation

Results

A total of 13,537 isolates were extracted from the previously described FSIS sampling program databases (chicken carcasses, parts and comminuted and turkey carcasses and comminuted) using our selection criteria, resulting in 13,241 recency-weighted strains. We were able to assign a cluster *via* random forest or supervised random forest to 9,578 of these weighted isolates. Using serovar assignments on the remaining strains, we allocated 13,106 (99%) weighted isolates to a cluster.

The weighted (recency, establishment and between products) proportion of strains in poultry is

provided for *k*=2 (Table 3) and *k* = 3 (Table 4). The proportions of strains in outbreaks were derived using our baseline scenario based on the previously described IFSAC transformation, outbreak recency weighting (5 years), and outbreaks definitively attributed to chicken and turkey. Serovars without assemblies were assigned to a cluster based on the previously described "proportion cluster" method.

Table 3: Multipliers for k = 2 (Estimate [bootstrap 95%CI])

	Cluster 1	Cluster 2	Not Assigned
Proportion in Poultry	0.33 [0.31, 0.35]	0.66 [0.64; 0.68]	0.010 [0.006; 0.017]
Proportion in outbreaks	0.71 [0.58; 0.83]	0.25 [0.14; 0.38]	0.039 [0.012; 0.081]
Multiplier	2.1 [1.7; 2.5]	0.38 [0.21; 0.58]	3.9 [1.1; 9.1]

Table 4: Multipliers for k = 3 (Estimate [bootstrap 95%CI])

	Cluster 1	Cluster 2	Cluster 3	Not Assigned
Proportion in Poultry	0.33 [0.31, 0.35]	0.42 [0.40; 0.44]	0.24 [0.23; 0.25]	0.010 [0.006; 0.017]
Proportion in outbreaks	0.71 [0.58; 0.82]	0.25 [0.14; 0.38]	0.002 [0.000; 0.021]	0.038 [0.012; 0.080]
Multiplier	2.1 [1.7; 2.5]	0.60 [0.34; 0.90]	0.01 [0.000; 0.088]	3.9 [1.1; 9.2]

Forty-seven of the 1,616 outbreaks included in this analysis are attributed to multiple serotypes based on samples from patients, food, and the environment, and some serotypes differ in cluster assignment. Of the 47 outbreaks linked to more than one serotype, 22 had at least two serotypes which sorted into different clusters (when *k*=2). Three of these 22 multi-cluster outbreaks were linked to chicken and none to turkey. Therefore, unique outbreak-serotype combinations, or "sub-outbreaks" were used to group the outbreak-associated illnesses (**Figure 3**). The total number of these sub-outbreaks extracted from the NORS databases was 1,690 - 216 of which were attributed to poultry (191 using our "definitively" definition). To each sub-outbreak, we assigned a cluster: first via random forest (*n*=51), then by supervised random forest (*n*=9), and then by assignment according to serotype (*n*=134), so that a total of 194 sub-outbreaks attributed to poultry were included (**Figure 3**). Applying the recency weight system reduced the influence of sub-outbreaks which occurred before 2017, so that the apparent weighted number of sub-outbreaks used were 118 (108 of these were definitively linked to poultry).

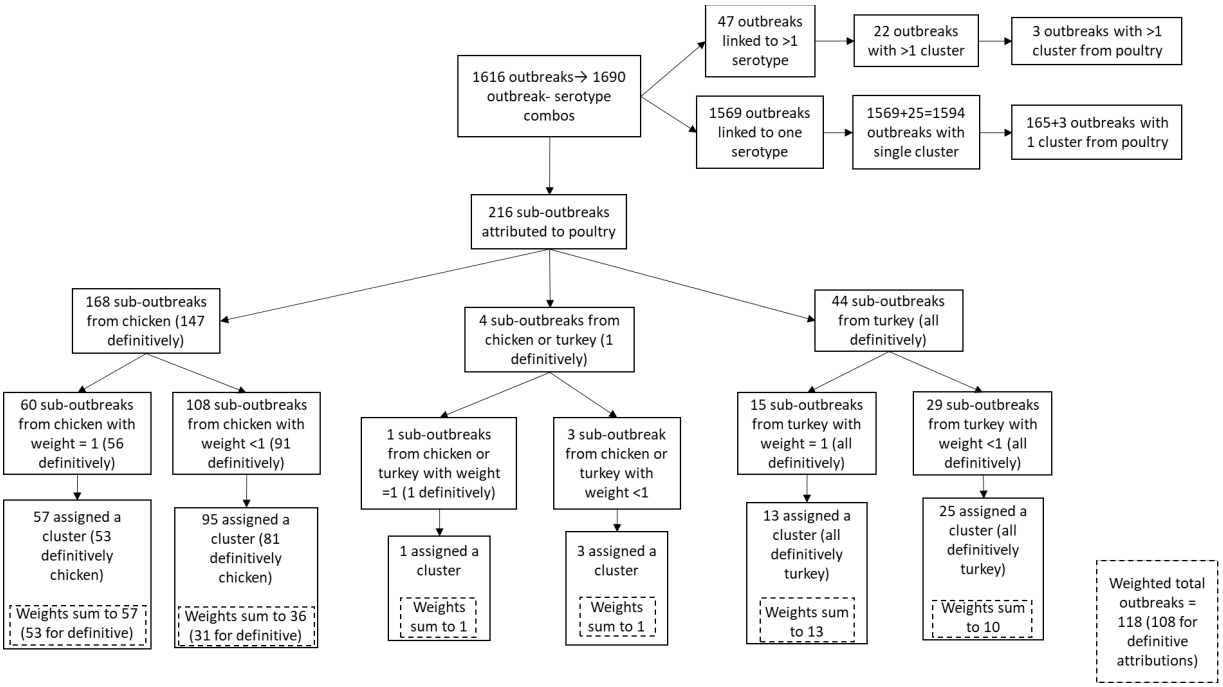


Figure 3: Distribution of outbreaks, sub-outbreaks, and weighted total outbreaks used from NORS data.

We collated 132,326 sporadic, domestically acquired cases the FoodNet database. Applying the recency weight led to 43,882 weighted cases as a majority of cases were recorded before 2017. We assigned a cluster *via* unsupervised random forest or supervised random forest for 6,133 of these weighted cases. Ultimately, using serovar, we assigned a cluster to 37,679 weighted cases (86%). For *k*=3, the proportion of cases attributed to each cluster was 63% [CI95%: 62; 63%] 37% [37; 38%] and 0.13% [0.09; 0.17%] for Cluster 1, 2, and 3, respectively, which is comparable to what was observed from the NORS outbreak data. For this estimation, we were unable to assign a large number of isolates (15%) to a given cluster.

Assessing sensitivity of risk multipliers to modeling options

We tested how the risk multipliers changed under the various options described for this analysis. The results were robust to modeling options, with the most impactful options being not using recency weighting in the data, or only using turkey data (Table 5).

Table 5: Sensitivity of risk multipliers to different modeling and data transformation options

	Cluster 1	**Cluster 2**
Baseline*	2.1 [1.7; 2.5]	0.38 [0.21; 0.58]
Outbreak counts transformation	2.0 [1.6; 2.4]	0.39 [0.24; 0.55]
Estimated Primary cases transformation	2.0 [1.3; 2.5]	0.51 [0.22; 0.84]
No recency weighting	1.8 [1.4; 2.1]	0.53 [0.36; 0.70]
Recency weight starting to decrease after 1 year	2.4 [1.8; 2.9]	0.32 [0.14; 0.58]
Turkey only	1.7 [0.77; 3.0]	0.65 [0.22; 1.10]
Chicken only	2.2 [1.8; 2.6]	0.32 [0.16; 0.52]
Do not weight different products	2.4 [1.9; 2.9]	0.36 [0.21; 0.56]
Outbreaks Definitively or Probably attributed to poultry	2.1 [1.8; 2.4]	0.39 [0.24; 0.55]
Use best Cluster	2.1 [1.7; 2.5]	0.38 [0.21; 0.59]

* Baseline: IFSAC transformation, recency weighting (5 years), use chicken data, use turkey data, use FSIS weights for different products, use outbreaks definitively attributed to chicken (resp. turkey), use proportion of cluster. Results from unattributed isolates not presented. Bootstrap used 1001 iterations.

Interpretation of risk multipliers

Recall Equation 1: $\Pr(\text{ill} \mid s \in C_i) = \Pr(\text{ill}) \times \frac{\Pr(s \in C_i \mid \text{ill})}{\Pr(s \in C_i)}$, that allows us to calculate the probabilities of illness following exposure to *Salmonella* from a given cluster.

The multipliers allow us to state that if we have an exposure from a strain belonging to Cluster 1 (e.g., Enteriditis), the risk of illness is 2.1 times [CI95%: 1.7; 2.5] higher than the probability of illness prior to knowing that the strain belonged to Cluster 1. Similarly, knowing that the strain is from Cluster 2 informs that the risk of illness is 1/0.38 = 2.63 times [CI95%: 1/0.58 = 1.74; 1/0.21 = 4.76] lower than the probability of illness without knowing the strain belonged to Cluster 2.

A note on FSIS' decision on the number of clusters to use for further analysis

Our team participated in weekly calls and discussed with FSIS the different results of the analysis. Particular attention was paid to the robustness of the allocation of serovars based on different number of clusters used, and how this translated in different risk multipliers. As described earlier, most of the serovars were stable and the allocations changed mostly for Infantis and Kentucky when increasing the number of clusters.

Using the information that we provided for 2-4 clusters combined with the risk multipliers, FSIS decided to proceed with the DR model adjustments for two clusters. Below we provide a summary of the serovars and multipliers for Cluster 1 ("higher virulence") and Cluster 2 ("lower virulence") (Table 6).

Table 6: Summary of the five most frequent serovars by cluster and cluster-specific multipliers

"Higher virulence" Cluster 1 (*n*=15,788) Risk multiplier: 2.1 [1.7, 2.5]	**"Lower virulence" Cluster 2** (*n*=24,250) Risk multiplier: 0.38 [0.21, 0.58]
Enteritidis, n=5,502	Kentucky, n=6,412
Typhimurium, n=3,403	Infantis, n=5,603
. Newport, n=2,724	Montevideo, n=1,531
. I 4,5,[5],12:i:-, n=970	Schwarzengrund, n=1,528
Dublin, n=696	Reading, n=1,273

Dose-response models

Figure 4 illustrates the fit of the DR model to outbreak data from Cluster 1, using data from Teunis et al. (2022)[14], which resulted in large uncertainty and very large variability in the DR models.

Here, "dose" is the parameter of a Poisson distribution considering serving-to-serving variability. Hence the dose is not an integer value as it represents an average or intensity parameter.

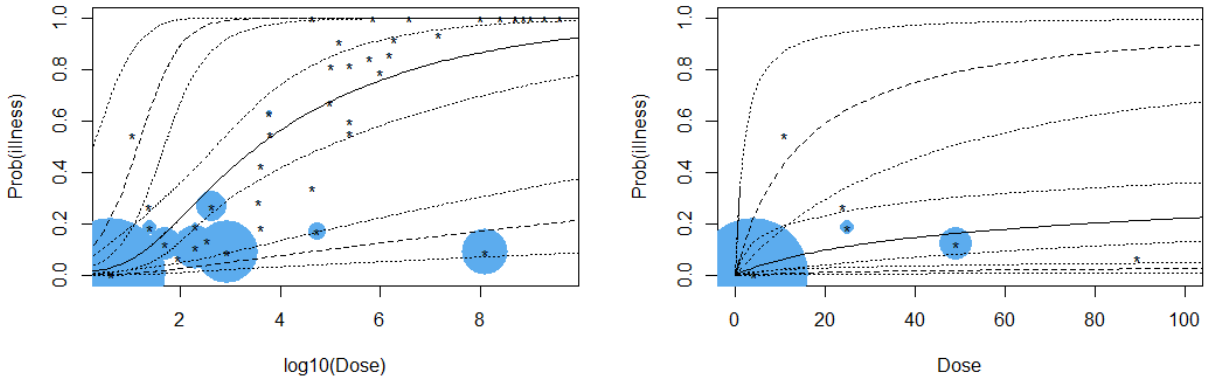


Figure 4: DR model fitted to Teunis (2022) data. Asterisks represent the proportion of individuals exposed to *S*. Enteritidis or *S*. Typhimurium that became ill from individual outbreaks. with blue radius proportional to the number of individuals in the outbreak. Curves (from bottom to top) represents: 1) the 2.5th uncertainty of the 2.5th variability, 2) the median of the 2.5th variability, 3) the 97.5th uncertainty of the 2.5th variability, 4) the 2.5th uncertainty of the median variability, 5 plain black) the median (uncertainty) of the median (variability), 6) the 97.5th uncertainty of the median variability, 7) the 2.5th uncertainty of the 97.5th variability, 8) the median uncertainty of the 97.5th variability, 9) the 97.5th uncertainty of the 97.5th variability. left: x is the $\log_{10}$(dose), right: x is the dose, up to 100 bacteria.

Figure 5 illustrates the DR model for *Salmonella* from Cluster 1 and Cluster 2. The value is the marginal

probability of infection after integration over strains within a cluster[14], over doses (Poisson distribution with intensity "dose") and over individuals[15]. We also include the FAO/WHO (2002) DR model for comparison.

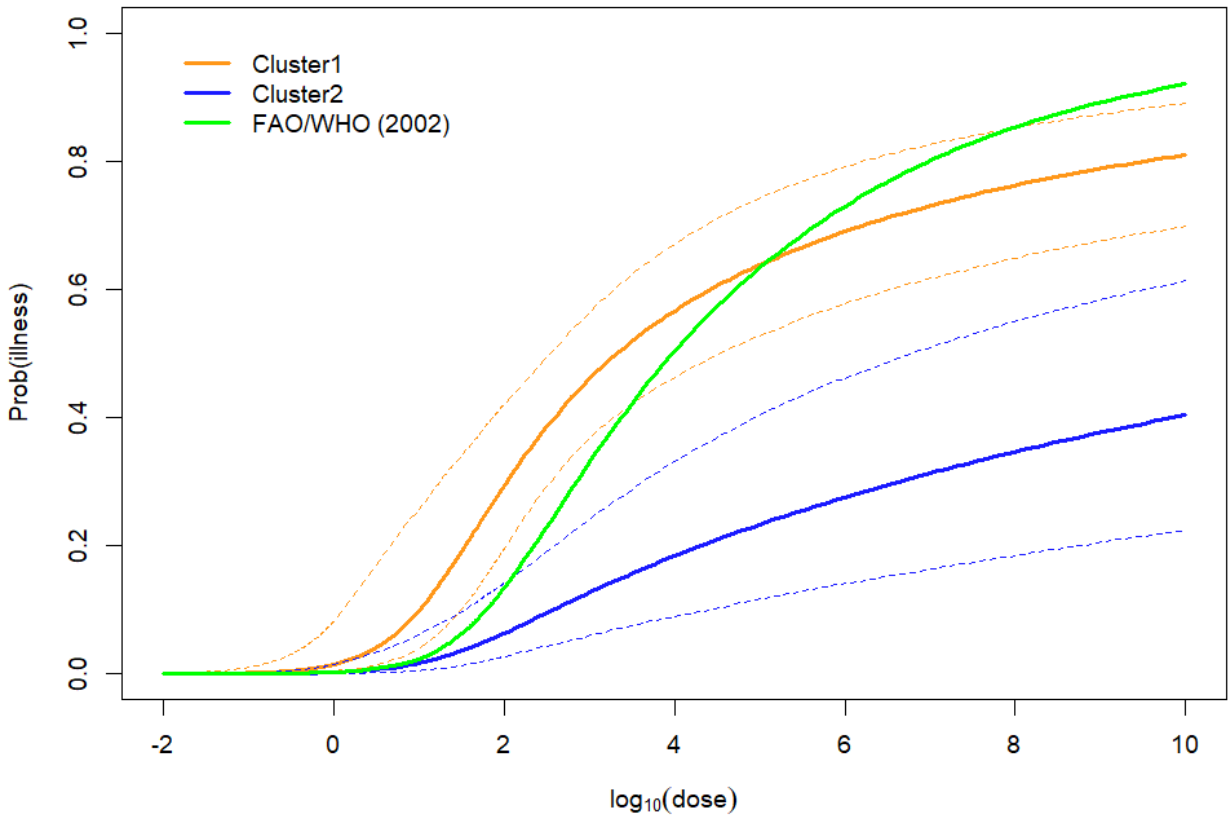


Figure 5: Average (over strains) DR model for Cluster 1 (orange), Cluster 2 (blue), compared to FAO/WHO (2002) dose response (green). Plain lines: best estimates (median of the values in the uncertainty dimension). Dotted line: 95% confidence intervals (2.5th and 97.5th quantiles in the uncertainty dimension).

Figure 6 provides a "zoomed-in" view of the DR model for lower doses and using a $\log_{10}$ y-axis. Note the (log-log) linearity of the DR model at these low doses, and the similarity between the Cluster 1 and FAO/WHO DR curves.

[14] As the model is integrated over strains, these values should not be used to estimate the expected number of cases for a given outbreak, where a single strain is involved.

[15] The integration over the dose (assuming a Poisson distribution) and over the individuals (assuming a beta distribution) are considered in the use of the underlying beta-Poisson DR function. The integration over strains within a cluster was done by averaging 5000 beta-Poisson DR models considering α and β strain-to-strain variability.

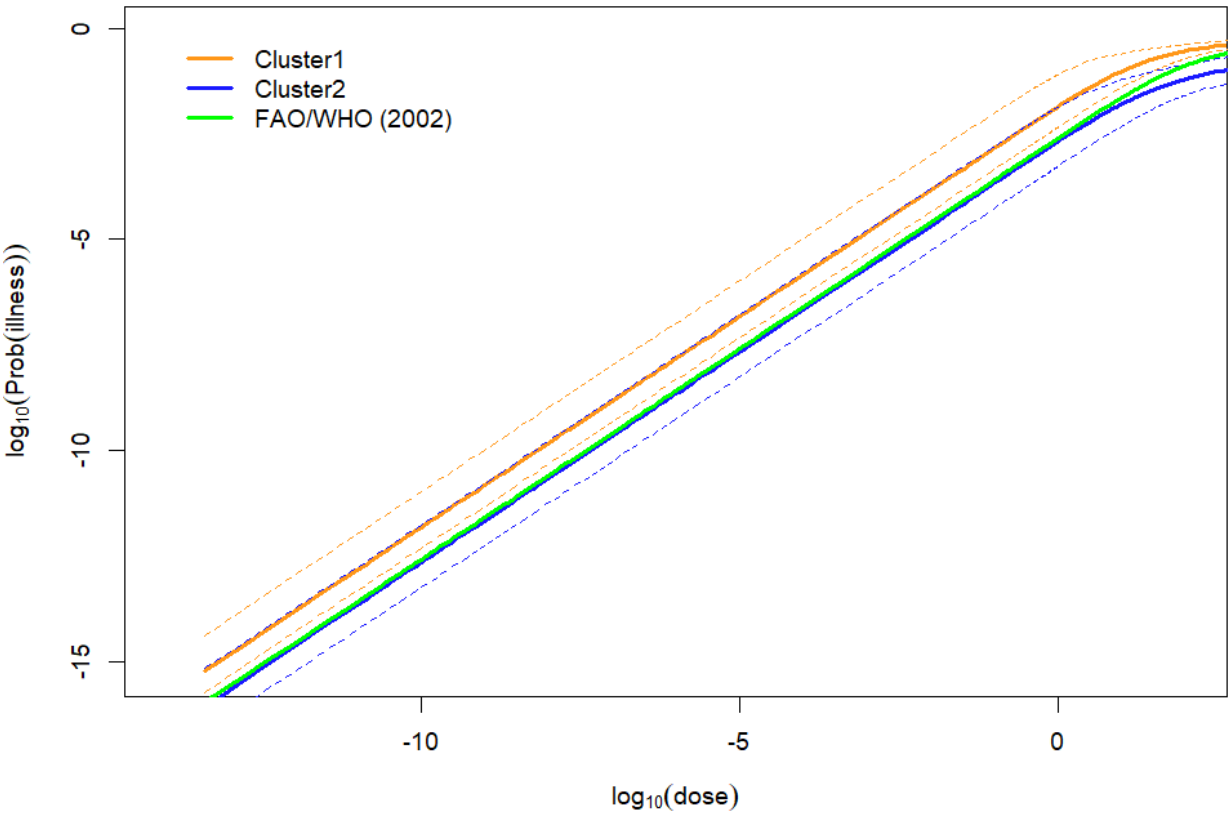


Figure 6: Average (over strains) DR model for Cluster 1 (orange), Cluster 2 (blue), compared to FAO/WHO (2002) DR (green). Note: the DR model for Cluster 2 is mostly hidden by the FAO/WHO DR. Plain lines: best estimates (median of the values in the uncertainty dimension). Dotted line: 95% confidence intervals (2.5th and 97.5th quantiles in the uncertainty dimension).

Table 7 provides the coefficients for a polynomial approximation that can be used to derive the DR model for Cluster 1 and Cluster 2 without having to repeat the inference. We checked the fit of the polynomial function for a dose ranging from 10^{-14} to 10^{10} cfu. Using these figures, the probability of illness for a given dose of *Salmonella* from a given cluster can be obtained using[16]:

Equation 3 $Prob(illness) = coef1 \times \ln(Dose+1) + coef2 \times (\ln(Dose+1))^2 + \ldots + coef9 \times (\ln(Dose+1))^9.$

Where "Dose" is the intensity parameter of the Poisson distribution describing the number of bacteria from serving to serving in the subpopulation of interest. To achieve reliable precision in this calculation, we recommend using the R functions provided as an output of this project.

[16] ln is $\log_e$ (logarithm, base e)

Table 7: Polynomial regression of the probability of infection as a function of the dose, for strains from Cluster 1 and 2 (2 Cluster). Validated from dose = 0 bacteria to dose = 1E10 bacteria.

	Cluster 1			Cluster 2/2		
	Estimate	Lower CI95%	Upper CI95%	Estimate	Lower CI95%	Upper CI95%
ln(Dose*+1)	1.677793E-02	8.028926E-03	1.044496E-01	2.547083E-03	8.850745E-04	1.635746E-02
$(\ln(Dose+1))^2$	-8.965997E-04	-8.483557E-03	2.859753E-02	-1.917973E-04	-7.198445E-04	6.578363E-03
$(\ln(Dose+1))^3$	9.486076E-03	8.090451E-03	-2.025913E-02	1.247473E-03	6.752766E-04	-2.176502E-03
$(\ln(Dose+1))^4$	-2.710090E-03	-1.421906E-03	5.193496E-03	-1.920443E-04	-5.763324E-05	5.905104E-04
$(\ln(Dose+1))^5$	3.473447E-04	9.426093E-05	-7.006729E-04	8.414753E-06	-5.246505E-06	-8.899356E-05
$(\ln(Dose+1))^6$	-2.460849E-05	-9.720838E-07	5.396930E-05	3.616453E-07	1.131561E-06	7.362400E-06
$(\ln(Dose+1))^7$	9.981556E-07	-1.788357E-07	-2.386504E-06	-4.844387E-08	-7.339216E-08	-3.394596E-07
$(\ln(Dose+1))^8$	-2.176171E-08	8.410247E-09	5.642269E-08	1.711068E-09	2.146805E-09	8.220825E-09
$(\ln(Dose+1))^9$	1.981074E-10	-1.155037E-10	-5.529915E-10	-2.095305E-11	-2.413967E-11	-8.166917E-11

* dose is the parameter of the Poisson parameter describing the distribution of dose from serving to serving in number of bacteria, ln is logarithm in base *e*. The formula has no intercept.

In addition to the figures illustrating the DR curves, we provide a summary of the probability of illness by cfu (dose) for the two clusters (Table 8). Only a finite set of dose values are provided in the table, but Equation 3 can be used to derive this probability for any dose.

Table 8: Probability of illness as a function of the dose for the two cluster.

Dose	Cluster 1			Cluster 2		
	Estimate	Lower CI95%	Upper CI95%	Estimate	Lower CI95%	Upper CI95%
1.00E-14	1.68E-16	8.02E-17	1.04E-15	2.55E-17	8.84E-18	1.63E-16
1.00E-13	1.68E-15	8.02E-16	1.04E-14	2.55E-16	8.84E-17	1.63E-15
1.00E-12	1.68E-14	8.03E-15	1.04E-13	2.55E-15	8.85E-16	1.64E-14
1.00E-11	1.68E-13	8.03E-14	1.04E-12	2.55E-14	8.85E-15	1.64E-13
1.00E-10	1.68E-12	8.03E-13	1.04E-11	2.55E-13	8.85E-14	1.64E-12
1.00E-09	1.68E-11	8.03E-12	1.04E-10	2.55E-12	8.85E-13	1.64E-11
1.00E-08	1.68E-10	8.03E-11	1.04E-09	2.55E-11	8.85E-12	1.64E-10
1.00E-07	1.68E-09	8.03E-10	1.04E-08	2.55E-10	8.85E-11	1.64E-09
1.00E-06	1.68E-08	8.03E-09	1.04E-07	2.55E-09	8.85E-10	1.64E-08
1.00E-05	1.68E-07	8.03E-08	1.04E-06	2.55E-08	8.85E-09	1.64E-07
1.00E-04	1.68E-06	8.03E-07	1.04E-05	2.55E-07	8.85E-08	1.64E-06
1.00E-03	1.68E-05	8.02E-06	1.04E-04	2.55E-06	8.84E-07	1.64E-05
1.00E-02	1.67E-04	7.91E-05	1.04E-03	2.53E-05	8.74E-06	1.63E-04
1.00E-01	1.60E-03	6.95E-04	1.02E-02	2.42E-04	7.84E-05	1.62E-03
1.00	1.38E-02	3.87E-03	8.05E-02	2.05E-03	4.78E-04	1.39E-02
10.00	9.96E-02	4.22E-02	0.26	1.66E-02	5.16E-03	6.08E-02
100.00	0.29	0.19	0.41	6.25E-02	2.61E-02	0.14
1000.00	0.46	0.36	0.56	0.13	5.87E-02	0.24
10000.00	0.57	0.47	0.68	0.18	9.00E-02	0.33
100000.00	0.64	0.53	0.74	0.23	0.12	0.40
1.00E+06	0.69	0.57	0.78	0.27	0.14	0.46
1.00E+07	0.73	0.62	0.83	0.31	0.16	0.51
1.00E+08	0.76	0.65	0.85	0.35	0.18	0.55
1.00E+09	0.79	0.67	0.87	0.38	0.20	0.58
1.00E+10	0.81	0.69	0.88	0.40	0.22	0.61

Assumptions, their implications, and justification

We provide a list of the main assumptions made to derive our estimates (Table 9). Next to each assumption we list their implications for the work and possible use in a quantitative risk assessment, and our justification(s) for making the assumption with references, where applicable.

Table 9: Table of assumptions made during phases of project with implications and justifications for making assumptions

Index	Assumption	Meaning / implications	Justification / Comment
Serovar clustering			
1	Serovar virulence cluster does not depend on host species (e.g., chicken vs. turkey)	Isolates from multiple species can be included in the same clustering analysis, to increase power of analysis without significantly affecting results.	Prior analysis indicates that when isolates from multiple species [3] (i.e., humans and bovine animals) are included in the clustering algorithm, isolates categorize into the same clusters regardless of the species of origin.
2	Only pre-assembled isolate contigs from NCBI were annotated. We assume isolates that were not pre-assembled were missing at random and would not change the results of the clustering.	By using only pre-assembled contigs, our analysis is faster, while still being complete.	Including NCBI isolates that were not pre-assembled would have taken weeks to months to download and assemble. Thus, this approach was unfeasible given the time frame of this project.
Derivation of Multipliers – Use of Multipliers as posterior probability			
3	Serotype proportion inferred from FSIS data are representative of the serotype proportion in the US poultry product supply	The risk multipliers derived from surveillance data apply to the entire US poultry product supply	See FSIS works.
4	Outbreak-based attribution of Salmonellosis to different food sources based on CDC data from NORS is representative from cases of *Salmonellosis* in the US population	The risk multipliers derived from surveillance data apply to the US population	See IFSAC works. There is empirical support for the comparability of sporadic and outbreak-associated foodborne illnesses [33]. Also, the CDC reports similar salmonella attribution to poultry using outbreak data, vs a genetic method using sporadic data [34].

Use of Multipliers as Dose-Response adjustment			
5	Concentration of *Salmonella* on product does not differ between serovar clusters	No need to adjust D-R calculations for different distributions of level	Sub-analysis of FSIS MPN data from ground beef and poultry (results not shown) indicates level of *Salmonella* did not differ significantly by cluster.
6	*Salmonella* inactivation and growth don't differ according to the cluster	The same attenuation factor was applied to NTS nontyphoidal *Salmonella* (NTS) in all clusters	Literature suggests that *Salmonella* resistance to heat treatment and growth parameter varies more within serotypes than between serotypes. We couldn't find in the literature any relevant research suggesting a significant differences in growth and/or inactivation according to serovar. Furthermore, as the type of NTS in product would be unknown, it is reasonable to assume that the process, cooking, and handling of animal meats would be the same regardless of the serovar present.
7	*Salmonella* inactivation and growth don't differ according to the product (chicken, turkey, parts, ground, carcasses)	There is no need to explicitly include this in the analysis, since it would be the same for all products	Note: because of a different harborage of clusters per product, this assumption has to be extended to these products (inactivation and growth shouldn't differ according to the product, because it would lead to a difference according to the cluster).
8	Data collected in Teunis(2022) are representative of Cluster 1 isolates in the U.S.	The risk multipliers can be applied to Cluster 1 serovars from Tenuis 2022, to adjust the DR models	In absence of similar studies focused solely in the US, Tenuis 2022 provides the best estimate available for a DR model relevant to Cluster 1 serovars [14]. The possibility of bias in the resulting DR models exists. However, no gold standard dose-response is available to test for such bias.

9	For a given strain, the DR model (probability of illness) follows a beta-Poisson DR	The beta-Poisson DR is applicable to *Salmonella* strains used in this study	FAO/WHO 2002 [35].
10	Strain variability impacts the mean $u = \alpha/(\alpha + \beta)$ of the underlying beta distribution in the beta-Poisson dose response DR model, but the parameter $v = \alpha + \beta$ is shared by all strains	These assumptions provide bounds in the estimation of beta-Poisson parameters	Assumption used in Teunis et al., 2010, and Thébault et al., 2013 [13,30].

References

1. Pouzou J, Fenske G, Pouillot R, Costard S, Taylor D, Zagmutt F. (2022, June 20-22) Classification of *Salmonella* serovars by genomic and epidemiological virulence traits. I3S2022 conference, St Malo, France.

2. Fenske G, Pouillot R, Pouzou J, Costard S, Taylor D, Zagmutt F. (2022, July 31- August 3) Identifying Sub-Populations in *Salmonella* Serovars from Genomic Virulence Markers. IAFP 2022. Pittsburgh Pennsylvania.

3. Fenske GJ, Pouzou JG, Pouillot R, Taylor DD, Costard S, Zagmutt FJ. The genomic and epidemiological virulence patterns of *Salmonella* enterica serovars. medRxiv; 2022. p. 2022.12.13.22283417. doi:10.1101/2022.12.13.22283417

4. Thiennimitr P, Winter SE, Winter MG, Xavier MN, Tolstikov V, Huseby DL, et al. Intestinal inflammation allows *Salmonella* to use ethanolamine to compete with the microbiota. Proceedings of the National Academy of Sciences of the United States of America. 2011;108: 17480–17485. doi:10.1073/pnas.1107857108

5. Ramachandran G, Panda A, Higginson EE, Ateh E, Lipsky MM, Sen S, et al. Virulence of invasive *Salmonella* Typhimurium ST313 in animal models of infection. PLOS Neglected Tropical Diseases. 2017;11: e0005697. doi:10.1371/journal.pntd.0005697

6. Jiang L, Wang P, Song X, Zhang H, Ma S, Wang J, et al. *Salmonella* Typhimurium reprograms macrophage metabolism via T3SS effector SopE2 to promote intracellular replication and virulence. Nat Commun. 2021;12: 879. doi:10.1038/s41467-021-21186-4

7. Cheng RA, Eade CR, Wiedmann M. Embracing Diversity: Differences in Virulence Mechanisms, Disease Severity, and Host Adaptations Contribute to the Success of Nontyphoidal *Salmonella* as a Foodborne Pathogen. Front Microbiol. 2019;10: 1368. doi:10.3389/fmicb.2019.01368

8. Brown E, Bell R, Zhang G, Timme R, Zheng J, Hammack T, et al. *Salmonella* Genomics in Public Health and Food Safety. EcoSal Plus. 2021;9. doi:10.1128/ecosalplus.ESP-0008-2020

9. Lerminiaux NA, MacKenzie KD, Cameron ADS. *Salmonella* Pathogenicity Island 1 (SPI-1): The Evolution and Stabilization of a Core Genomic Type Three Secretion System. Microorganisms. 2020;8: 576. doi:10.3390/microorganisms8040576

10. Carden SE, Walker GT, Honeycutt J, Lugo K, Pham T, Jacobson A, et al. Pseudogenization of the Secreted Effector Gene sseI Confers Rapid Systemic Dissemination of S. Typhimurium ST313 within Migratory Dendritic Cells. Cell Host Microbe. 2017;21: 182–194. doi:10.1016/j.chom.2017.01.009

11. Kuo C-H, Ochman H. The extinction dynamics of bacterial pseudogenes. PLoS Genet. 2010;6: e1001050. doi:10.1371/journal.pgen.1001050

12. Risk assessments of *Salmonella* in eggs and broiler chickens. [cited 13 Dec 2022]. Available: https://www.who.int/publications-detail-redirect/9291562293

13. Teunis PFM, Kasuga F, Fazil A, Ogden ID, Rotariu O, Strachan NJC. Dose–response modeling of *Salmonella* using outbreak data. International Journal of Food Microbiology. 2010;144: 243–249. doi:10.1016/j.ijfoodmicro.2010.09.026

14. Teunis PFM. Dose response for *Salmonella* Typhimurium and Enteritidis and other nontyphoid enteric salmonellae. Epidemics. 2022;41: 100653. doi:10.1016/j.epidem.2022.100653

15. Yoshida CE, Kruczkiewicz P, Laing CR, Lingohr EJ, Gannon VPJ, Nash JHE, et al. The *Salmonella* In Silico Typing Resource (SISTR): An Open Web-Accessible Tool for Rapidly Typing and Subtyping Draft *Salmonella* Genome Assemblies. PLOS ONE. 2016;11: e0147101. doi:10.1371/journal.pone.0147101

16. Liu B, Zheng D, Jin Q, Chen L, Yang J. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. Nucleic Acids Res. 2019;47: D687–D692. doi:10.1093/nar/gky1080

17. Olson R, Assaf R, Brettin T, Conrad N, Cucinell C, Davis J, et al. Introducing the Bacterial and Viral Bioinformatics Center (BV-BRC): a resource combining PATRIC, IRD and ViPR. Nucleic Acids Res. 2023;51:D678-D689. doi:10.1093/nar/gkac1003

18. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics. 2012;28: 3150–3152. doi:10.1093/bioinformatics/bts565

19. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30: 2068–2069. doi:10.1093/bioinformatics/btu153

20. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics. 2010;11: 119. doi:10.1186/1471-2105-11-119

21. Liaw A, Wiener M. Classification and Regression by randomForest. 2002;2.

22. Ward JH. Hierarchical Grouping to Optimize an Objective Function. Journal of the American Statistical Association. 1963;58: 236–244. doi:10.1080/01621459.1963.10500845

23. Murtagh F, Legendre P. Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? J Classif. 2014;31: 274–295. doi:10.1007/s00357-014-9161-z

24. Hennig C. Cluster-wise assessment of cluster stability. Computational Statistics & Data Analysis. 2007;52: 258–271. doi:10.1016/j.csda.2006.11.025

25. Batz MB, Richardson LC, Bazaco MC, Parker CC, Chirtel SJ, Cole D, et al. Recency-Weighted Statistical Modeling Approach to Attribute Illnesses Caused by 4 Pathogens to Food Sources Using Outbreak Data, United States - Volume 27, Number 1—January 2021 - Emerging Infectious Diseases journal - CDC. [cited 13 Dec 2022]. doi:10.3201/eid2701.203832

26. Scallan E, Hoekstra RM, Angulo FJ, Tauxe RV, Widdowson M-A, Roy SL, et al. Foodborne illness acquired in the United States--major pathogens. Emerg Infect Dis. 2011;17: 7–15. doi:10.3201/eid1701.p11101

27. Teunis PFM, Ogden ID, Strachan NJC. Hierarchical dose response of E. coli O157:H7 from human outbreaks incorporating heterogeneity in exposure. Epidemiol Infect. 2008;136: 761–770. doi:10.1017/S0950268807008771

28. de Valpine P, Turek D, Paciorek CJ, Anderson-Bergman C, Lang DT, Bodik R. Programming With Models: Writing Statistical Algorithms for General Model Structures With NIMBLE. Journal of Computational and Graphical Statistics. 2017;26: 403–413. doi:10.1080/10618600.2016.1172487

29. EBEL ED, WILLIAMS MS. When Are Qualitative Testing Results Sufficient To Predict a Reduction in Illnesses in a Microbiological Food Safety Risk Assessment? Journal of Food Protection. 2015;78: 1451–1460. doi:10.4315/0362-028X.JFP-15-042

30. Thebault A, Teunis PFM, Le Pendu J, Le Guyader FS, Denis J-B. Infectivity of GI and GII noroviruses established from oyster related outbreaks. Epidemics. 2013;5: 98–110. doi:10.1016/j.epidem.2012.12.004

31. McMillan EA, Wasilenko JL, Tagg KA, Chen JC, Simmons M, Gupta SK, et al. Carriage and Gene Content Variability of the pESI-Like Plasmid Associated with *Salmonella* Infantis Recently Established in United States Poultry Production. Genes (Basel). 2020;11: 1516. doi:10.3390/genes11121516

32. Aviv G, Tsyba K, Steck N, Salmon-Divon M, Cornelius A, Rahav G, et al. A unique megaplasmid contributes to stress tolerance and pathogenicity of an emergent *Salmonella* enterica serovar Infantis strain. Environmental Microbiology. 2014;16: 977–994. doi:10.1111/1462-2920.12351

33. Ebel ED, Williams MS, Cole D, Travis CC, Klontz KC, Golden NJ, et al. Comparing Characteristics of Sporadic and Outbreak-Associated Foodborne Illnesses, United States, 2004–2011 - Volume 22, Number 7—July 2016 - Emerging Infectious Diseases journal - CDC. [cited 16 Dec 2022]. doi:10.3201/eid2207.150833

34. Pettengill J, Carleton H, Tolar B, Lindsey R, Batz M, Bazaco M, et al. (2022, July 31- August 3) Predictive Analytics within Food Safety: Source Attribution of *Salmonella* Using Whole-Genome Sequence Data and Random Forest. IAFP; 2022. Pittsburgh, Pennsylvania.

35. Conducting the Dose–Response Assessment. Quantitative Microbial Risk Assessment. John Wiley & Sons, Ltd; 2014. pp. 267–321. doi:10.1002/9781118910030.ch8

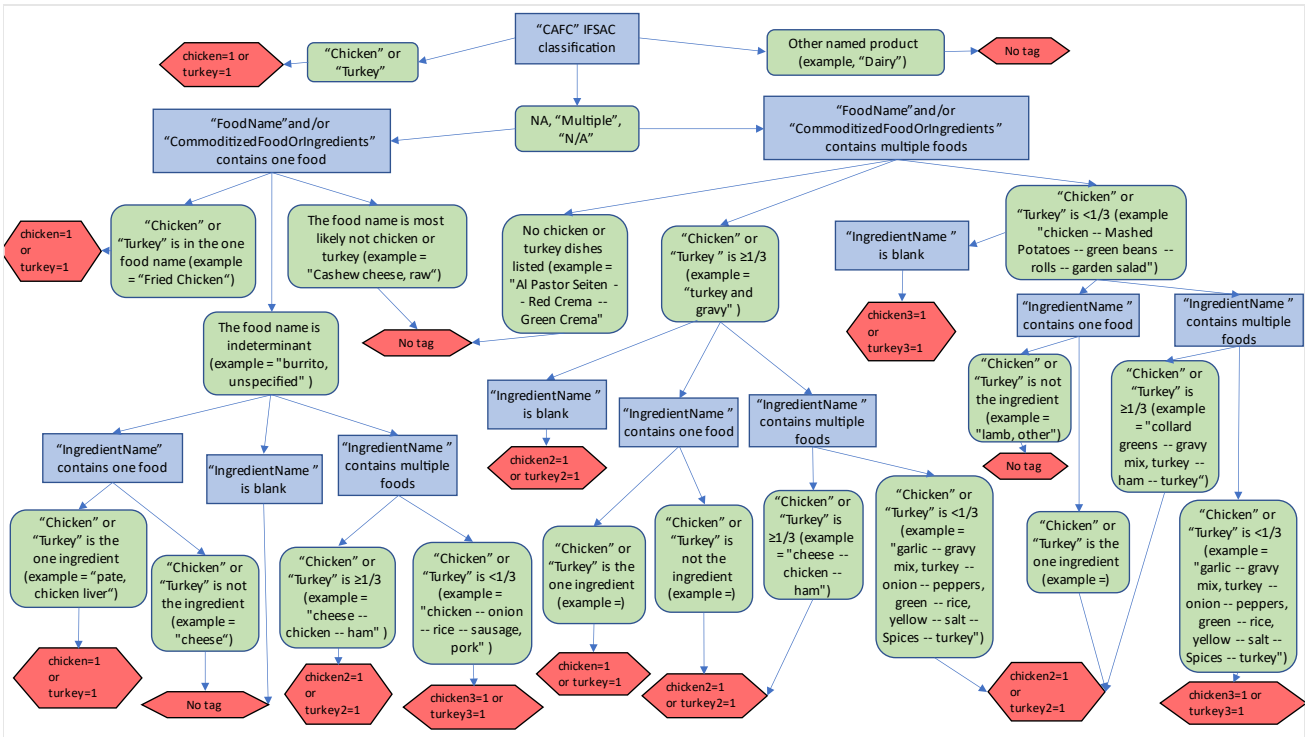
36. Clark CE. Letter to the Editor—The PERT Model for the Distribution of an Activity Time. Operations Research. 1962;10: 405–406. doi:10.1287/opre.10.3.405

Supplementary materials

This section contains more details on the methodology and computations used in our analysis. This section is not self-explanatory and should be reviewed in combination with the full report.

S1: Outbreak attribution

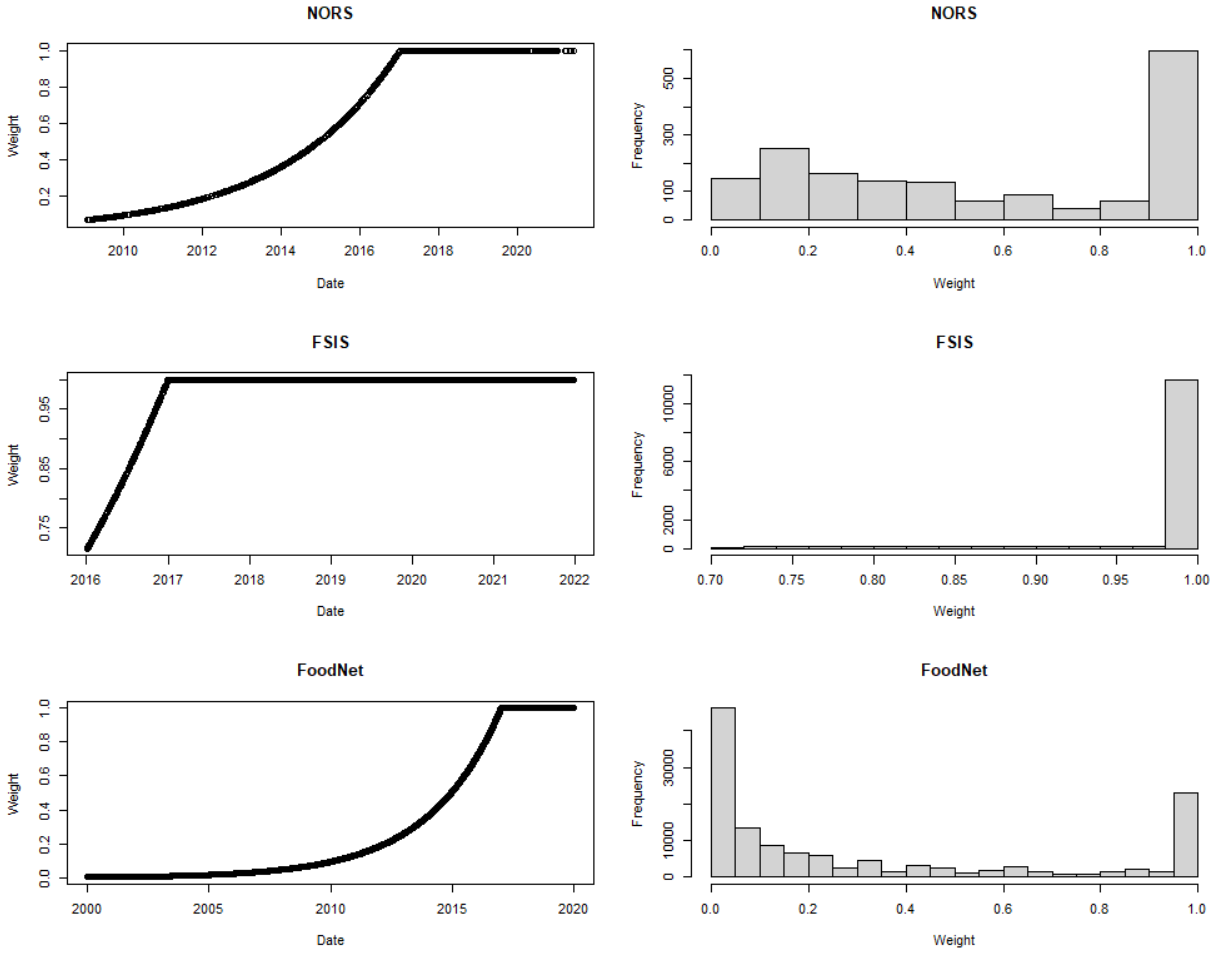
Chicken or turkey-attributed outbreaks were categorized as “definitive, “probable” or “possible” depending on the following NORS dataset fields: “CAFC”, “FoodName”, “CommoditizedFoodOrIngredient”, and “IngredientName. Supplemental figure 1 shows the flow diagram used to attribute an outbreak to poultry and assign a level of certainty.



Supplemental figure 1: Criteria for NORS attribution to poultry. Chicken/turkey= definite attribution, chicken2/turkey2= probable attribution and chicken3/turkey3=possible attribution

S2: Recency Weight

The equation [25] is written $w(x) = \begin{cases} (5/7)^{\left(\frac{maxDate-x}{365}\right)-5} & \text{if } \left(\frac{maxDate-x}{365}\right)-5 > 0 \\ 1 & \text{if } \left(\frac{maxDate-x}{365}\right)-5 \leq 0 \end{cases}$, where *maxDate* is the number of days between a reference day and 12/31/2021, and *x* is the number of days between this reference day and the date considered. Supplemental figure 2 illustrates the evolution of the weight with the date in the NORS, the FSIS and the FoodNet data.



Supplemental figure 2: left: weight as a function of the date, right: histogram of weights for the NORS (top), FSIS (middle) and FoodNet (bottom) data.

S3: Modeling outbreak cases

As cases per outbreak can vary widely, we adapted the method used by Batz et al., 2021 [25] to provide a more robust prediction of outbreak cases (i.e., smoothing extreme values). We used the following linear mixed model:

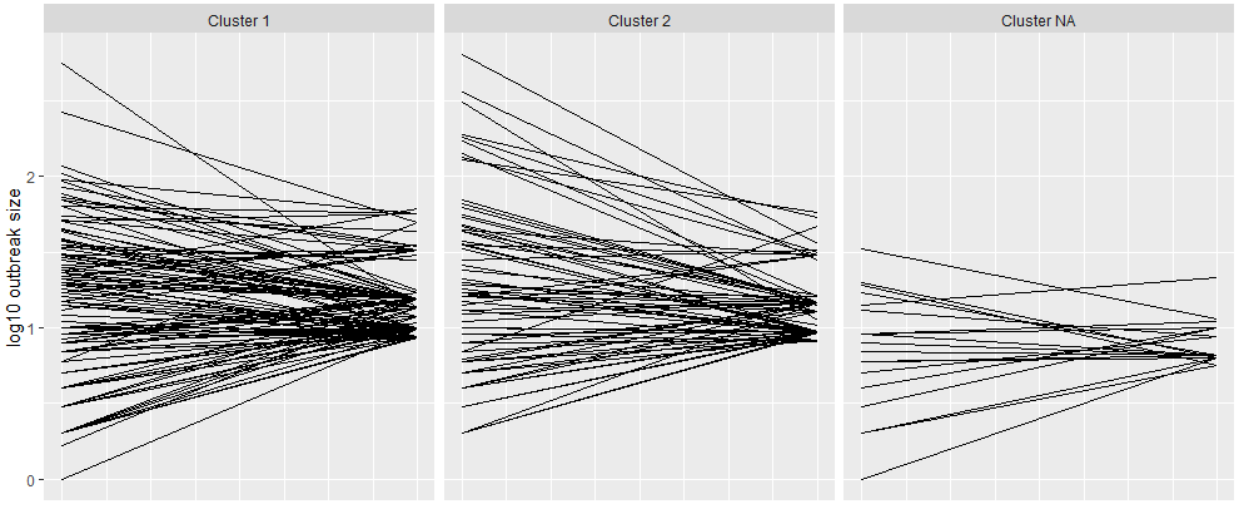
$$\log(Y_{i,j}) = \mu + FC + TP + MS + \zeta_i + \varepsilon_{ij}$$

with :

- *Y* the estimated number of Primary Cases (eventually weighted for sub-outbreaks) associated to cluster *i*; Note that we are interested in primary cases (rather than total cases, as in Batz et al. (2021)[25] ;
- *FC*: the food category (fixed effect). We used 17 IFSAC categories, including "multiple" and "NA"s;
- *TP*: the type of preparation (fixed effect) using 5 categories as described in Batz et al, 2011;

- *MS*: multistate (Binary);
- $\zeta_i \sim N(0, \sigma_1^2)$ is a random effect associated to the cluster *i*;
- $\varepsilon_{ij} \sim N(0, \sigma^2)$ is the error, independently.

Supplemental figure 3 illustrates how the model shrinks extreme values to more central ones.



Supplemental figure 3: number of cases (log_{10}) before (left) and after (right) use of the model, according to the Cluster (example: 3 clusters. Cluster 3 is not represented as no NORS outbreak was assigned to this cluster).

S4: Underdiagnosed cases according to severity

In order to consider the differential underdiagnosis according to case severity, we used the method developed by Scallan et al. (2011)[26]. Following this method, we assume that the sensitivity of laboratory test followed a Pert (min=0.6, mode=0.7, max=0.9)[36]; that the proportion of clinical laboratories routinely testing stool samples for *Salmonella* followed a Pert(0.94, 0.97, 1); that the proportion of respondents who submitted a stool specimen among persons with bloody diarrhea followed a Pert(0.11, 0.36, 0.62); that this proportion among persons without bloody diarrhea followed a Pert(0.12, 0.19, 0.25); that the proportion of individual who sought medical care among persons with bloody diarrhea followed a Pert(0.19, 0.35, 0.51); and those without bloody diarrhea followed a Pert(0.15, 0.18, 0.20). We also apply this differentiated underreporting factor (on average: 1 case out of 13 for bloody diarrhea vs. 1 case out of 44 for non-bloody diarrhea) to the various clusters. For the proportion of bloody diarrhea, we use, as an uncertainty distribution, a Beta distribution under the assumption of a prior proportion of bloody diarrhea equal to a Beta(0.5, 0.5) (i.e., Jeffrey's prior).

S5: Bayesian Inference model for Cluster 1

The model is written as following. For each outbreak *o*:

$$p_o = 1 - {}_2F_1(\alpha_o, r_o, \alpha_o + \beta_o, -d_o / r_o)$$

$$x_o \sim \text{binomial}(\text{size}=f n_o, \text{prob}=p_o)$$

with r_o, d_o, n_o and x_o, the data provided in appendix, and

$\alpha_o = u_o \times v_o$

$\beta_o = (1-u_o) \times v_o$

$u_o = \exp(\omega_o) /(1+\exp(\omega_o))$

$v_o = \exp(\zeta_o)$

ω_o ~ normal(mean= ωm, sd= ωsd)

ζ_o ~ normal(mean= ζm, sd= ζsd)

The prior distributions were flat, centered around the values estimated from Teunis et al (2010):

ωm ~ normal(mean= -5.9, sd=8)

ζm ~ normal(mean= 1.15, sd=8)

ωsd ~ uniform(min=0, max=4)

ζsd ~ uniform(min=0, max=4)

In Teunis (2022), one outbreak appears to have a right-censored dose (“>2.4E5” cfu, and one has a left-censored dose (‘"<3.60E3" cfu). We adapted our model to deal with these censored values (considering a flat lognormal(11, 10) prior distribution for the doses).

Appendix B Data and Data Analysis

Data Sources

Data used throughout the risk assessment are summarized below. Prevalence estimates were based on the most recent calendar year (i.e., 2021) of data available at time of analysis for parts and comminuted chicken products. All carcass estimates relied on the 2022 young chicken carcass exploratory sampling program. Available *Salmonella* serotype data from January 1, 2016 through December 31, 2021 was used by EpiX Analytics to develop a dose-response model as outlined in **Appendix A**

1998-2019 FSIS Chicken Carcass Verification Program

Data spanning the entirety of chicken carcass verification testing was used to assess how the mixture of serotypes has changed over time. The analysis considers only samples that tested positive for *Salmonella* and were successfully serotyped, of which there were 14,928. The primary use of the data is to illustrate the change in Infantis over time. This trend is compared to the trend in human cases attributed to Infantis between 1996 and 2019. These results are used to motivate the choice of the number of serotype clusters used in the risk assessment model.

2022 Young Chicken Carcass Exploratory Sampling Program

FSIS began its exploratory data collection program in late April of 2022 and samples collected up to the end of October 2022 are included in this analysis. This dataset is based on FSIS' ongoing performance standard sampling program, but it adds sample collection at the rehang location of the slaughter process. Additional laboratory testing was performed to quantify indicator organism (AC and Enterobacteriaceae) at both locations. Enumeration of *Salmonella* via quantitative PCR was performed, with the first enumerated samples being collected in early August. These data will primarily be used to characterize establishments as they relate to observed changes in indicator organisms.

At the time of this analysis, 5,425 samples had been collected from 206 slaughter establishments. Of these samples, a total of 4,517 samples from 200 establishments had AC level data at both the rehang and post-chill locations. The subset of data with enumerated AC data at both locations was used to assess relationships between multiple explanatory variables and the occurrence of *Salmonella.* Of the total samples, 2,823 samples screened positive at rehang and 216 screened positive at post-chill. Level data was available for 214 collected at rehang and 14 collected at post-chill.

2007-2008 Young Chicken Microbiological Baseline Study

The sampling frame for the young chicken study contains establishments that slaughter 99.99% of the total head of young chickens slaughtered in the U.S. under Federal Inspection during fiscal year 2006 (FSIS, 2009a). FSIS personnel collected samples at 182 broiler chicken slaughter facilities from July 2007 and June 2008. There were a total of 6550 samples, with an equal number (3275) of samples collected at pre-evisceration and post-chill (FSIS, 2009a). For each sampling event at an establishment, the inspector randomly selected one pre-evisceration carcass and one post-chill carcass from the same grow-out flock.

Samples were collected using a 400 mL rinse sampling method with a 1 min agitation time. For the *Salmonella*, each sample was first tested using a qualitative test with a theoretical LOD of 0.03 cfu/ mL of rinse fluid. The screen test for *Salmonella* uses the DuPont Qualicon BAX system.

For the samples that tested positive in the screening test for *Salmonella*, the levels of *Salmonella* were estimated using a MPN experiment (Cochran, 1950; Haas, 1989) using a 3-tube, 3-dilution experiment and dilution volumes of 10, 1 and 0.1 mL of rinsate in each tube (LOQ ¼ 0.03 cfu/mL).

For the AC and GEC level, 1 mL of rinsate was added to 9.0 mL of a diluent blank (10^{-1}) and vortexed. Serial dilutions from 10^{-1} to 10^{-4} were plated onto Petrifilm and enumerated following incubation.

2012 Chicken Parts Microbiological Baseline

The chicken parts microbiological baseline study is the only time that FSIS has estimated *Salmonella* levels in parts samples. Data describing microbial contamination of chicken parts were generated by an FSIS microbiological baseline survey (FSIS, 2013). Samples were collected from January through August of 2012. Although this survey collected samples from many different types of chicken parts, only the data related to breast, wing and leg sampling were used here. We chose these parts because they had similar levels of *Salmonella* and *Campylobacter* contamination and constituted about 90% of the chicken parts produced in the U.S. (FSIS, 2015). In this microbiological baseline study, each sample consisted of 4 lbs. of a particular part. We assume testing for the parts performance standard will either consist of 4 lb. samples comprising a single part or a mixture of parts. In the latter case, it is desirable that pooling of different types of parts should be among parts with similar pathogen occurrence. This is the case for wing, breast and leg samples as opposed to, for example, neck or giblets samples.

There were 384 establishments in the microbiological baseline survey with at least one breast, wing, or leg sample. An average of 4.4 such samples per establishment (range 1 – 13) were collected. For the purposes of weighting the establishment sampling data, the annual production volume (in pounds) for each sampled establishment was captured from FSIS' records. The level in each *Salmonella*-positive sample was estimated using a 3-tube, 3-dilution MPN, with tube volumes of 10, 1 and 0.1 mL.

The chicken parts dataset consisted of 1,681 samples of which 458 were positive on the *Salmonella* screening test.

2016-2021 FSIS Chicken Parts Verification Program

As part of its ongoing *Salmonella* performance standards, FSIS has collected samples of chicken parts since late March of 2015. Given changes to sampling and laboratory methods during the period of this study (Williams, 2022), data from January 1, 2016 through December 31, 2021 is considered here.

Samples were collected under a continuous sampling approach where FSIS inspection personnel in regulated establishments were tasked with collecting samples at regular intervals throughout the year. The collection rate is stratified by production volume, with the largest establishments being assigned weekly sample collection tasks. Lower volume establishments are assigned bi-weekly or monthly sample collection tasks and the performance standard is adjusted to slightly reduce the probability of incorrectly classifying an establishment as failing (Williams, 2022).

Sample collection within the establishment occurred after all interventions are applied. These interventions often consist of the application of organic acids to inactivate bacteria, with the compound and method of application varying across the industry (Rasekh, 2005). For each sample, 1.8 kg (4 lbs) of either chicken breasts, legs or wings were aseptically placed into a sterile bag with 400 mL of buffered peptone water and shaken for one minute. The rinsate was shipped to one of three FSIS laboratories by an overnight delivery service FSIS laboratory methods used for *Salmonella* detection are available online (FSIS, 1998).

Data from January 1, 2021 through December 31, 2021 are used in microbial contamination estimates throughout this risk assessment. The parts dataset consisted of 14,192 samples of which 1081 were positive on the *Salmonella* screening test.

Data from January 1, 2016 through December 31, 2021 was shared with the FSIS Cooperator EpiX Analytics and used in their development of a dose-response model.

2016-2021 FSIS Comminuted Chicken Verification Program

As part of its ongoing *Salmonella* performance standards, FSIS has collected samples of comminuted chicken since 2015, with consistent collection across all establishments commencing in 2016. Each sample used in this study was subjected to a *Salmonella* test and was based on a 325 g aliquot of comminuted poultry. Each sample underwent enrichment and incubation prior to testing for the presence of *Salmonella*. The screening test was a BAX PCR test for the declaration of presumptive positive samples. All presumptive positive samples underwent further testing to confirm the presence of viable *Salmonella*.

A subsample of these samples were enumerated using a 3-tube, 5-dilution the Most Probable Number (MPN) method analysis (Cochran, 1950). The MPN analysis was performed on all samples at one of the FSIS three laboratories. These samples are assumed to be a random subsample of comminuted chicken sample because the location of sample collection does not determine the laboratory at which the samples are analyzed (i.e., the overnight courier service ships all samples to a central location in Memphis, Tennessee and then ships the samples to one of the three laboratories. Thus, there is no advantage to having to a laboratory serve a specific geographic region of the country).

MPN analysis was not consistently performed during 2020 and no MPN analyses were performed after 2020.

The comminuted chicken dataset consisted of 1,815 samples of which 387 were positive on the screening test. Given the lower priority of MPN analysis, 21 *Salmonella*-positive samples were not analyzed due to limited staffing in the laboratory. These missing results were addressed by using imputation by simple random sampling ((van Buuren, 2011)) from samples where the MPN analysis was performed.

Production Volume

In addition to the testing data, FSIS maintains production volume estimates for each establishment. These estimates are derived by assigning each establishment to a daily production volume class and multiplying the production volume represented by each class by the monthly days of production in the

establishment. These volumes are regularly updated to reflect changes in production volume in each establishment.

Data Analysis

Accuracy of Quantitative PCR Enumeration Methods

Since the inception of the original PR/HACCP prevalence-based performance standards in the mid-1990s (FSIS, 1996b), FSIS has set performance standards such that establishments that are truly passing the standards are not misclassified as failing (i.e., a low probability of a false positive result). For these performance standards, FSIS ensured that each sample that was positive on the screen test truly contained viable *Salmonella* using an extensive confirmatory analysis process (FSIS, 2022b), so the chance of a false positive results was theoretically zero. Given the nearly perfect specificity of the test, the probability of incorrectly classifying an establishment was determined by assessing the binomial probability of observing more than the allowable number of positive samples when the establishment's underlying prevalence was below the target threshold.

For example, all the original FSIS performance standards were chosen so that an establishment whose prevalence was at the maximum allowable prevalence for each product had at least an 80% chance of passing the standard (i.e., a specificity of at least 0.8). While the performance standards for parts and carcasses were not developed to achieve a given specificity, the estimated specificity ranges from 0.82 to 0.92 for high-volume establishments that receive 52 samples per year (Ebel, 2020). The performance standards were made less stringent to improve the specificity of low-volume establishments that are sampled less frequently (FSIS, 2016).

If new standards are to be implemented where the condition of passing a standard is dependent on the level of *Salmonella* observed in a sample, it is necessary to assess the quantification assay's ability to accurately determine sample levels. Different scenarios apply to post-chill and rehang samples, as described below. A description of how FSIS has historically enumerated *Salmonella* is presented below, as well as a description of the qPCR enumeration method used by FSIS at the time of this analysis.

How FSIS enumerates Salmonella
Methods based on cell cultures have historically been used to determine the level of microorganisms in a sample (i.e., for bacteria, cells per unit of volume or weight). The population is enumerated or estimated by either using plate count methods or MPN. MPN is utilized when organisms may be at lower levels (i.e., <10 cfu/g or mL) or when an enrichment is needed to address resuscitation, possibly from injury. Bacterial injury can occur in response to processing stresses (i.e., heat, sanitizers, storage conditions, etc.) (Petran, 2015).

MPN analysis requires multiple test portions to be analyzed from the same sample, requiring significant resources of time and supplies (media and laboratory disposables) and associated costs. Time-to-result is dependent upon the detection system used for determining the presence of *Salmonella*.

In recent years, methods have been developed using quantitative polymerase chain reaction (qPCR) for enumerating *Salmonella* n poultry. The cycle threshold (CT) value, which is inversely related to the amount of *Salmonella* in the sample, is used to quantify the number of colony-forming units (cfu). Many of these methods require that the sample is enriched in specified media for a prescribed time and temperature (NACMCF, 2023). One method has been developed without enrichment that uses

centrifugation of the sample to concentrate cells (bioMérieux Gene-Up Quant *Salmonella*) to provide test results more rapidly than traditional enumeration methods. Two of these methods, Hygiena's SalQuant™ and bioMérieux's GENE-UP® QUANT *Salmonella*, have been validated by the Association of Official Agricultural Chemists (AOAC) Performance Tested Method certification for level of *Salmonella* in 2021 and 2022 (NACMCF, 2023).

In July 2022, FSIS awarded a contract to bioMérieux to incorporate its non-enrichment quantification system for *Salmonella*, 'GENE-UP™ QUANT *Salmonella*,' into the agency's laboratory system. FSIS evaluated commercially available quantification systems and determined that this technology is the most appropriate for use in the high throughput FSIS laboratory environment.

Post-chill analysis

The concept of assessing an instrument's ability to correctly classify contaminated samples is essentially the epidemiological equivalent of determining the assay's positive predictive value (PPV) (Altman, 1994), which considers the sensitivity and specificity of the assay as it relates to the level distribution. The relationships that define PPV for this application are

$$PPV = \frac{\text{the fraction of true positive samples above the LOQ}}{\text{the fraction of positive samples declared above the LOQ}} = \frac{\text{true positives}}{\text{true positves+false positives}}.$$

The fraction of false positives is estimated first.

The bioMérieux's GENE-UP® QUANT *Salmonella* limit of quantitation (LOQ) is 10cfu/mL, but the maximum level observed in both the previous FSIS carcass and parts microbiological baseline studies was 11 cfu/mL (FSIS, 2009b; FSIS, 2013). As there has been additional reduction in AC of roughly 1 log between the previous microbiological baseline studies and current conditions, it is logical to assume that current MPN level estimates would likely be lower than what was observed in the microbiological baselines. Consequently, most screen-test positive samples are expected to have levels between the 0.033 cfu/mL LOD for the screen test and the instrument's LOQ.

The certification process for the bioMérieux's GENE-UP® QUANT instrument and related documentation does not specifically address the accuracy of the binary decision to classify a sample as being above or below a specific level that is near the instrument's LOQ. Thus, it is necessary to assess the probability that samples below the LOQ generate estimates that exceed the LOQ and if so, by how much. This value can be compared to the portion of samples that are expected to truly exceed the LOQ.

While the probability of a sample testing positive at any given level is not known, the documentation for the AOAC certification of the instrument provides data that can be used to estimate the probability of the level in the sample being high enough to assign a level value of >10 cfu/mL (i.e., samples whose crossing point values are short enough to be declared positive and will be enumerated based on that time (Joelsson, 2022). The testing data describes the number of positive tests out of either 5 or 20 trials across 3 different food matrices (pea powder, protein powder and cookie dough). Experiments with 20 trials were carried out at what was classified as low levels (0.45 to 0.9 cfu/g), while the experiments with 5 trials were carried out on samples classified as high levels (0.75 to 4.43 cfu/g). A logistic regression model, denoted by $P(+ \mid y)$ was fitted to the data to estimate the probability of a positive sample as a function of log10 level (denoted *y*)(**Figure 47**). Note that the instruments LOQ is determined by choosing a level when some reasonably high fraction of the samples (i.e., > 70-80%) would be declared positive.

The other component needed to assess the instrument's ability to correctly classify samples with low levels is the distribution of sample levels. The level distributions for each commodity can be used, but because only samples that test positive on the screening test are subjected to quantification, only the portion of the distribution that falls between the LOD and LOQ are considered. This distribution will be referred to as the normalized contamination distribution, which is depicted by the dashed red line in **Figure 47**, and will be denoted by

$$\tilde{f}(y) = f(y) / \int_{\log10(1/30)}^{\log10(10)} f(y)\,dy..$$

The probability of incorrectly classifying these low level samples as samples with levels above 10 cfu/mL is determined by finding the expected value of $P(+ \mid y)$.

The probability of interest is determined by evaluating

$$P(\hat{y} > LOQ \mid LOD < y < LOQ) = \int_{\log10(1/30)}^{\log10(10)} P(+ \mid y)\tilde{f}(y)\,dy,$$

where $\hat{y}$ is the qPCR-estimated level.

r carcasses, parts, and comminuted product, the probability of incorrectly classifying the samples that fall between the LOD and LOQ of the assay is 0.0468, 0.0232, and 0.0408, respectively. These values represent the estimated proportion of false positives.

A similar mathematical integration argument can be made to determine the fraction of true positive samples. Nevertheless, a simpler approximation exists that assumes perfect sensitively to the assay, which in this case is

$$P(\hat{y} > LOQ > LOQ) \cong 1 - F\left(LOQ, \hat{\mu}, \hat{\sigma}\right).$$

This approximation overestimates the proportions of true positives and has values of 0.00034, 4.5×10^{-5}, 0.0079 for carcasses, parts, and comminuted samples, respectively.

Combining these results leads to the approximation

$$\tilde{PPV} = \frac{1 - F\left(LOQ, \hat{\mu}, \hat{\sigma}\right)}{P(\hat{y} > LOQ \mid LOD < y < LOQ) + 1 - F\left(LOQ, \hat{\mu}, \hat{\sigma}\right)},$$

Where $\tilde{PPV}$ is an overestimate the PPV, as defined above. These estimated PPV values are 0.0073, 0.0020, and 0.1628. The values of PPV range from 0 to 1, with desirable values of PPV being near 1. These values for carcasses and parts are low enough to be unacceptable in any epidemiological of medical application.

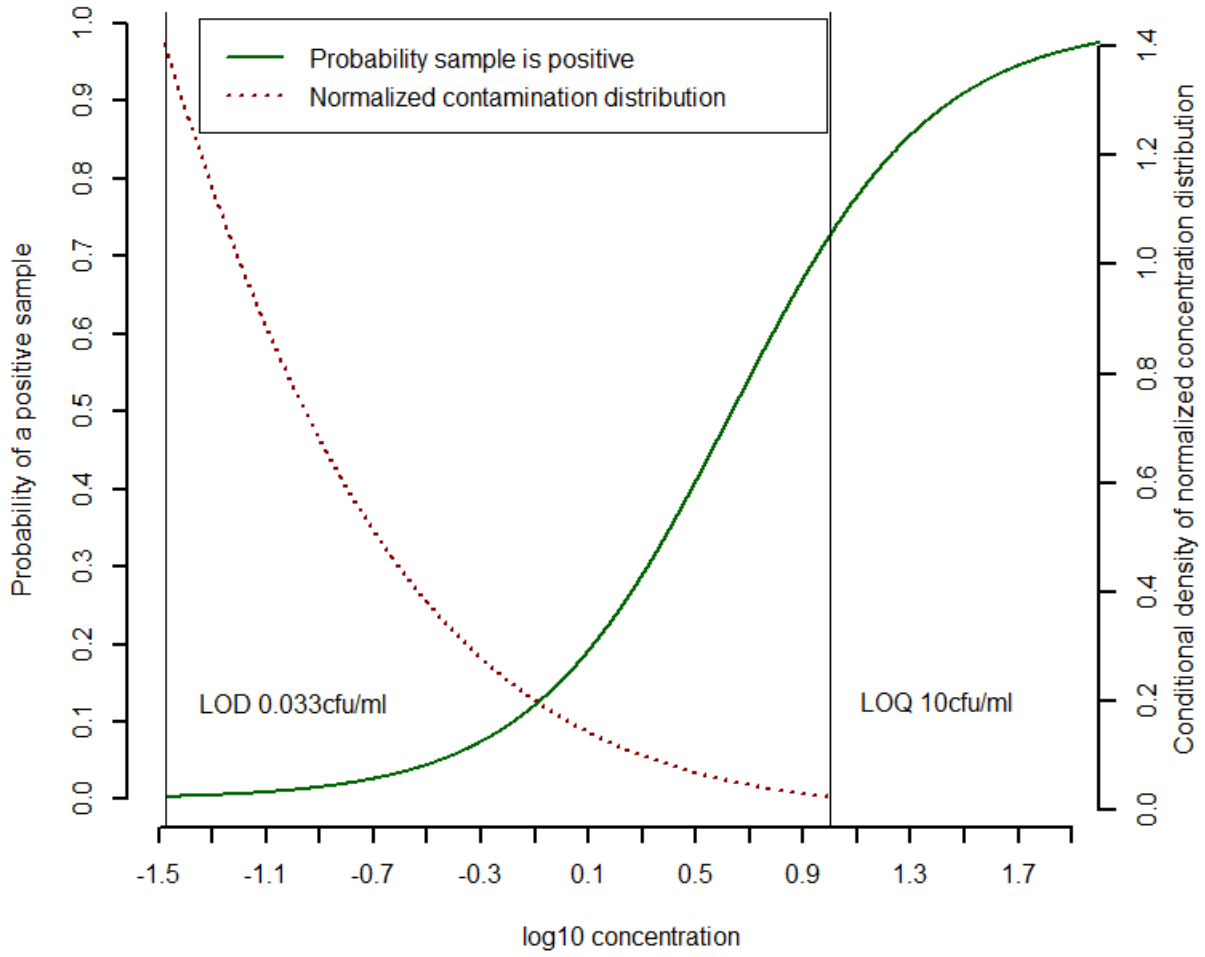


Figure 47: Overlap of positive sample probability and current normalized contamination distribution for *Salmonella* in chicken carcasses.

Effect of Misclassification at Post-Chill

To appreciate the effect of misclassification, these values can be compared to the fraction of samples that are expected to exceed the LOQ and can be assumed to have level values that are above the LOQ.

This ratio is given by

$$P(\hat{y} > LOQ \mid y < LOQ) / \left(1 - F\left(LOQ, \hat{\mu}, \hat{\sigma}\right)\right).$$

This ratio suggests that carcass and parts samples whose estimated levels are above the LOQ, but whose actual level lies between the LOD and LOQ, are 135 and 505 times more likely to occur than samples whose level is truly above the LOQ. Comminuted chicken, which has a larger fraction of samples with higher levels, has a ratio of 5.1. Nevertheless, the ratio for an assay that performed well in this

application would be around 0.2 (i.e., 4 out of 5 sample could be reasonably assumed to indicate that the true level was above the LOQ).

The performance of the current qPCR technologies in this application is due to a confluence of the following factors:

- An LOD near the maximum levels expected in the population.
- The much higher proportion of samples whose true level lies between the LOD and LOQ than the fraction of samples above the LOQ.
- The non-negligible probability of samples with low levels being declared test-positive (i.e., samples whose crossing point values are short enough to be declared positive so their crossing point value will then be used for levels).

These issues can be further be summarized by noting that an assay can only have a high PPV when the majority of the positive samples have levels above the LOQ of the assay (Hazra, 2017). Of the poultry commodities where FSIS has collected samples of final product, only *Campylobacter* samples collected in the mid-1990s, which was prior to the implementation of the PR/HACCP rule, met these conditions (FSIS, 1996a; FSIS, 1997a; Williams, 2021). The effect of high levels of contamination on the estimated PPV can be understood by contrasting the level distribution and probability of detection for this product pathogen pair (**Figure 48**) to the *Salmonella* in chicken carcasses graph (**Figure 48**). In this comparison, the dashed line for the normalized level distribution line is increasing, rather than decreasing, across the range of level values.

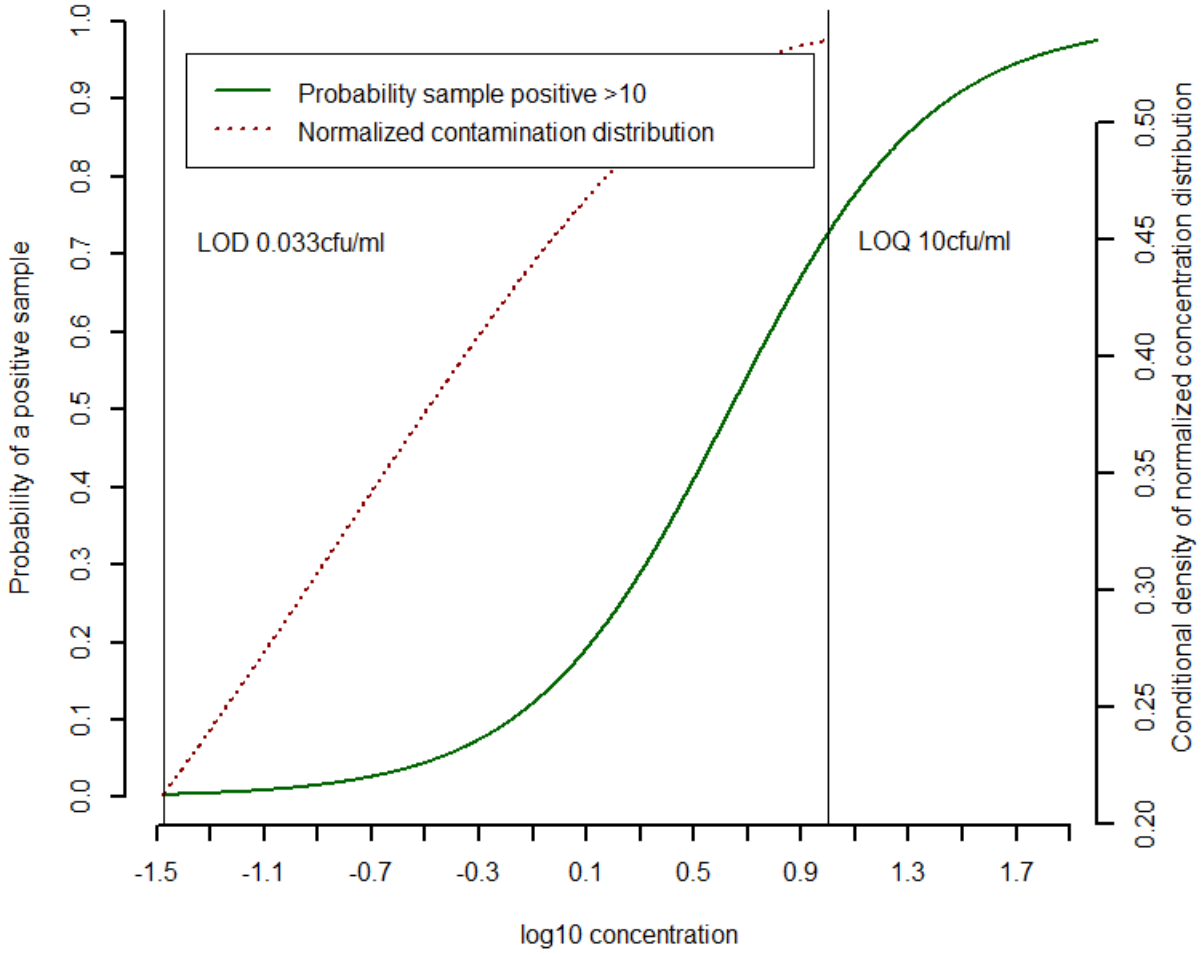


Figure 48: Overlap of positive sample probability and 1990s *Campylobacter* in chicken carcasses normalized contamination distribution.

In summary, while the analyses presented here are limited, they indicate that nearly all final product carcass and parts samples that are assigned enumerated values above the LOQ of 10 cfu/mL using qPCR would be incorrectly classified. This is a challenge to achieving meaningful public health benefits using a performance standard that determines the acceptability of lots based qPCR enumeration. Furthermore, findings from current and future risk assessments that evaluate the public health impact of *Salmonella* level using data from any assay with a low PPV should be evaluated before use.

Rehang Analysis

Figure 49 (adapted from (Chaney, 2022)) depicts the precision of a typical qPCR assay. The horizontal green lines across the graph intersect the clusters of points that coincide with aliquots with known and differing levels. Note that the upper horizontal line intersects the cloud of points representing both samples containing 1 and 2 logs of *Salmonella*. Similarly, the lower green line intersects the clusters for

samples containing either 3 or 4 logs of *Salmonella.* The red line separating clusters of points with either 1 or 2 logs and the clusters for samples with 3 or 4 logs indicates that the instrument has reasonably good discriminatory power to differentiate between samples <2 and >2 logs. Nevertheless, about 99% of rehang samples are expected to have levels < 2 logs ($F(2, \hat{\mu}_{rehang}, \hat{\sigma}_{rehang}) = 0.988$)

This phenomenon indicates that qPCR technology has poor discriminatory power for samples with levels that differ by up to 1 log. The benefit of qPCR technologies is that the enrichment step needed for assays such as MPN and counting colonies on petri film at various dilutions is not necessary. However, as was presented above, challenges remain regarding the precision of the estimates.

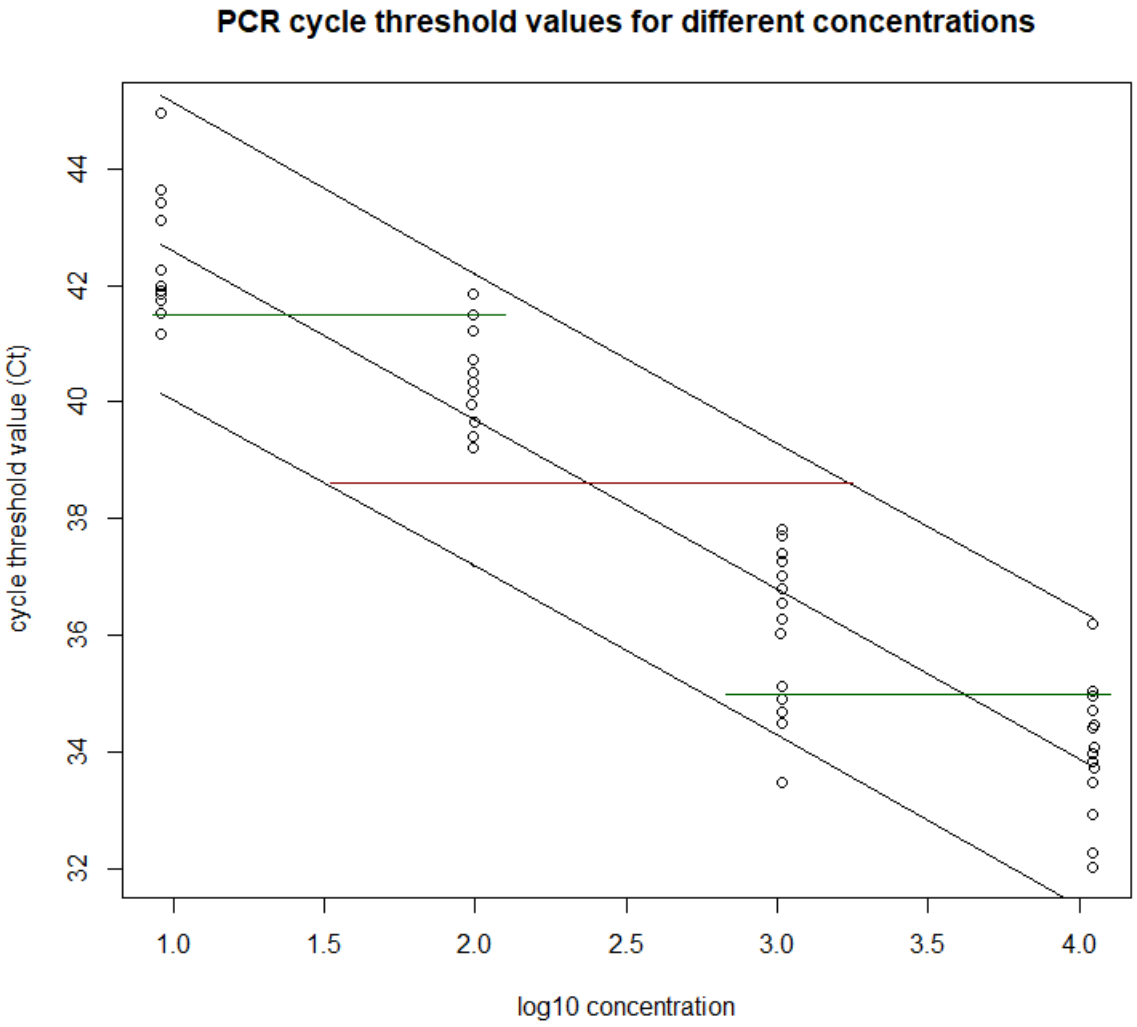


Figure 49: Precision of a typical qPCR assay adapted from (Chaney, 2022).

Flock Size

Many of the risk management options consider the effects of specific actions taken on individual flocks

or daily production lots of parts or comminuted product. FSIS assigns a daily production volume category and an estimate of the number of days of production per month for the parts and comminuted establishments. The daily production volume for these product types is used to estimate the volume of product in each establishment that would be affected by the risk management options that require removal or diversion of a production lot.

FSIS records the daily broiler slaughtered in every establishment and it also maintains records on the number of broiler chickens per flock. While total slaughter numbers are routinely monitored and are accurate, the recorded flock information has occasional incorrect values. For example, there are occasional records where the reported flock size exceeds the establishment's weekly production volume.

While it is not possible to correct the data or remove erroneous entries, the mean flock size, and its variability, were estimated using the following logic. We assume that the majority of entries for an establishment are accurate and developed a list of all flocks and their size. The influence of outliers is mitigated by determining the median reported flock size as well as the flock sizes representing the 40^{th} and 60^{th} quantiles of the distribution. A log10 transform was applied to these values and the parameters of a normal distribution are estimated using the percentiles match with the quantiles using the rriskDistributions (Belgorodski, 2016) library in R.

The estimated annual number of flocks processed each year is 240,311. Flock size is strongly correlated with an establishment's annual production volume (**Figure 50**), with high-volume establishments processing flocks of that range from approximated 16,000 to 79,000 (4.2 to 4.9 on the log10 scale). Establishments producing less than 1 million broilers per year have average flock sizes ranging from roughly 100 to no larger than approximated 3500 broilers. The average number of flocks processed per day shows similar patterns, with the highest volume establishments having multiple flocks and establishments producing less than one million broilers per year generally only processing one flock per day (**Figure 51**).

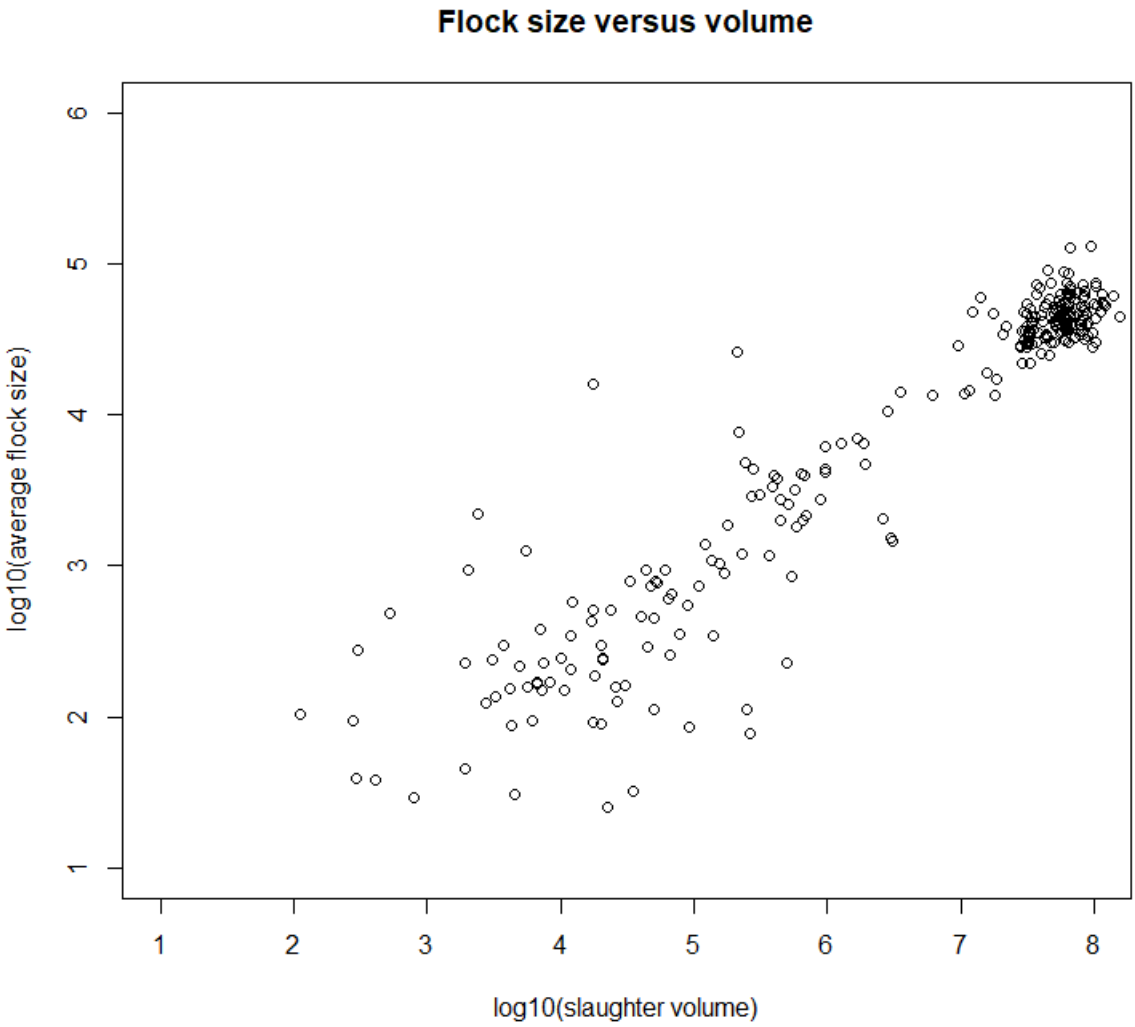


Figure 50: Average flock size as it relates to total annual slaughter volume. Flock size clearly increases as a function of annual production.

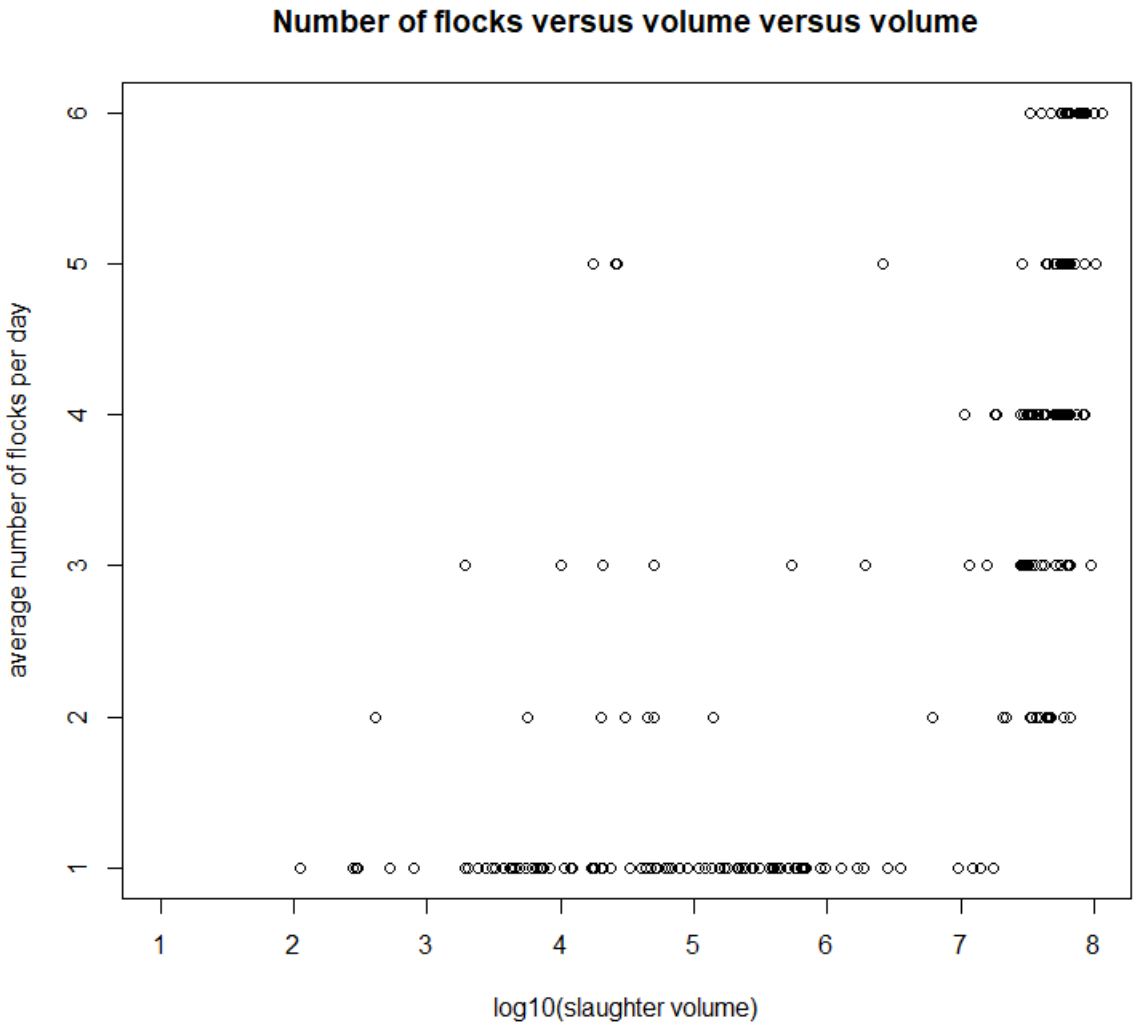


Figure 51: Average number of flocks processed per day as a function of annual production volume. With few exceptions, establishments producing less than 1 million broilers per year process single flocks.

NHANES Chicken Consumption Data Analysis

Data on the consumption of chicken in the U.S. were obtained from the National Health and Nutrition Examination Survey (NHANES). The NHANES program suspended field operations in March 2020 due to the 2019 coronavirus disease (COVID-19) pandemic. As a result, data collection for the NHANES 2019-2020 cycle was not completed and the collected data are not nationally representative. Therefore, data collected from 2019 to March 2020 were combined with data from the NHANES 2017-2018 cycle to form a nationally representative sample of NHANES 2017-March 2020 pre-pandemic data.

All NHANES participants are eligible for two 24-hour dietary recall interviews. The first dietary recall interview is collected in-person in the Mobile Examination Center and the second interview is collected by telephone 3 to 10 days later. In the 2017-March 2020 pre-pandemic sample, 12,634 participants provided complete dietary intakes for Day 1. Of those providing the Day 1 data, 10,830 provided

complete dietary intakes for Day 2. The NHANES Day 1 weights, adjusted for non-response and daily variability and the SAS code examples given on the NHANES website were used in the analysis (CDC, 2022b).

The full results of this analysis are below.

Results

Table 47 has the percent of the U.S. population consuming chicken commodities on an average day according to NHANES day 1 records. **Table 48** has the servings per day and percent of the U.S. population consuming chicken commodities on an average day according to NHANES day 1 records. According to **Table 47**, about 44.8% of the population consumes chicken of which 28% consists of chicken parts with 15.5% comminuted chicken and with the remainder of ground chicken. Or alternatively and arguably more accurate is the percentage of 44.8% chicken broken down to 28% parts and 16.8% combined ground and comminuted chicken.

Table 47: Percent of U.S. population consuming chicken commodities per day.

Commodity	Consumers	SEM	Percent
CHICKEN	144,335,194	9,791,077	44.8
PARTS	90,103,575	5,710,953	28.0
GROUND	4,275,885	749,686	1.3
COMM	49,955,734	3,332,843	15.5
GC=GROUND+COMM	54,231,619	3,667,564	16.8
PARTS+GROUND+COMM	144,335,194	9,791,077	44.8
PARTS+GC	144,335,194	9,791,077	44.8
US POPULATION	322,281,961		100.0

According to **Table 48**, about 52.6% of servings consumption population or 100% of servings of chicken is equal to 169,585,818 servings per day of which about a third consists of chicken parts with an average of 107,074,562 servings per day with 63.1% of servings as comminuted chicken at 57,964,750 servings per day 34.2% of servings and with the remainder of ground chicken at less than 5 million serving per day. Or alternatively and arguably more accurate is the average 169,585,818 servings per day of chicken broken down to 107,074,562 servings per day of parts and 62,511,256 servings per day of combined ground and comminuted chicken.

Table 48: The servings per day and percent of the U.S. population consuming chicken commodities.

Commodity	Servings per Day	Standard Deviation	Population%	Servings%
CHICKEN	169,585,818	11,503,970	52.6	100.0
PARTS	107,074,562	7,655,109	33.2	63.1
GROUND	4,546,506	804,414	1.4	2.7
COMM	57,964,750	4,395,304	18.0	34.2
GC=GROUND+COMM	62,511,256	4,676,879	19.4	36.9
PARTS+GROUND+COMM	169,585,818	11,503,970	52.6	100.0
PARTS+GC	169,585,818	11,503,970	52.6	100.0
US POPULATION	322,281,961			

Table 49 shows average daily chicken consumption in grams of chicken commodity on a population basis. The average over all chicken containing food codes as a high and low estimate average taken as all the "chicken or turkey" food codes (21.6% of total 415 chicken food codes) are all chicken (GRM) or alternatively contain 50% chicken (GRM50). This means the total average grams equals 148.2 with the averages for parts, ground, and comminuted summing to that value. **Table 50** shows the average percent with an overall average of 64.7% for chicken parts, 33% for comminuted chicken, and 2.2% for ground chicken. Or alternatively and arguably more accurate is the percentage of 100% chicken broken down to 64.7% parts and 35.3% combined ground and comminuted. The risk assessment used the GRM50 estimates to derive empirical probability of illness estimates.

Table 49: Daily chicken consumption in grams of chicken commodity.

	Average Daily Consumption from Day 1 NHANES for Population					
Commodity	**GRM[a]**	**SEM[a]**	**GRM50[b]**	**SEM[b]**	**AVEGRM[c]**	**SEM[c]**
TURKEY	120.3	3.7	79.7	2.3	100.0	2.2
PARTS	14.5	1.8	10.9	1.6	12.7	1.2
GROUND	6.9	1.3	4.1	0.7	5.5	0.7
COMM	98.9	3.6	64.7	2.2	81.8	2.1
GC=GROUND+COMM	105.8	3.9	68.8	2.2	87.3	2.3
PARTS+GROUND+COMM	120.3	2.5	79.7	1.6	100.0	1.5
PARTS+GC	120.3	4.1	79.7	2.5	100.0	2.4

a Grams consumed per day without subtracting chicken or turkey food codes
b Grams consumed per day subtracting 50% of grams per day for chicken or turkey food codes
c Average grams per day for a and b

Table 50: Daily consumption (%).

Average Daily Consumption% from Day 1

Commodity	GRM Percent	GRM50 Percent	AVE Percent
CHICKEN	100.0	100.0	100.0
PARTS	64.2	67.4	64.7
GROUND	2.6	1.9	2.2
COMM	35.0	30.7	33.0
GROUND + COMM	37.6	32.6	35.3

Table 51 shows average daily consumption in grams chicken commodity by consumer domain. This means that the denominator of the average is only from the part of the U.S. population that consumed the parts, ground, or comminuted chicken.

Table 51: Average daily consumption for commodity domain.

	Average Daily Consumption from Day 1 NHANES for Commodity Domain					
Commodity Domain	**GRM[a]**	**SEM[a]**	**GRM50[b]**	**SEM[b]**	**AVEGRM[c]**	**SEM[c]**
CHICKEN	157.4	2.4	139.0	2.4	148.2	1.7
PARTS	142.1	3.7	135.7	3.5	138.9	2.5
GROUND	124.1	23.0	78.9	12.1	101.5	13.0
COMM	144.0	4.6	111.5	3.3	127.8	2.8
GC=GROUND+COMM	143.7	4.4	110.0	3.4	126.8	2.8

a Grams consumed per day without subtracting chicken or turkey food codes
b Grams consumed per day subtracting 50% of grams per day for chicken or turkey food codes
c Average grams per day for a and b

Table 52 shows the percentiles of average daily chicken consumption as high, low, and average values.

Table 52: The percentiles of average daily chicken consumption as high, low, and average values.

	ChickenGRM		**Chicken GRM50**		**AVE ChickenGRM**	
PERCENTILE	**GRAMS**	**SEM**	**GRAMS**	**SEM**	**GRAMS**	**SEM**
1%	9.7	1.0	6.4	0.8	8.1	0.9
2.5%	18.0	1.2	13.1	0.8	15.5	1.0
5%	28.3	1.1	19.6	2.2	24.0	1.7
10%	38.6	1.3	33.7	1.0	36.2	1.2
20%	59.7	1.6	50.8	0.8	55.3	1.2
50%	119.3	3.3	106.5	2.6	112.9	2.9
Mean	157.4	2.4	139.0	2.4	148.2	2.4
80%	239.8	7.3	209.3	4.7	224.6	6.2
90%	332.9	11.5	285.0	8.3	308.9	10.1
95%	419.6	10.1	359.7	7.5	389.6	8.9
97.5%	505.0	16.8	432.0	11.2	468.5	14.3

99%	626.9	28.4	526.8	14.5	576.8	22.6

Figure 52 shows the distribution approximation, generated in (@Risk version 7.6.1*@Risk The Decision Tools Suite*), to the average total daily grams chicken consumption. The best fit is for a gamma distribution. The percentiles are shown in **Table 53**. The percentiles do not exactly match those in Table 5 because the distribution takes the average uncertainty of all or 50% chicken in the "chicken or turkey" food codes to be 75%. These percentiles and means are nearly identical to the average of the two separate chicken total grams and chicken grams 50 distributions.

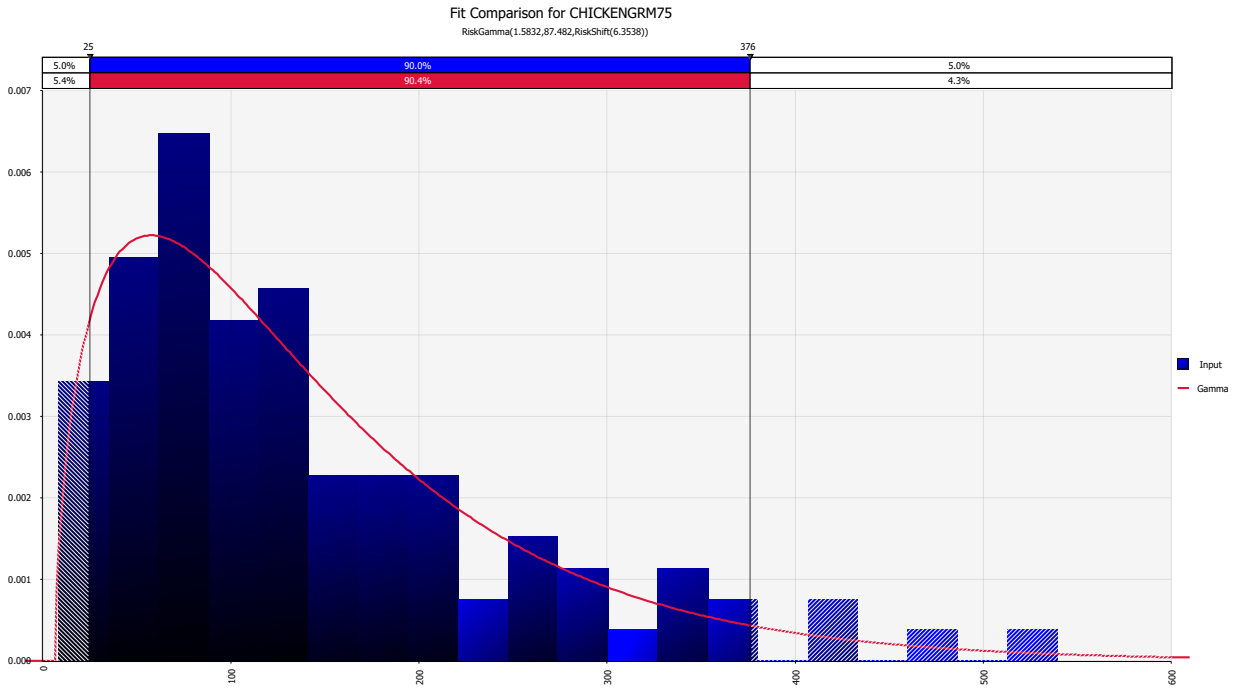


Figure 52: The distribution approximation to the average total daily grams chicken consumption.

Table 53: The gamma distribution percentiles.

	Chicken Grams75										
Stats	Input	Gamma	Pecentiles	Input	Gamma	Pecentiles	Input	Gamma	Pecentiles	Input	Gamma
Minimum	8.4	6.4	1%	8.4	12.5	40%	90.7	94.5	80%	220.1	219.4
Maximum	539.5	∞	5%	25.0	24.1	45%	101.1	105.4	85%	264.4	249.6
Mean	144.9	144.9	10%	37.3	35.1	50%	112.0	117.0	90%	313.0	291.2
Mode	≈134.67	57.4	15%	48.4	45.2	55%	127.9	129.6	95%	376.1	360.7
Median	112.0	117.0	20%	57.7	54.8	60%	135.1	143.3	99%	539.5	517.1
Std Dev	110.3	110.1	25%	65.3	64.4	65%	151.9	158.4			
Skewness	1.4	1.6	30%	68.6	74.1	70%	174.6	175.6			
Kurtosis	4.6	6.8	35%	81.4	84.1	75%	198.4	195.5			

Appendix C Theory

The methods used as basic components of the risk assessment model are defined in this appendix.

Methods for estimating the *Salmonella* prevalence

It is necessary to estimate the prevalence of sample units (i.e., carcasses, parts, and comminuted samples) where *Salmonella* is present at levels above the LOD of the assay. This requires weighting the sampling information from each establishment to account for the large range of establishment production volumes. Using broiler carcasses as an example, the target population consists of the V carcasses produced during the period of interest. Associated with each carcass is one or more attributes of interest, denoted Y_k, and the production volume for the establishment where slaughter occurred. The objective of a survey is to estimate some function of the population total, which is defined as

$$T_y = \sum_{k=1}^{V} Y_k .$$

For most food-safety applications the target parameter is the population mean $\frac{T_y}{V}$. When estimating the prevalence of a pathogen, $Y_k = 1$. when the pathogen is present and 0 otherwise. In other applications, it could be the information regarding serotype.

ere are two approaches to describing the estimation strategy using a design-based inferential paradigm. The sample design assumes that a sample of slaughter establishments (clusters) is selected for testing and samples of the commodity of interest are collected from each selected establishment. This is a typical application of two-stage cluster sampling, where establishments represent the clusters (Cochran, 17; Särndal, 1992). The sample design for selecting a sample from M . establishments will define a first-stage probability of selection,

$$P\left(\text{establishment } j \text{ is selected}\right) = \pi_{1j}, j = 1,...M .$$

The Horvitz-Thompson estimator (Cochran, 1977; Fuller, 2009) can be used to estimate a population tal and is given by

$$\hat{Y} = \sum_{j=1}^{m} \frac{\hat{Y}_j}{\pi_{1j}},$$

where m . is the number of establishments sampled, π_{1j} represents the probability of selecting establishment j. , and $\hat{Y}_j$ is the estimator for the total of the target parameter in the establishment. Ne, in this case, $\pi_{1j} = 1$ because samples will be collected from all M establishments.

Within an establishment, FSIS personnel randomly select a time for sample collection. For carcass sampling, an initial carcass is selected at that time. From this carcass, FSIS personnel count backward or forward five carcasses and select the next carcass (FSIS, 2021).

For all sampling programs, FSIS collected samples within an establishment at regular intervals, so an assumption of systematic sampling is reasonable. Sampling within establishment j yields n_j samples and sample unit (e.g., a chicken carcass) i has a second-stage probability of inclusion of $\pi_{2i} = \frac{n_j}{V_j}$, where V_j is the total number of units produced by that establishment. The key difference between simple random sampling and the more systematic nature of the FSIS sample design is that the joint inclusion probability for all samples within the sample period (e.g., weekly) is $\pi_{2,ii'} = 0$. . The Horvitz-Thompson estimator (Cochran, 1977; Fuller, 2009) of the total for the test outcomes in establishment j is

$$\hat{Y}_j = \sum_{i=1}^{n_j} \frac{y_{ij}}{\pi_{2i}}.$$

r this example, $y_{ij} = 1$ when a sample tests positive for a *Salmonella* and 0 otherwise, so $\hat{Y}_j$ is the estimator of the total number of test-positive carcasses across the entire volume of production in establishment j. Therefore, in the case where y_{ij} is binary, the estimator of the proportion of test-positive carcasses is

$$\hat{P}_j = \widehat{Y}_j = \frac{\hat{Y}_j}{V_j}.$$

Alternatively, if y_{ij} . is the pathogen count per unit volume (e.g., *Salmonella* colony forming units per milliliter (cfu/mL)), then $\widehat{Y}_j$ is the average microbial count per unit volume across all sample units produced by establishment j.

The total across all establishments is

$$= \sum_{j=1}^{m} \sum_{i=1}^{n_j} \frac{y_{ij}}{\pi_{2i}\pi_{1j}}.$$

Given that $\pi_j = 1$ because samples are collected from all establishments, the population total for y is estimated by

$$\hat{Y} = \sum_{j=1}^{M} \sum_{i=1}^{n_j} \frac{y_{ij}}{\pi_{2i}\pi_{1j}} = \sum_{j=1}^{M} \sum_{i=1}^{n_j} \frac{V_j y_{ij}}{n_j}$$

and

$$\hat{Y} = \frac{1}{V}\sum_{j=1}^{M}\sum_{i=1}^{n_j}\frac{V_j y_{ij}}{n_j}$$

is a design-unbiased estimator of the mean. When the target parameter is the prevalence of an indicator organism of pathogen, the estimator can be written as

$$\hat{P} = \sum_{j=1}^{M}\frac{V_j}{V}\hat{P}_j = \frac{1}{V}\sum_{j=1}^{M}\frac{V_j s_j}{n_j},$$

Where s_j is the number of positive samples in establishment j.

When the estimation strategy is viewed as an application of two-stage cluster sampling, the variance estimator for the population total is given by

$$var\left[\hat{Y}\right] = M\left(M-m\right)\frac{\hat{\sigma}^2_{between}}{m} + \frac{M}{m}\sum_{j=1}^{m}V_j\left(V_j - n_j\right)\frac{\hat{\sigma}^2_{within}}{n_j}.$$

Noting that $M = m$ and $V_j \gg n_j$ for all FSIS performance standards applications yields

$$var\left[\hat{Y}\right] = \sum_{j=1}^{M}\left(\frac{V_j}{V}\right)^2\frac{\hat{\sigma}^2_{within}}{n_j}$$

because the contribution of the between-cluster sampling variance is zero.

If the population parameter of interest is the proportion of *Salmonella*-positive carcasses in the population, the variance estimator is given by

$$var\left[\hat{Y}\right] = \sum_{j=1}^{M}\left(\frac{V_j}{V}\right)^2\frac{p_j\left(1-p_j\right)}{n_j - 1},$$

where p_j is the proportion of positive samples in establishment j.

Methods for estimating the contamination distribution

Due to differences in testing methods and data availability, the estimation approach used for each product type varied.

Describing the levels of *Salmonella* found on each product (e.g., carcasses) is more difficult than the estimation of proportion of positive samples because a negative test result is not necessarily indicative of the absence of the pathogen. Rather it indicates that *Salmonella* can be at some level below what can be detected by the combination of the sample collection method (e.g., rinse sample) and assay. The solution is to summarize the data with a distribution function $f\left(D_{test}\right)$ that allows for the possibility of negative samples while also adequately modeling the possibility of high levels of contamination. For estimating the *Salmonella* level on carcasses, the level distribution $f\left(D_{test}\right)$ relies on the following

assumptions about the data and assay.

- The proportion of samples where *Salmonella* is detected decreases dramatically as broiler chicken carcasses move through the slaughter process (De Villena, 2022). At rehang a large fraction of samples are positive for *Salmonella* and nearly 10% have sufficiently high levels of *Salmonella* to be enumerated (i.e., >10 cfu/mL). In contrast, the current fraction of all post-chill broiler carcasses whose level of contamination exceeds the 0.033 cfu/mL LOD of the assay is 0.031 (i.e., roughly 96.9% of observation are censored/negative for *Salmonella*). The fraction of samples whose level exceeds the theoretical LOQ for the current qPCR level technology was 0.0015. Both proportions are significantly below the recommended lower bounds for the guidelines provided by previous studies (Helsel, 2005; Williams, 2013a) for the fitting of a contamination distributions to multiply censored data, though this does not obviate the need to estimate a level distribution for the risk assessment.

- The qPCR assay used for level of carcasses differs from methods such as plate counts and the MPN technique in the sense that it lacks a well-defined probabilistic model to explain the enumeration (level) process as a function of the observed PCR cycle times. If such a model existed, the parameters of a contamination distribution could be estimated using methods such as maximum likelihood (Pouillot, 2013) or Markov chain Monte Carlo (Williams, 2012a; Williams, 2012b). While the mathematical basis for qPCR is well understood for estimating the level for a single sample (Kralik, 2017), the methods needed to summarize the data at a population level have yet to be developed and an attenuation bias of roughly ½ log in average level always occurs when the measurement error in the estimated level is ignored (Armstrong, 1998; Williams, 2012a). While the appropriate estimation methods are being developed, the additional variability associated with qPCR values being estimates of the level, rather than true values, are partially accounted for by treating the integer valued qPCR levels as the realization of Poisson process. The Poisson process assumption most likely still underestimate the variability at low levels because the qPCR estimates have estimated standard deviations on the order of 0.5 log10 at levels of 2 log10 cfu/mL in chicken rinse samples ((Joelsson, 2022)).

The specifics of the Poisson lognormal distribution used for the carcass data is as follows. The integer *Salmonella* level results reported by qPCR are assumed to follow a Poisson distribution whose rate parameter (λ) varies according to a lognormal distribution with parameter vector $\theta = (\mu, \sigma)$. This distribution is the Poisson-lognormal distribution (Bulmer, 1974; Engen, 2002; Izsak, 2008). Using the natural log scale, the probability of observing x organisms is

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma x!}\int_0^\infty e^{-\lambda}\lambda^{x-1}e^{-\frac{(\ln\lambda-\mu)^2}{2\sigma^2}}\,d\lambda. \quad (0.1)$$

Assuming that N samples are collected randomly, the level estimates can be treated as independent random variables. In this case, the likelihood function for fitting the Poisson-lognormal to data consisting only of samples that are above the LOQ for an assay is the product of the individual probability density functions:

$$\Lambda_x = \prod_{i=1}^{N} p(x_i).$$

When the screening test is added to the laboratory protocol, the three possible outcomes for a test are:

1. the organism is not detected ($\delta = 0$; denoted),
2. the organism is detected by the screening test, but the cycle time for the PCR test exceeds the threshold for declaring greater than 10 cfu/mL ($\delta = 1, x = 10$; denoted $qual$),
3. the organism is detected by the screening test and the qPCR estimated level is greater than 10 ($\delta = 1, x \geq 10$; denoted $quan$)

The total number of samples, N, can be subset $N = N_{nd} + N_{qual} + N_{quan}$ to account for the three possible outcomes and the different detection and quantitation limits. Using these results and assuming independence between samples, leads to the likelihood function

$$\Lambda_{MLE} = \left[p(\delta = 0) \right]^{N_{nd}} \times \left[p(x = 10) - p(\delta = 0) \right]^{N_{qual}} \times \prod_{i=1}^{N_{quan}} p(x_i).$$

Maximum likelihood estimates of the parameters describing the underlying lognormal distribution $(\hat{\mu}_{MLE}, \hat{\sigma}_{MLE})$ are derived by maximizing Λ_{MLE}. The $p(x_i)$ values are calculated from Poisson lognormal probability density function, while determining the probability of a negative screening test requires additional work.

The *Salmonella* screening test using by FSIS employs a $v_{screen} = 30ml$ aliquot of rinsate so the LOD is $1/v_{screen} = 0.03$. . The adjustments for incorporating screening test results into the Poisson-lognormal likelihood are different from the approach taken using a continuous distribution function like the lognormal (Helsel, 2005; Shorten, 2006) because the lognormal distribution naturally scales to values less than one. The Poisson-lognormal, however, requires rescaling the lognormal contamination distribution to account for the change in sample volume.

erive the adjustment, let λ describe the average level per mL and assume it follows a lognormal distribution with parameters (μ, σ). Let γ be a scalar transform of λ that represents conversion between colony-forming-units of the target bacteria per gram (cfu/mL) and cfu within the larger mass of the screening test sample $(\gamma = v_{screen}\lambda)$; then the cumulative probability distribution for γ is

$$F\left(\lambda < \frac{\gamma}{v_{screen}} \right) = \Phi\left(\frac{\ln(\gamma / v_{screen}) - \mu}{\sigma} \right) = \Phi\left(\frac{\ln(\gamma) - \left(\mu + \ln\left(v_{screen} \right) \right)}{\sigma} \right),$$

where Φ is the standard normal distribution. Consequently, $\gamma \sim \text{lognormal}\left(\mu + \ln(v_{screen}), \sigma \right)$.

efore, in the case of the Poisson-lognormal, the μ parameter is shifted to the right to account for the increased mass of the sample used in the screening test.

For a screening sample with volume v_{screen}, the probability of a negative test is

$$p(\delta = 0) = \frac{1}{\sqrt{2\pi}\sigma}\int_0^{\infty} \frac{e^{-\lambda}}{\lambda} e^{-\frac{(\ln(\lambda)-(\mu+\ln(v_{screen})))^2}{2\sigma^2}} d\lambda.$$

The production volume and sampling intensities vary substantially across establishments, so the parameter estimates $\hat{\mu}, \hat{\sigma}$ for the true distribution of pathogen levels are derived using weighted maximum likelihood estimation (Williams, 2013b), where the full likelihood is

$$\Lambda_{weighted}(\mu, \sigma) = \prod_{j=1}^{M} \frac{E[\pi_j] p^*(x_j \mid \mu, \sigma)}{\pi_j}.$$

Bias adjustment for whole carcass levels

The bias in the parameters of the level distribution has a substantial impact on the upper tail of the contamination distribution. The upper-right tail of the distribution most likely accounts for servings with a higher probability of causing illness, so the following approach to bias adjustment commodities was applied.

While the precision of the qPCR level data is low, the FSIS prevalence estimator $\widehat{P}$ is relatively precise because the sample size is large (~10,000 samples per year) and the assay used to determine presence of *Salmonella* by the screening test relies on enrichment and all screen-test positive samples undergo confirmation to avoid false positives. For example, an assessment of the power of the prevalence estimator derived from the annual chicken carcass performance standards finds that the survey has approximately 80% power to detect a change in prevalence of 22% (i.e., a change from the annual prevalence of approximately 0.007 has an 80% chance of being declared statistically significant).

The approach for adjusting the parameters of the lognormal contamination distribution starts with the observation that previous simulation study work provides reasonable estimates of the magnitude of the bias in $\hat{\mu}$ and $\hat{\sigma}$ parameters. These two parameters also provide an estimate of the prevalence, which is given by $\tilde{P} = 1 - F(LOD, \hat{\mu}, \hat{\sigma})$, where F is the lognormal cumulative distribution function. The method employed to correct the bias in the carcass distribution is to first choose a bias correction for the $\hat{\mu}$ parameter based on the previous simulation study. A value of -0.4 was chosen because the bias is expected to lie somewhere between -0.35 and -0.65. The lower end of the range was chosen because the Poisson component of the Poisson-lognormal distribution is likely to account for some of the measurement error in the underlying level.

The next step is to use the biased adjusted μ parameter ($\tilde{\mu} = \hat{\mu} - 0.4$) and the estimated prevalence $\widehat{P}$ to determine a new prevalence from the lognormal distribution, which is given by $\tilde{P} = 1 - F(LOD, \tilde{\mu}, \hat{\sigma})$. The bias-adjusted σ . parameter is chosen using a nonlinear optimization routine to choose $\tilde{\sigma}$ so that $\tilde{P} = 1 - F(LOD, \tilde{\mu}, \tilde{\sigma}) = \hat{P}$ Using this approach, the parameters describing the level of *Salmonella* on chicken carcasses is $\tilde{\mu} = -4.512$ and $\tilde{\sigma} = 1.624$

Methods that scale rinsate levels to whole carcass levels

In order to scale the level distribution of *Salmonella* per mL of rinsate up to an estimate of the total number of *Salmonella* per carcass, the μ_{mL} parameter of the lognormal distribution was adjusted. This adjustment accounts for the 400 mL rinse volume and the estimated removal rate (r) of *Salmonella* from a broiler chicken carcass during rinse sampling:

$$\mu_X = \mu_{mL} + \log_{10}\left(\frac{400}{r}\right).$$

Rinse sampling only removes a portion of the microbes on the carcass. Few studies have focused on determining the proportion of microbes removed by rinsing. An extensive study was performed in the late 1980s (Lillard, 1988), which estimated approximately 14% of *Enterobacteriaceae* and 10% of aerobic bacteria are removed using an automated shaking machine and an agitation time of 60 seconds. Attachment characteristics of *Salmonella* are assumed to be consistent with the general class of *Enterobacteriaceae*, so the 14% removal rate was chosen with uncertainty in this estimate characterized by a $r \sim beta(14,86)$.

Methods that scale whole carcass levels to a per serving level

Given the whole carcass level, the final step is to adjust the level distribution to a per serving basis. This requires an estimate of the number of servings per carcass. An analysis of the 2019 NHANES data finds that the number of chicken servings consumed per day in the U.S. is 169,585,818 which gives a total number of servings per year of 61.90 billion. The number of broiler chickens slaughtered 2019 was 9.224 billion, which provides an estimate of the average 6.71 servings per carcass. This adjustment accounts for the 400 mL rinse volume and the estimated removal rate (r) of *Salmonella* from a broiler chicken carcass during rinse sampling:

$$\mu_{serving} = \mu_X + \log_{10}\left(\frac{1}{6.71}\right).$$

The carcass levels can also be converted to a per gram basis.

Concentration estimation for chicken parts.

Concentration information for chicken parts was only collected during the 2012 chicken parts microbiological baseline study (FSIS, 2013) using a 3-tube, 3-dilution MPN method. That microbiological baseline study also found a weighted prevalence of 0.24 for *Salmonella*. A subset of these data consisting only of samples of legs, breast or wings was used to set the performance standards. The prevalence of this subset was $\hat{P}_{2012} = 0.28$. Fitting a volume weighted lognormal distribution to the legs, breasts, and wings level data resulted in parameter estimates of $\hat{\mu}_{parts,2012} = -2.131$ and $\hat{\sigma}_{parts,2012} = 1.027$. Since these data were collected, the prevalence has declined substantially, with the prevalence for calendar year 2021 being $\hat{P}_{2021} = 0.067$ (Williams, 2022).

As was the case with the carcass data, the contamination distribution requires adjustment to reflect the change in prevalence. The reduction in contamination on chicken parts has been largely achieved by the application of additional antimicrobials during processing, so it is assumed that the change in the contamination distribution would be almost exclusively represented by a reduction in the μ parameter for the distribution describing the contamination distribution in the 2012 microbiological baseline. Using this logic, the $\hat{\sigma}_{2012}$ parameter was held constant and a nonlinear optimization routine was used to choose $\hat{\mu}_{2021}$ such that $\hat{P}_{2021} = 1 - F\left(LOD, \hat{\mu}_{2021}, \hat{\sigma}_{2012}\right)$. The parameter estimates representing the level of *Salmonella* per mL for chicken parts are $\hat{\mu}_{parts,2021} = -3.07$ and $\hat{\sigma}_{parts,2021} = 1.027$.

The chicken parts samples are derived from a 400 mL rinse sample consisting of 4 pounds of chicken parts, which is assumed to be the equivalent weight of the average chicken carcass. Thus, the same adjustment approach for the effect of the subsampling of the 400 mL and the removal rate parameter for a 1-minute rinse sample were applied to convert the level per milliliter to a per gram basis are applied. The weight of the parts sample is 1816 grams (i.e., 4 lbs), which is assumed to be equivalent to the weight of the average broiler chicken carcass. Specifically, the adjustment factor of $\log 10\left(1/\left(454 \times 4\right)\right)$ was applied to the μ parameter. The same adjustment factor can be applied to the carcass level distribution to convert that level distribution to a per gram basis.

Concentration estimation for comminuted chicken

FSIS has enumerated a subset of comminuted chicken and turkey performance standards samples since 2015. The subsample of enumerated samples consisted of samples that were analyzed at one of the FSIS three laboratories. These samples are assumed to be a random subsample of comminuted chicken samples because the location of sample collection does not determine the laboratory at which the samples are analyzed (i.e., the overnight courier service ships all samples to a central location in Memphis, Tennessee and then ships the samples to one of the three laboratories. Thus, there is no advantage to having to a laboratory serve a specific geographic region of the country).

Samples collection in 2015 and 2020 were not used in the analysis because not all establishments were consistently tested prior to the implementation of performance standards in 2016 and the MPN analysis was not consistently performed. The comminuted chicken dataset consisted of 1,815 samples of which 387 were positive on the screening test.

Given the lower priority of MPN analysis, some samples were not analyzed due to limited staffing in the laboratory. The number of samples that were not analyzed was 21 and 1 for the comminuted chicken and turkey datasets respectively. These missing results were addressed using the imputation by simple random sampling method (van Buuren, 2011).

A weighted maximum likelihood routine was used to fit a lognormal distribution to the comminuted chicken dataset. The estimated parameters for were $\hat{\mu}_{comm,chick} = -3.700$, $\hat{\sigma}_{comm.chick} = 1.949$. The implied prevalence for comminuted chicken, derived from the cumulative distribution of the lognormal evaluated at the LOD=1/325, is 0.271. This estimate is similar to the prevalence estimate for calendar year 2021, which was 0.280.

Methods for combining contamination distributions

Consumer exposures to *Salmonella* differ for the three different product types. A single distribution describing contamination, on a per gram basis at the end of production, can be constructed for all chicken. The mean and standard deviation of a mixture of the three distributions is calculated using

$$\mu_{all} = \sum_{j=1}^{3} w_j \mu_j$$

and

$$\sigma_{all} = \sqrt{\sum_{j=1}^{3} w_j ((\mu_j - \mu)^2 + \sigma_j^2)}, .$$

with w_j is the proportion of serving for each of the three product types. The estimated parameters for were $\mu_{true} = -0.333$ and $\sigma_{true} = 0.516$.

Modeling the relationship between prevalence and level

Risk assessments that evaluate the difference between performance standards approaches based on prevalence or level can give the impression that two approaches are inherently different (Lambertini, 2019). Nevertheless, both approaches are related because reductions in prevalence, once a flock enters slaughter, are achieved by applying interventions that result in an additional log reduction compared to the baseline scenario (e.g., the addition of an organic acid spray). Similarly, if incoming levels of *Salmonella* are similar at two establishments, but the establishments achieve different log reductions in pathogen (***Figure 53***), the establishments will have different prevalence which are determined by $1 - F\left(\log 10\left(1/30\right), \mu, \sigma\right)$.

Figure 53*: Log 10 transformed level distributions with means of -2.5 and -1.*

This phenomenon can also be expressed probabilistically as an application of Bayes Theorem using the following argument. Consider a distribution that describes the log10-transformed contamination distribution with parameters μ and σ and assume that σ remains roughly constant. To simplify the notation, express this distribution as $P\left(\mu\right)$. Next consider that a sample has level x that is greater than a threshold value denoted by d. Bayes Theorem yields the relationship

$$P\left(\mu \middle| x \rangle d\right) = \frac{P(x > d \mid \mu) P\left(\mu\right)}{P(x > d)}.$$

The duality of the relationship between level and prevalence is demonstrated by noting that the probability of $x > d$ can be replaced by $x > LOD$, which is the event that the sample is positive on the screening test, with $P(x > LOD) = P\left(test+\right)$ and $P\left(test+\right) = 1 - F\left(LOD, \hat{\mu}, \hat{\sigma}\right)$.

Methods for modeling illnesses

FSIS is interested in evaluating the effectiveness of different approaches to reducing illnesses associated with *Salmonella* contaminated poultry. The first step in the risk assessment process is to define new probabilistic models to address potential risk management scenarios. For simplicity, the symbols for each variable used in the development of the methodology are treated as fixed values. The parameterization of the probability distributions used to describe variability in factors such as the consumed dose are summarized later.

The starting point for the risk assessment is the concept that the annual number of illnesses is the product of the probability of illness per serving times the number of servings, so

$$I = N_{\text{servings}} P(ill),$$

Where $N_{servings}$ is the number of serving of chicken consumed per year and $P(ill)$ is the probability of illness per serving. The total number of illnesses from chicken are determined by the estimated total number of domestically acquired foodborne cases of salmonellosis (Scallan, 2011) multiplied by the attribution fraction (IFSAC, 2019). The number of servings can be estimated using the estimated per capita weight of chicken available for consumption (USDA-ERS, 2021) times the average serving size (Appendix B). Given that the motivation for revised performance standards is driven by a lack of observed changes in overall cases of salmonellosis reported by FoodNet, the probability of illness per servings should logically be directly tied to CDC illness estimates, which imposes the requirement that

$$P(ill) \approx I / N_{servings}.$$

This formulation for the probability of illness will be referred to as the attribution-based probability of illness per serving.

Interest lies in addressing specific serotypes or groups of serotypes, indexed by $g = 1,...G$, so the illnesses are decomposed by

$$I = I_1 + I_2 + ..., I_G = N_{servings,1} P_1(ill) + N_{servings,2} P_2(ill) +, ... N_{servings,G} P_G(ill).$$

A reasonable method for estimating the number of serving contaminated with serotype g is to use the fraction *Salmonella*-positive samples where serotype s_g is identified divided by the number of all *Salmonella*-positive samples s, so

$$N_{servings,g} = \frac{s_g}{s} N_{servings} = p_g N_{servings}.$$

Estimates of I_g can be derived by considering the attribution fraction for different serotypes/serogroups, as was the case in the original attribution study (Painter, 2013b).

While the probability of illness per servings will be required to "match" the observed value $P(ill) = I / N_{servings}$, the risk assessment model will need an additional level of detail so that changes in the levels of contamination can be assessed.

There is no evidence to suggest that any flock of broiler chickens is truly free of *Salmonella* contamination (De Villena, 2022), thus the preferred parameterization of $P(ill)$ assumes that all servings have the potential for some level of contamination, so that the random variable describing dose D describes the average number of pathogens in each serving. Note that because D describes an average level, it is possible for these level values to be much less than 1 organism per serving. The average level of *Salmonella* follows a distribution with probability density $f(D)$. The probability that a random person will become ill, given a microbial dose of average level D, is $P(ill|D)$. Averaging across all possible doses yields the probability of a person becoming ill. When D describes an average dose, the probability of illness given exposure described by a continuous dose distribution is

$$\tilde{P}(ill) = \int_0^\infty P(ill \mid D) f(D)\, dD = \int_0^\infty R(D) f(D)\, dD,$$

Where $R(D)$ is the dose-response function and the ~ sign indicates that this probability of illness is derived from a dose-response function.

Epi*X* Analytics supplied beta-Poisson dose-response functions that are appropriate for continuous dose distributions where the input variable is the average level per serving. What will be unique for this risk assessment is that specific serotypes will be grouped into a small number of clusters based on the estimated pathogenicity of the serotype. Let's assume there are three clusters representing high, medium, and low pathogenicity. Then the number of illnesses associated with a highly virulent serotype is

$$N_{servings,s}\tilde{P}_s(ill) = \int_0^\infty R_s(D) f_s(D)\, dD.$$

An assessment of the broiler microbiological baseline data (FSIS, 2009b) found insufficient evidence to reject the hypothesis of significant differences in the levels of contamination between serotypes assumed to be in the high pathogenicity cluster compared to the low pathogenicity cluster, so it is reasonable to assume $f_g(D) = f_{g'}(D)$ for all serotype clusters.

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.

Models for describing consumption dose distribution.

Data to directly estimate the parameters of the dose distribution at consumption (i.e., $f(D) = f(D_{consump})$) are typically only available for a small number of outbreaks. This risk assessment will use data collected at the end of production, which is represented as $f(D_{test})$. The lognormal distribution is appealing for describing microbial data from different locations in the food chain (Chen, 2001; Commeau, 2012; Gonzales-Barron, 2011; Pouillot, 2013; Williams, 2015). Furthermore, the lognormal distribution is mathematically convenient for scaling the level to account for sampling volumes and efficiencies (Williams, 2010; Williams, 2011b), modeling the effects of cooking (Bassett,

2010), growth (Shorten, 2006) and cross contamination (Chen, 2001). A lognormal distribution is obtained asymptotically even if intermediate processes that modify a lognormal distribution are not themselves lognormal (Mitzenmacher, 2003). This result is important because even if some intermediate processes are not lognormally distributed, it is reasonable to assume that $f\left(D_{consump}\right)$ follows a lognormal distribution.

If the focus of a risk assessment is to determine changes in risk due to measurements taken at the end of production, then the lognormal distribution allows modification of $f\left(D_{test}\right)$ through a single component by modeling $\log_{10}(D_{consump}) = \log_{10}(D_{test}) + \log_{10}(D_{atten})$, where $f\left(D_{atten}\right)$ is a lognormal distribution describing the cumulative change in average microbial level between production and consumption (i.e., it combines the effects of mixing, growth, partitioning, cooking and other processes). Assuming independence between level at the end of production and magnitude of attenuation, the mean and standard deviation of the consumption distribution can be computed directly from the means and variances of the distributions for D_{test} and D_{atten}. (i.e., $\mu_{test}, \mu_{atten}, \sigma^2_{test}, \sigma^2_{atten}$), where $\mu_{consump} = \mu_{test} + \mu_{atten}$ and $\sigma_{consump} = \sqrt{\sigma^2_{test} + \sigma^2_{atten}}$. In this case, μ and σ refer to the mean and standard deviation of the $\log_{10}$ transform of the random variable (e.g., $\log_{10}\left(D_{test}\right) \sim Normal\left(\mu,\sigma\right)$ and $D_{test} \sim Lognormal\left(\mu,\sigma\right)$).

Appendix D Analysis of FSIS' *Salmonella* Initiative Program (SIP) Data

Data Description

Exploratory Chicken Carcass Sampling Data:

In this analysis, we used data from FSIS' exploratory chicken carcass sampling data (hereafter Exploratory data) collection program that started in April 2022. This dataset is based on FSIS' ongoing performance standard sampling program, but it adds sample collection at the rehang location of the slaughter process. Additional laboratory testing was also performed to quantify indicator organism (Aerobic Count (AC) and Enterobacteriaceae (EB)) at both locations. Enumeration of *Salmonella* via quantitative PCR was also performed, with the first enumerated samples being collected in early August 2022.

As of October 31, 2022, a total of 5,425 samples had been collected from 206 slaughter establishments. Of these samples, 4,594 samples (85%) from 200 establishments had AC level data at both the rehang and post-chill locations. A subset of this data with enumerated AC level data at both locations and the occurrence of *Salmonella* were used to assess relationships between Explanatory and *Salmonella* Initiative Program (SIP) data.

Salmonella Initiative Program (SIP) Data:

The FSIS SIP program data available for this analysis spans April of 2015 through February of 2022 and contains over 122,000 observations in total. This data was collected by industry and contains samples from the rehang and post-chill locations of the slaughter process. Data consists of AC level, *Salmonella* occurrence, and EB at both locations. The exact methods to quantify indicator organisms was not specified in the dataset. The SIP data was reported for 63 total establishments. However, there were only 48 common establishments that had data that could be compared with the Exploratory dataset. Therefore, a subset of the SIP that had 84,700 samples was used in this analysis.

Methods:

Exploratory Data:

We used FSIS' Exploratory data to calculate average $\log_{10}$ AC values and average *Salmonella* prevalence (percent of positive occurrences) by establishments at rehang and post-chill sampling locations of the slaughter process. Samples with missing data for different establishments were removed before calculating average *Salmonella* prevalence at establishment level. The AC contained many reported values below the level-of-detection (LOD) which were termed as "censored" data for this analysis (Helsel, 2009; Helsel, 2010). The censored data was modeled using a log-normal distribution with parameters calculated by maximum likelihood estimation for each establishment (Williams, 2014).

For comparison with SIP data, we extracted data for 48 matching establishments from the Exploratory data. The subset of Exploratory data for 48 establishments (23.3% of all the establishments) had 1,274 samples at rehang (26.17% of the total samples) and 1,231 samples at post-chill (23.47% of the total samples).

SIP Data:

We used industry-reported SIP data to calculate log_{10} AC and *Salmonella* prevalence. The AC data contained many reported values below the level-of-detection (i.e., censored data). In particular, the post-chill AC data contained a significant percentage of censored data. Because it was not clear from the submitted data what the level of detection was for AC data at each establishment, all entries below 10 cfu/g were classified as censored. The same modeling technique was used for censored AC SIP data as was used for Exploratory data. The SIP data used for comparison contained 53,987 samples at rehang (63.76% of the total samples) and 57,875 samples at post-chill (68.35% of the total samples).

Comparison of Exploratory and SIP data:

In theory, exploratory and SIP data should closely match. While this was the case for some establishments, there was a significant disparity in calculated values for many other establishments. This was particularly pronounced for some of establishments with AC data at post-hang. To demonstrate the differences between the data, the SIP and Exploratory data are displayed together with the SIP data on the y-axis and the Exploratory data on the x-axis. The line representing identical data values has a slope of 1 going through the origin. For analyzing agreements and disparity between both datasets by establishments, we plotted two additional lines on the graphs for average AC values and average *Salmonella* prevalence. For example, we plotted two parallel lines with the same slope but with y-intercept values of 0.5 and -0.5 (see **Figure 54** and **Figure 55**). All data points between these two outer parallel lines were considered to have reasonably consistent values.

All the AC values were log_{10} transformed. Pearson's correlation coefficient (r) was used for quantifying association between average AC values and average *Salmonella* prevalence between Exploratory and SIP data at rehang and post-chill locations; statistical significance level was determined at $\alpha = 0.05$ ($P < 0.05$ significant, and $P \geq 0.05$ not significant).

Results:

The average log_{10} AC level at rehang for the Exploratory data was 4.264 (± 0.310) whereas for SIP data average log_{10} AC was 4.2% (± 0.644). The average AC level at post-chill for Exploratory data and SIP data were 1.237 (± 0.849) and -0.399 (± 1.240), respectively (**Table 54**).

The average *Salmonella* prevalence at rehang was 64.1% (±19.660) for Exploratory data and 71.2715% (±37.279) for SIP data. The average *Salmonella* prevalence at post-chill was 3.405% (±3.034) for Exploratory data and 2.747% (±12.960) for SIP data (**Table 54**).

Table 54: Summary of Exploratory and SIP data analysis results by establishment (n = 48 common establishments). SD is standard deviation.

	Exploratory data (2022) Mean (SD)		**SIP data (2015-2021) Mean (SD)**	
	Rehang	Post-chill	Rehang	Post-chill
Average Log_{10} AC	4.264 (0.310)	1.237 (0.849)	4.150 (0.644)	-0.399 (1.240)
Average *Salmonella* prevalence (%)	64.088 (19.660)	3.405 (3.034)	71.271 (37.279)	2.747 (12.960)

There was an average 2.888 log_{10} reduction in AC level from rehang to post-chill locations in the 48

establishments subset of the Exploratory data. However, the average $\log_{10}$ reduction in AC level in SIP data was unexpectedly high, 4.548, for the same sampling locations.

Below we describe results from the comparison of average $\log_{10}$ AC level and average *Salmonella* prevalence between Exploratory and SIP data at rehang and post-chill sampling locations of the chicken slaughter process.

Relationship between AC levels at rehang:

Figure 54 shows the relationship between FSIS and SIP data average $\log_{10}$ AC levels for 48 common establishments. No statistically significant relationship was detected, and the correlation was poor (correlation coefficient, r = 0.263; *P* = 0.071). Approximately 73% of the SIP establishments (n = 35) had better agreement with $\log_{10}$ AC levels in FSIS data; they were within 0.5 $\log_{10}$ AC values (between the black parallel lines in Fig. 1).

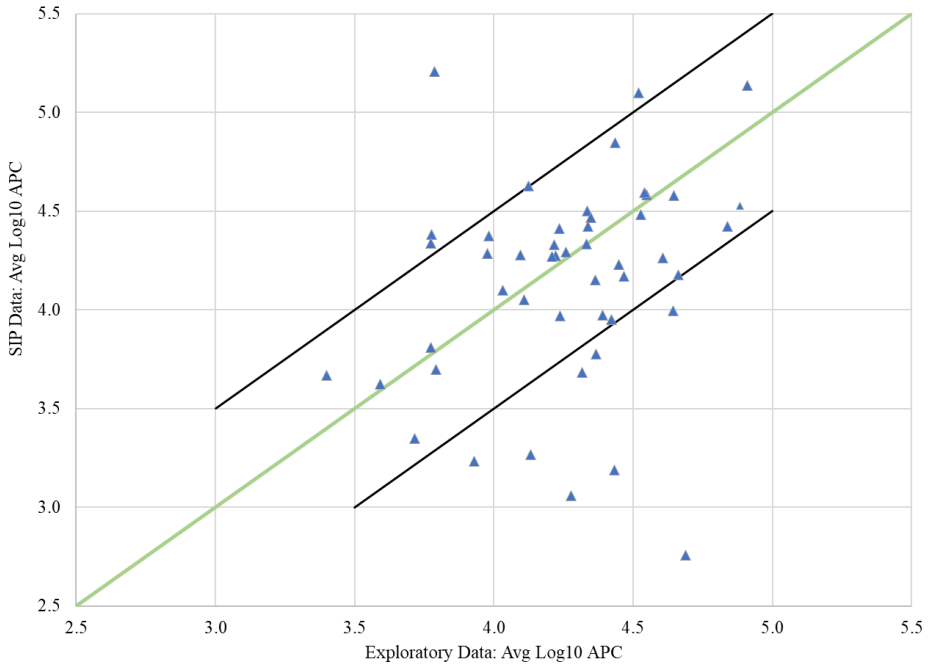


Figure 54: Relationship between per establishment Exploratory and SIP data average $\log_{10}$ AC levels at rehang.

Relationship between Salmonella prevalence at rehang:

Figure 55 shows the relationship between FSIS and SIP data average *Salmonella* prevalence for 48 common establishments. A statistically significant positive relationship was found for *Salmonella* prevalence between Exploratory and SIP data (r = 0.331; *P* = 0.022). However, only 37.5% of the SIP establishments (i.e., 18 out of 48) had a better agreement with *Salmonella* prevalence in FSIS data; they were within 10% prevalence (between the black parallel lines in Fig. 2).

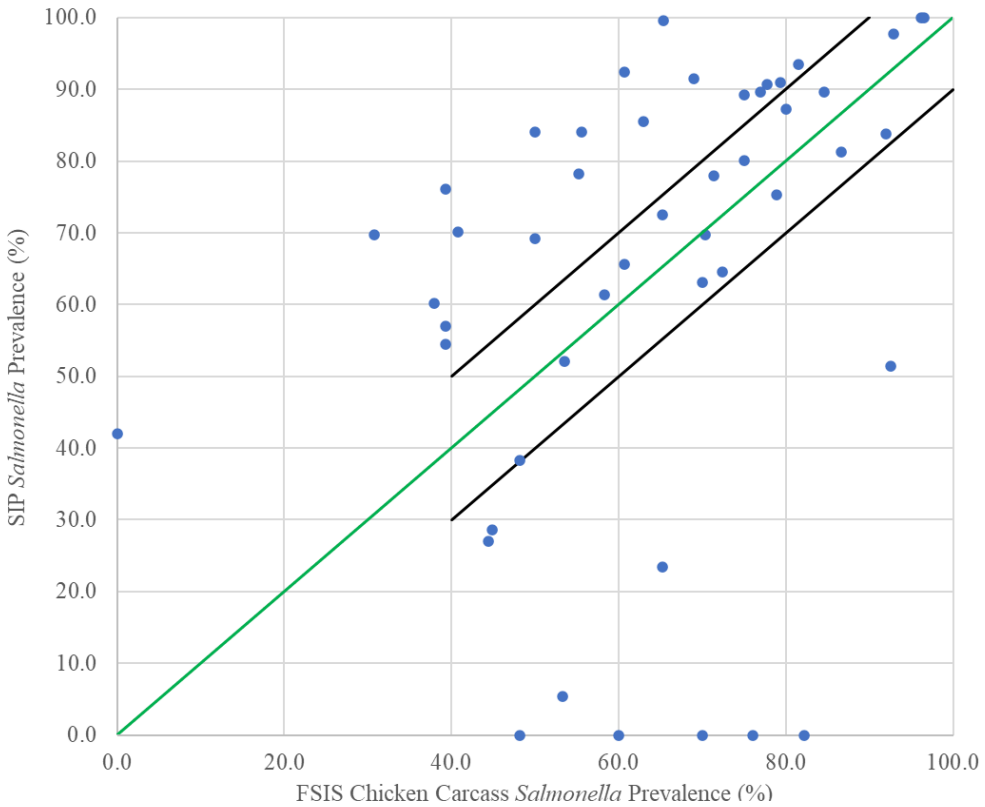


Figure 55: Relationship between per establishment Exploratory and SIP data average *Salmonella* prevalence (%) at rehang.

Salmonella Relationship between AC levels at post-chill:

Figure 56 and **Figure 57** show the relationship between FSIS and SIP data average $\log_{10}$ AC levels for the 48 common establishments. We removed two outlier establishments (at the lower right corner of Fig. 3) for calculations and better visualization of results (see Fig. 4). No statistically significant relationship between the FSIS and SIP data was found and the correlation was poor (r = 0.134; *P* = 0.374). Only 30% of the SIP establishments (14 out of 46) were in better agreement with $\log_{10}$ AC levels in FSIS data; they were within 0.5 $\log_{10}$ AC values (between the black parallel lines in Fig. 4).

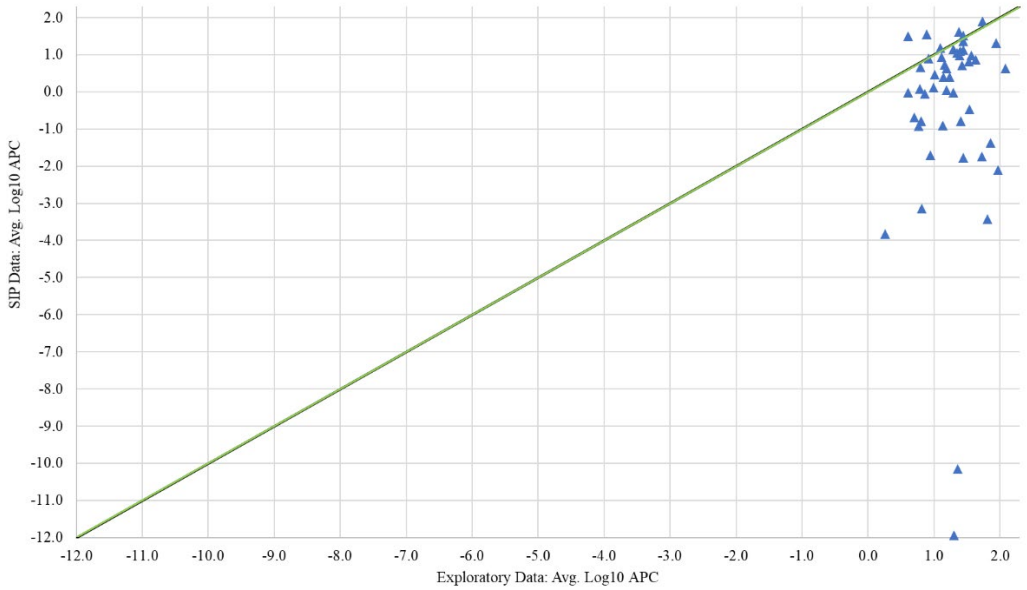


Figure 56: Relationship between per establishment Exploratory and SIP data Average log_{10} AC levels at post-chill.

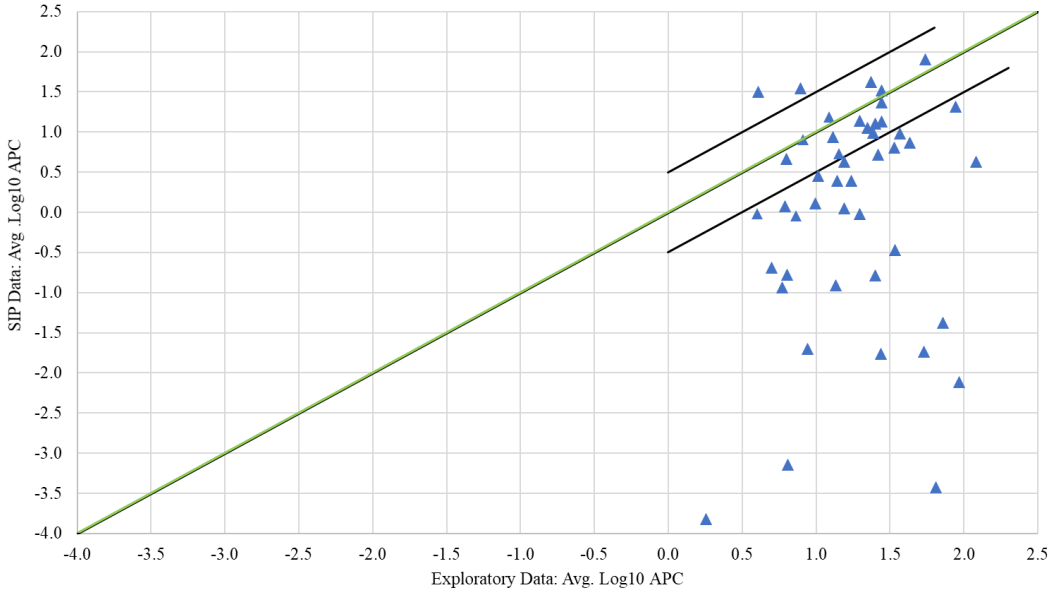


Figure 57: Relationship between per establishment Exploratory and SIP data Average log_{10} AC levels at post-chill after removing two outlier establishments.

Relationship between Salmonella prevalence at post-chill:

Figure 58 shows the relationship between FSIS and SIP average *Salmonella* prevalence for the 48 common establishments. No statistically significant relationship was found for *Salmonella* prevalence between Exploratory and SIP datasets (r = 0.103; *P* = 0.485). Only 31% of the SIP establishments (i.e., 15 out of 48) had a better agreement with *Salmonella* prevalence in FSIS data; they were within 2% prevalence (between the black parallel lines in **Figure 57**).

Because the FSIS dataset did not contain many observations for *Salmonella* post-chill prevalence and many of the values were identical or similar, some of the values in **Figure 58** appear to form a vertical line. It is particularly noticeable for values of 0% on the x-axis.

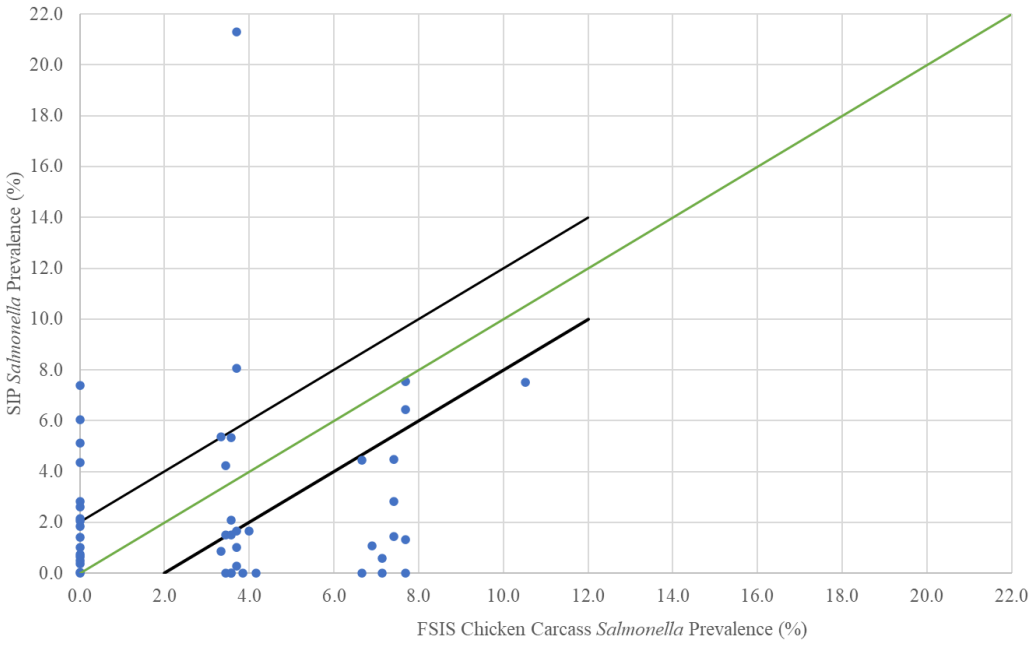


Figure 58: Relationship between per establishment Exploratory and SIP data *Salmonella* prevalence (%) at post-chill.

Figure 59 shows a comparison of the percentage of censored AC observations at post-chill at the 48 common establishments from most similar (on the left-hand side of the Figure) to least similar (on the right-hand side of the Figure). This high percentage of censored observations for SIP data contributed to implausibly low AC level averages at post-chill at some of the establishments.

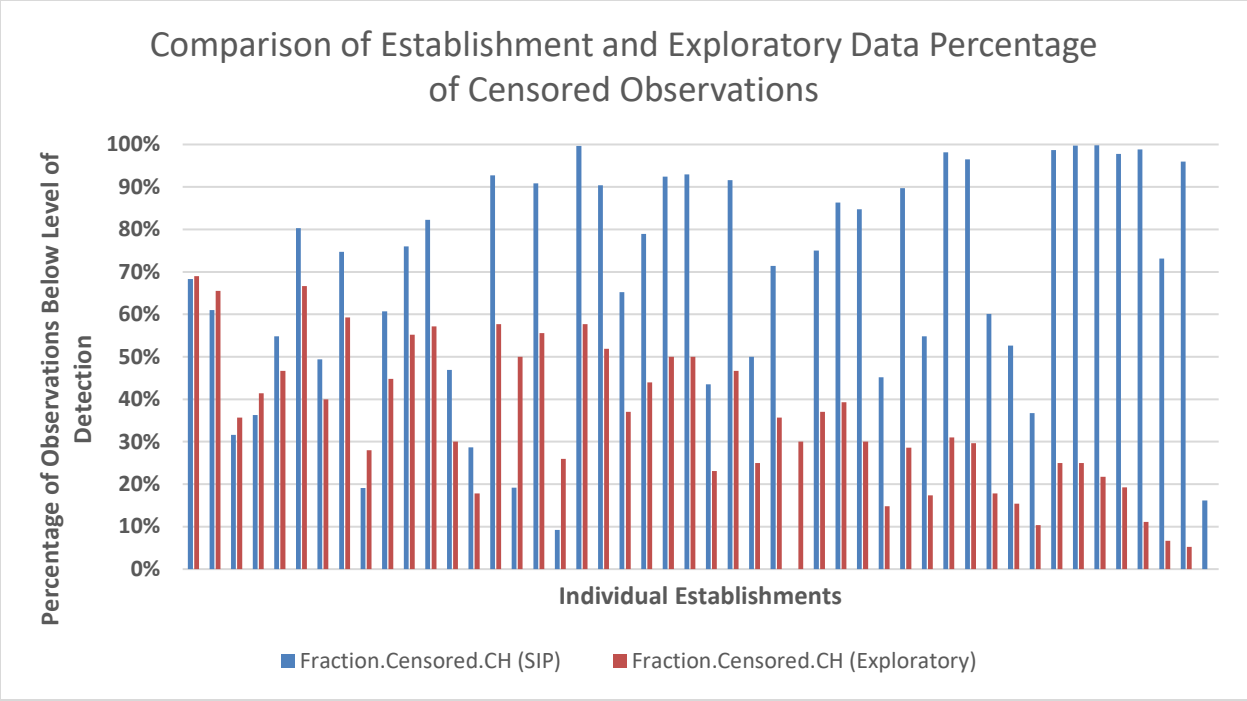


Figure 59: Percentages of censored AC observations at post-chill in Exploratory and SIP data by establishments.

Discussion and Conclusions:

The SIP data had limitations that prevented extensive analyses and comparison with FSIS's Exploratory data. The average log10 AC level at rehang for SIP data was slightly lower than the Exploratory data. However, SIP data had higher average *Salmonella* prevalence than Exploratory data at the rehang **(Table 54**). Both datasets had significantly lower average log10 AC levels and average *Salmonella* prevalence at post-chill compared to rehang. However, SIP data had improbably lower average log10 AC level at post-chill compared to Exploratory data **(Table 54)**. As a result, SIP data collected by industry could not be used for modeling AC level and *Salmonella* prevalence in this risk assessment (Williams, 2015).

While the results from this analysis identify differences between the 2 datasets, there are many possible explanations for these differences. For example, while substantial data is shared with FSIS through the SIP, specific details about the sampling itself, such as the identification of the specific point of sampling collection (e.g., rehang v. pre-chill) are often not included in the data sharing. The absence of this metadata makes it challenging to discern why these differences exist. It was beyond the scope of this analysis to fully elucidate why these differences exist. However, as a part of FSIS' Cooperative Agreement with the University of Maryland (FSIS-02152022), the Agency is working closely with industry partners to better understand these data gaps and develop data sharing criteria that would allow for a greater understanding of differences between industry and FSIS data, when they exist.