IMPROVING THE UTILITY OF *SALMONELLA* THERMAL INACTIVATION RESEARCH FOR THE VALIDATION OF LOW MOISTURE FOODS PREVENTIVE CONTROLS

By

Ian Michael Hildebrandt

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Biosystems Engineering – Doctor of Philosophy

2024

ABSTRACT

The modern U.S. food safety system relies on risk characterization, reduction, and management to provide safe food for consumers. Much of the responsibility for managing risks associated with foodborne illnesses is placed on the food industry, which must incorporate and validate preventive controls whenever a known hazard is reasonably likely to occur. However, because of the biological hazards involved, much of the information needed for a successful preventive control validation originates from independent laboratory-based research groups. The resulting practical challenges include: (1) Science-based evidence for pathogen inactivation developed in laboratory-scale environments often has critically different features from industry-scale applications; (2) Guidelines for predictive modeling- and surrogate-based validations often fail to address how researchers can improve the utility of their research for industrial application; and (3) There is no known standard for how robust a preventive control must be, beyond the expectation that it must be "statistically valid." Therefore, the overall goal of this dissertation was to improve the utility of food safety research for application in preventive control validations. The specific objectives were to: (1) Quantitatively evaluate current practices for development and application of predictive inactivation models for *Salmonella* in low-moisture foods and develop a framework for improved practices; and (2) Develop and demonstrate improved statistical tools for the application of *Salmonella* surrogates in low-moisture preventive control validations.

The first step was to define criteria for a robust "statistically valid" standard for evaluating preventive control validations and then evaluate how well current practices are meeting these criteria. Based on an extensive review of the literature, it was demonstrated that the current state-of-practice in predictive microbiology is not yielding models ready for use in

preventive controls. This was partially by design, as most instances of inactivation models in relevant literature fill a descriptive rather than predictive function. For models that were clearly intended to inform, or be utilized in, preventive control validations, there were persistent issues of underreporting key model components that would maximize the utility of such models. The issues limiting experimental utility across studies/labs were investigated via a multi-laboratory thermal inactivation study, and a standard template for future thermal inactivation studies was developed to improve research synergy. Subsequently, a widely disseminated, model-based approach to low-moisture food validations, which ignored several core principles of predictive microbiology, was critically evaluated for a baking case study. The results demonstrated significant "fail dangerous" errors resulting from inappropriate application of inactivation models that do not incorporate the critical effect of moisture on *Salmonella* thermal resistance in low-moisture foods.

Non-pathogenic surrogate organisms are commonly used for validating preventive controls in industry; however, much like predictive models, there is no consensus or standard for statistically evaluating a preventive control validation. Therefore, a statistical framework for evaluating reduction performance criteria was developed and tested, including a statistical foundation for translating surrogate-based validation results into likely pathogen outcomes. This framework was demonstrated in a case study encompassing a large dataset comprised of thermal process validations in the almond industry.

Overall, this work demonstrated key approaches to improve standards of practice for predictive microbiology and food safety surrogate research, which should improve utility for application to process validations. Additionally, this work demonstrated methods that may help industry improve the design, robustness, and costs of their preventive control validations.

ACKNOWLEDGEMENTS

This work and I stand on the shoulders of giants. I have been able to accomplish this PhD only through the assistance and guidance of so many of those around me. Words on a page cannot do justice to the gratitude I hold for everyone who helped me; however, it is the best I can do right now, so they will have to accept it.

I would like to thank each of my committee members for their support: Dr. Elliot Ryser, Dr. Kirk Dolan, Dr. Sanghyup Jeong, Dr. Jade Mitchell, Dr. Nathan Anderson, and most importantly, Dr. Bradley Marks. I want each of you to know that you were a role model for me up to now, and I will continue to hold each of you as a role model as I continue my journey. When I first arrived at MSU, I was truly a novice who was eager to learn. I was lucky to have Dr. Marks, Dr. Jeong, and Dr. Ryser as mentors from day one. Through them I was able to appreciate the complex boundary between food engineering and food microbiology that I find myself in today.

As I lost my green, I owed my continued development to Dr. Dolan, Dr. Mitchell, and Dr. Anderson. They helped me focus my attention on the critical roles that biology-based engineering can fill. At this point, I would like to acknowledge the guidance I received from Dr. Susanne Keller (I am glad she was finally able to retire) and Dr. Elizabeth Grasso-Kelley. While I was working with the FDA, they were there to challenge me and push me to be better.

There are a multitude of undergraduate, graduate, and graduated students that I would like to thank! I have been around MSU for a while and have had the opportunity to work with many people near and far. These people are one reason I am so proud to work in this field; because it attracts so many kind and interesting people. I fear to list anyone specific here out of

dread that I would forget someone essential. If you are reading this, and you are not one of my committee members, then know that I am probably talking about you.

I need to thank Nicole Hall and Mike James by name for making working at MSU so easy and enjoyable that it made me never want to leave our research team.

My research, beyond what was included in this dissertation, was funded by many different organizations. From the summer I was an EnSURE student, to my role as an ORISE research assistant, to my current position funded by a USDA SAS grant. I could only have done what I have done because I was paid to do it. I am not independently wealthy enough to do this for fun.

On a personal note, I want to thank libraries. Any institution that focuses on freely disseminating knowledge and encouraging personal growth deserves respect from everyone. As someone who had to rely heavily on the resources made available through MSU libraries to complete extensive literature reviews, I appreciate the lengths that the MSU library and online databases have gone through to improve the utilization of their systems.

And now I would like to extend my deepest debt of gratitude to my best friend and wife, Mattie Anne Cook-Hildebrandt. For some reason, she decided to date, marry, and have a child with a perpetual student. Hopefully I can live up to all the promises I made her about being married to a doctor. Regardless, Mattie has always been there supporting me. From the humblest beginnings to slightly less humble times now, and for all time, I know she will be there for me. I would like to acknowledge all members of my family, especially my father, Dr. John Hildebrandt, for pushing me to get involved in research so long ago, and my mother, Cheryl, for always being interested in what I was doing for research, although I am not certain she always

understood it. Finally, I would like to thank my newest little advocate, Beatrix Pearl Hildebrandt. She is only four months old, yet she gives me strength to push forward each day.

The work presented in this dissertation was supported by: (1) The U.S. Department of Agriculture, National Institute of Food and Agriculture, Coordinated Agricultural Projects award no. 2015-68003-23415, (2) The U.S. Department of Agriculture, Agriculture and Food Research Initiative, Sustainable Agricultural Systems Program grant no. 2020-68012-31822, and partially by (3) An Oak Ridge Institute for Science and Education (ORISE) fellowship administered by the Center for Food Safety and Applied Nutrition through an interagency agreement between the U.S. Department of Energy and the U.S. Food and Drug Administration.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

1.1 The Challenge

In 2012, Scharff *(207)* estimated that the annual burden of foodborne illnesses in the U.S. was approximately $77 billion, with *Salmonella* contributing the most illnesses and hospitalizations. *Salmonella* continues to be associated with various meat, poultry, and low-moisture food products *(232)*. The United States Food and Drug Administration (FDA) required, with the passage of the 2011 Food Safety Modernization Act (FSMA), the food industry to have "science-based minimum standards for conducting a hazard analysis, documenting hazards, implementing preventive controls, and documenting the implementation of the preventive controls." *(234)*. A preventive control should reduce the risk of illness from a foodborne pathogen by limiting pathogen propagation or by pathogen reduction. With the enactment of FSMA, food safety regulations shifted focus from reactive to preventive, requiring science-based preventive controls to reduce the risk of illness. While requiring further documentation and verification of preventive controls, the FDA did not provide any prescriptions for specific tools or methodologies. Because no standardized approach for preventive control research and validation existed prior to the FSMA, the body of knowledge required for the food industry to adhere to the FSMA regulations has been developed largely ad hoc.

Much of the science-based evidence around implementing and documenting effective preventive controls has been developed in academic and government laboratories. For example, processing guidelines for *Salmonella* inactivation in meats/poultry (United States Department of Agriculture Food Safety Inspection Service (USDA FSIS) "Appendix A") and almonds (Almond Board of California (ABC) "Guidelines for Using *Enterococcus faecium* NRRL B-2354 as a Surrogate Microorganism in Almond Process Validation" and "Guidelines for Validation of Dry

Roasting Processes") were developed based largely on laboratory-scale experiments, which were then adapted for industry-scale application *(121, 131)*. Despite the guidelines issued by the USDA FSIS and the ABC, no standardized protocols exist for the development or validation of preventive controls for pathogen inactivation. As a result, science-based evidence for pathogen inactivation developed in laboratory-scale environments often does not accurately reflect industry-scale applications. Furthermore, several studies have indicated that the use of different methodologies in laboratory-scale research has significantly impacted the results and subsequently weakened the utility of results combined from multiple studies *(108, 123, 239)*. Without proper diligence, researchers may produce unintentionally biased and conflicting studies, which then may misinform the validation of industrial preventive controls.

Published guidelines for the validation of preventive controls are scarce; however, there are important examples *(8, 13, 16, 43, 110)*. Despite the availability of these resources, the demand has not been met, and some validation guidelines have been stretched to applications beyond their original intent. As a result, validation protocols established for a specific product may be misappropriated for other products. For example, when ABC published "Guidelines for Using *Enterococcus faecium* NRRL B-2354 as a Surrogate Microorganism in Almond Process Validation," there was a collection of research that specifically examined *Salmonella* and *E. faecium* in almonds *(1, 13, 68, 102, 120, 121)*. As a result, other low-moisture industries attempted to use *E. faecium* as a surrogate for *Salmonella* without performing the necessary research for both organisms in the different products *(31)*. Subsequent research has suggested that *E. faecium* may be an appropriate surrogate for *Salmonella* in various low-moisture foods *(32, 44, 45, 71, 252)*; however, universal acceptability has not been established. This example demonstrates the need for process validation solutions, and the absence of standardized

approaches for translating information for a successful validation. As a result, the depth of knowledge varies widely by product and technology – with a substantial portion of work relying on research developed for other pathogens/products/technologies.

Process validation guidelines exist that address FSMA compliance *(35, 43)*, surrogate-based process validations *(110)*, general low-moisture foods *(8, 99)*, specific low-moisture foods *(16)*, and even technology-specific low-moisture foods *(9)*. These documents provide insight into information that should be considered prior to process validation and when composing a validation report; however, those guidelines generally do not provide sufficient information on developing a statistically-sound validation and analysis of results. For example, the guidelines range from failing to mention the terms statistic(s), error, and replication(s) *(99)* to emphasizing that statisticians should be included in the design and analysis of the process validation *(8, 35)*. None of the referenced validation guidance documents report standardized methodologies for performing or analyzing predictive model- and surrogate-based process validations.

1.2 Goal and Objectives

There is a food safety "knowledge market" in which research yields a supply of information and industry provides a demand for answers. Although the supply of specific research that improves predictive model- or surrogate-based validations may seem plenty, it does not provide sufficient utility for food processors to reliably and robustly solve their process validation problems. Therefore, the goal of this dissertation was to improve the utility of food safety research for improved application in predictive model- and surrogate-based preventive control validations for low-moisture foods. The specific objectives were to:

1. Quantitatively evaluate current practices for development and application of predictive inactivation models for *Salmonella* in low-moisture foods and develop a framework for improved practices.
2. Develop and demonstrate improved statistical tools for the application of *Salmonella* surrogates in low-moisture preventive control validations.

CHAPTER 2: REVIEW OF LITERATURE

Chapters 3 and 6 of this dissertation contain extensive reviews of literature specifically related to predictive inactivation modeling and surrogate-based validations, respectively. To limit the redundancy of information, the literature review in chapter 2 contains basic information for both topics as well as a more in-depth discussion of topics not extensively covered in chapter 3 or 6.

Despite the magnitude of scientific effort that goes into developing the foundation of process validations, there are no known standards for methodologies at the various stages of this research-to-application domain. Several studies have been published on data collection and regression techniques *(28, 59, 67, 84, 95, 123, 162)*; however, little is known about the impact of these resources have had on improving practices. The state-of-practice for developing and reporting predictive inactivation models varies widely, with sections of the field using unique methodologies. Assumptions about the relationship between pathogen and surrogates often rely on predictive models; therefore, the state-of-practice for surrogate utilization shares the same limitations.

2.1 Salmonella

One of the pathogenic members of the Enterobacteriaceae family, *Salmonella enterica* is a facultative anaerobe, oxidase-negative, catalase-positive, Gram-negative, rod-shaped bacteria *(58)*. *Salmonella* is most known for causing foodborne illness with gastroenteritis symptoms, though some *Salmonella* (such as serovar Typhi) have been known to cause systemic illness. The infectious dose of *Salmonella* is known to be impacted by several intrinsic and extrinsic factors, resulting in outbreaks with illness caused by only a few ingested cells *(58, 83)*.

In 2011, nontyphoidal *Salmonella* was estimated to be the most burdensome foodborne pathogen to the U.S. economy, with an estimated $4.4 billion cost *(207)*. This may be an underestimate of the economic burden, as it focused on estimating quality-adjusted life years and did not account for impact of food loss due to product recalls. Because *Salmonella* may persist in low-moisture foods for months to years in a desiccated state *(1, 34, 78, 107, 136, 236)*, and when considering how many low-moisture foods are common ingredients in other foods (e.g., spices, flour, nuts), there is a large, interconnected network of products that can be impacted by a single outbreak and/or recall.

From a food safety management perspective, *Salmonella* is often the target for reductions in low-moisture food preventive controls because it is often considered the most resistant pathogen of concern *(8)*. Therefore, preventive controls designed to sufficiently reduce the risk of *Salmonella* may also sufficiently reduce the potential risk of other foodborne pathogens (e.g., *Escherichia coli* and *Listeria monocytogenes*). *Salmonella* resistance has been shown to exponentially increase with decreasing water activity *(126, 213, 218)*, which may be at least partially attributable to *Salmonella* stress response to desiccation having a cross-protective effect for other lethal stressors (e.g., radiation, heat) *(83)*. The relative impact of temperature and water activity on *Salmonella* inactivation is also product dependent. For example, Jin et al. (2018) examined the impact of temperature and water activity on the inactivation of *Salmonella* in two model low-moisture foods (i.e., plain biscuit and cracker) composed of identical ingredients in different amounts *(126)*. Despite being formulated to identical water activities, the thermal resistance of *Salmonella* was often greater in one matrix than the other, depending on the specific temperature and water activity combination tested.

2.2 Model Forms

Within the field of predictive microbiology, Whiting and Buchanan *(266)* proposed the structural nomenclature of primary, secondary, and tertiary models. Primary models describe the microbial response to time within a single set of conditions. Most primary models interpret this as the change in microbial populations under constant conditions. Conversely, an end-point-type model could interpret "a single set of conditions" to encapsulate a predefined treatment unit that can include dynamic environmental factors. Secondary models describe the effects of environmental factors on the primary model parameters. For example, a parameter within the primary model may be described via a function, or secondary model, that accounts for changes in temperature. Due to the intimate connection between primary and secondary models, certain model forms are often associated with each other. The distinction of primary, secondary, and tertiary models is a unique practice for predictive microbiology and food science; therefore, literature supporting the development of predictive microbiology models is highly specialized. As a result, predictive microbiology models may be isolated from innovations or perspectives that advance other predictive fields.

Using the structure proposed by Whiting and Buchanan *(266)*, predictive inactivation models can be generically structured using Eqns. 2.1-2.3.

$$log\left(\frac{N}{N_o}\right) = f\left(t|\beta_{primary}\right) + \epsilon_{primary} \quad \text{Eqn. 2.1}$$

$$\beta_{primary} = f\left(X|\beta_{secondary}\right) + \epsilon_{secondary} \quad \text{Eqn. 2.2}$$

$$log\left(\frac{N}{N_o}\right) = f\left(t|f\left(X|\beta_{secondary}\right)\right) + \epsilon_{model} \quad \text{Eqn. 2.3}$$

Eqn. 2.1 represents the expression of primary models where the survivor ratio, the logarithmic transformation of the microbial concentration (N) at any time (t) divided by the microbial

concentration at time 0 (N_0), is described by a function of independent variable *t* and primary model parameters $\beta_{primary}$. The secondary model (Eqn. 2.2) describes $\beta_{primary}$ as a function of independent variables besides time (e.g., temperature, water activity, sublethal injury, pH, etc.). Each parameter in $\beta_{primary}$ may have a secondary model. The combination of primary and secondary models yields a complete model form (Eqn. 2.3), which then can be integrated over the independent variables to yield a mean predicted lethality of the pathogen of concern. Model residuals, $\varepsilon_{primary}$, $\varepsilon_{secondary}$, and ε_{model} typically are not reported; however, estimates of model error, such as Mean Square Error (MSE) or Root Mean Square Error (RMSE), have been reported for primary and complete model forms Eqn. 2.4.

$$MSE = \frac{1}{n-p}\sum \varepsilon^2 = \frac{1}{n-p}\sum \left(log\left(\frac{N}{N_0}\right)_{predicted} - log\left(\frac{N}{N_0}\right)_{observed} \right)^2 \quad \text{Eqn. 2.4}$$

Where *n* represents the number of data points used in the regression of the model with *p* parameters. Model error always depicts the variability of the model for the response variable, which means Eqn. 2.4 is not formulated accurately for secondary models. Instead, the predicted and observed values are primary model parameters, and large collections of model parameters (which are not typically included in predictive microbiology studies) would be needed to derive a practical interpretation of secondary model error. Analyses based solely on secondary models are more common in meta-analyses *(22, 239)*.

While models may be used to describe phenomena, the purpose of predictive inactivation models is to predict pathogen inactivation. Model regression results in parameters and error estimates, the quality of which may depend on regression methodologies *(108, 162)*, which then

can be used to predict the mean and spread of outcomes when supplied with the appropriate inputs. Both model uncertainty and variability are required to determine the range of possible outcomes from a model *(176)*.

Additional information on commonly utilized inactivation models can be found in chapter 3. Predictive microbiology may consist of many different forms of primary models *(146)*; many primary model forms are specific to growth models. When considering the state-of-practice, only a few inactivation model forms are commonly used: linear, log-linear, and Weibull.

2.3 Data Collection Requirements

The estimation of model parameters can be categorized as an inverse problem, where data are collected to estimate the parameters of a model describing the phenomena that generated the data collected. As a result, the quality of the estimated model depends on the quality and quantity of the data used in its estimation. Most literature on optimal experimental design focuses on improving model quality by minimizing model parameter variance and covariance, such as the *D*-optimal design *(28, 59, 84, 95)*. As a result, primary model parameters can rapidly converge with as few as four data points per replication if the data are collected at precisely the optimal times. For such approaches, a major requirement is that the model form describing the phenomenon is known and the relative value of the model parameters are known prior to data collection. Additionally, because of minimum data collection, there is little ability to accurately judge the quality of fit.

Other approaches to data collection go beyond determining how many data points are needed and focus on determining optimal conditions wherein data points should be collected. Scaled Sensitivity Coefficients (SSC) is an experimental design tool that is more generalizable

than the *D*-optimal design *(67)*. An SSC indicates the magnitude of change in the model response due to a perturbance in parameters. A plot including the model response and SSCs can then illustrate how perturbances in the parameters impact the predicted values during an experiment (Figure 2.1). If data are collected where the magnitude of the SSC response is large, then the model parameter is more easily discernable. Furthermore, if multiple parameters are included in

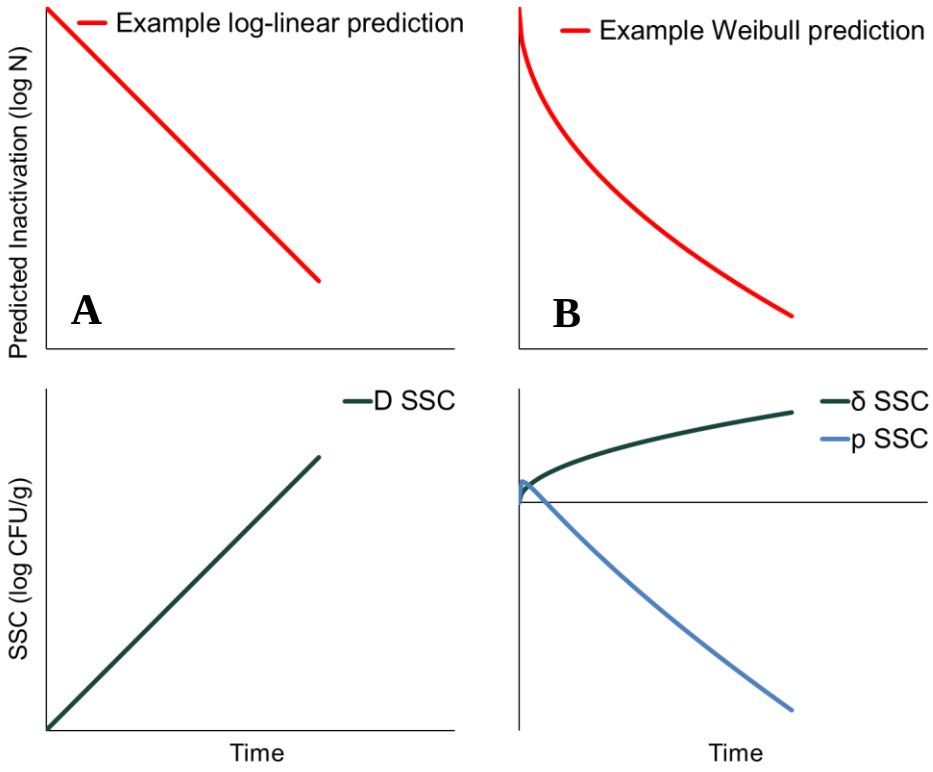


Figure 2.1 Side-by-side comparison of SSCs for log-linear (A) and Modified Weibull models (B) for isothermal conditions. See chapter 3 Eqn. 3.1 and Eqn. 3.3 for log-linear and Weibull model equations, respectively.

the SSC analysis, any correlation between SSCs can reduce the ability of data collected in the correlated regions to enable estimation of either parameter. While assessing SSC plots is more qualitative than methodologies that optimize data collection by minimizing variance and

covariance of model parameters, the SSC analysis has the benefit of not requiring the exact parameters to be known.

Independent of the model form, there are practical considerations to data collection. Common limitations involve accounting for temperature, selection of relevant treatment conditions, and accounting for high/low background and/or inoculated population levels. Isothermal inactivation experiments must consider the impossibility of having perfectly constant-temperature conditions. These experiments often exclude the come-up-time (time needed for samples to reach treatment temperature) to collect data during the pathogen's response to a constant temperature. Some isothermal experiments correct the data to account for the lethality accumulated during the come-up-time *(23)*, whereas studies focusing on end-point-type models often ignore come-up-time *(68, 102, 202)*. Additionally, when developing secondary models, both isothermal and dynamic-temperature experiments must consider the applicability domain of the secondary model parameters. Finally, a similar application domain principle can be applied to the lethality achieved during an experiment. A model developed using lethality estimates that are similar to or exceed the desired pathogen reduction would limit extrapolation errors. Previous research has demonstrated that pathogen inactivation kinetics are initial-population independent and not impacted by artificially high inoculation levels *(109)*. However, model regression is significantly impacted by the presence of low-plate counts, due to limits-of-detection *(90)*.

2.4 Model Regression and Reporting Error

With the structure of models in predictive microbiology typically following the classification of primary and secondary models *(266)*, regression strategies are often either one- or two-step procedures. A two-step regression involves a series of primary model regressions from which the model parameters are used as data in the secondary model regression. A one-step

regression estimates all parameters in the complete model form, simultaneously using all the data points *(66, 123)*. Two-step regressions are limited to studies with constant conditions, as model estimation using dynamic profiles require simultaneous estimation of primary and secondary model parameters *(195)*. Most reported parameter estimation methodologies include a two-step regression; however, this is not the best regression strategy *(94, 108, 123, 162)*. Studies comparing one-step and two-step regressions reported larger RMSEs, larger parameter uncertainty, and poorer measures of goodness-of-fit for models resulting from a two-step regression. Overall, the two-step regression was accompanied by inferior prediction power. Two-step regressions were identified as adequate for preliminary analysis, but were insufficient for complete model characterization, due to an overemphasis on fitting primary model parameters rather than representing the data as a whole *(108, 123)*. When examining the impact of one and two-step regressions on the joint confidence range for parameters, Silva et al. *(94)* reported that the parameter estimate regions for the two regression strategies were mutually exclusive.

Reproducibility of resultant model parameters is not well defined, partially because reproducibility of data is likely pathogen- and product-dependent *(185)*. Pflug *(197)* examined the reproducibility of *D*-values for bacterial spores measured using a Biological Indicator Evaluator Resistometer (BIER) device across twelve laboratories. While this is a rare example of experimental reproducibility quantification, this study was industry-driven and limited in statistical merit. For example, that study performed no statistical comparisons between *D*-values and tried to attribute all nominal variations in *D*-values to difference in treatment temperature. Literature exists that seeks to estimate the biological variability in the response parameters for predictive models *(22, 92, 93, 239)*, which assumes that data collected using identical methodologies are reproducible and the resulting error is due to biological variability. However,

theoretically similar, but perhaps dissimilar in practice, experimental methods have been demonstrated to not produce similar parameter estimates *(108, 239)*, making it unknown how likely results from two independent research groups are comparable.

2.5 Model Selection

Chapter 3 will detail common practice for model selection in predictive microbiology, and this section will go into theory for several practices.

While selecting models, the model with the better goodness-of-fit is typically the more correct model for the data; however, with nonlinear regression, the model selection process is more complex *(176)*. A model with more parameters can account for variability in the data and will more likely produce a better goodness-of-fit; however, more parameters introduce more uncertainty, which can reduce the predictive ability of the model.

Model selection based on a single statistical criterion or outcome from a model validation often considers the predictive ability of the model to distinguish the more-correct model. Model selection criteria reported in literature include the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), F-test, t-test, R^2_{adj}, and likelihood ratio test *(4, 132, 230, 238, 241)*. These model selection criteria are not interchangeable, and some consideration is needed when identifying appropriate model selection criteria *(176)*. Metrics strictly evaluating the goodness-of-fit, such as the adjusted coefficient of determination (R^2_{adj}), are considered an inappropriate statistical selection criterion for nonlinear models *(156)*. Model selection tests such as AIC or BIC are intended to compare un-nested model forms *(176)*. The F-test, t-test, and likelihood ratio test can be used to compare nested models or eliminate insignificant parameters from linear models *(176)*.

Criteria like the t-test analyze model uncertainty by performing significance testing on model parameters. The log-linear model is nested within the Weibull model when the shape parameter is equal to 1; therefore, a t-test comparison confirming that the shape parameter is not equal to 1 suggests that the Weibull model is the more correct model *(176)*. Most model selection criteria (AIC, BIC, f-test, likelihood ratio) emphasize evaluation of variability, often through the analysis of residuals, in order to support one model over another *(176)*. AIC and BIC compensate for model uncertainty by penalizing models with more parameters, because each additional parameter reduces the statistical power to estimate parameters, therefore increasing uncertainty *(176)*. This is not a perfect approximation of model uncertainty, because the selection penalty is dependent on the number of parameters and not the actual measures of uncertainty.

Model selection criteria typically evaluate the likelihood that the model fits accurately to the development dataset; however, it may be better to select predictive models based on accuracy of the model applied to a different dataset (i.e., predictivity). The ultimate test of a model's predictivity and robustness, which incorporates both model uncertainty and variability, would be a model validation against independent data. Hildebrandt et al. *(108)* reported that the log-linear model was more robust than the modified Weibull model when predicting the lethality of *Salmonella* in ground beef, despite model selection criteria identifying the modified Weibull model as the more likely correct model. The utility of a predictive model in the intended environment should be considered when selecting a model for preventive control validations; such utility can be informed by a statistical test and verified by a model validation. Furthermore, the objectivity of a model selection test is dependent upon whether data collection practices allowed for accurate estimation of each model independently.

2.6 Model Validation

McClure et al. *(166)* defined two components of a model validation as the "model validation," which applies to verifying that the model is mathematically appropriate and makes biological sense, and the "food validation," which applies to verifying that the model is appropriate for predictions in real foods. Despite the scope of the McClure et al. *(166)* review, information on the methodologies and metrics to consider when validating a predictive model is limited. For example, expected levels of standard error were listed as "0.1-0.3 log CFU/mL" for the primary model and "15-20%" for secondary models. With no definition of standard error provided, it was unclear whether the standard error for primary and secondary models was referring to the same type of metric. Additionally, model validations were suggested to include metrics of goodness-of-fit, testing for parsimony, robustness, and an assessment of the distribution of residuals, all with limited information about how to quantify these metrics. For "product validation," even less information was provided. Other reviews of predictive models reported limited information on model validations *(117, 167)*.

Other fields of research may use other terms to define different perspectives on model validation; one promising classification system considers the use of internal and external validations *(219, 220)*. Within the context of predictive microbiology, internal validations would be a model validation against a dataset with similar-to-identical data collection and experimental context. Examples could include various forms of cross validations *(29, 120, 198)* or validations of isothermally-derived models against dynamic temperature profiles for the identical microorganism and product *(111, 112, 227, 238)*. Internal validations may be seen as a robust goodness-of-fit metric as it verifies that the model fit is appropriate for describing microbial kinetics in the environment used to derive the model. In contrast, external model validations seek

to challenge the validity of the model when applied to "external" datasets, or those that are sufficiently different from the model development dataset. For models developed in laboratories with the intention of being applied in model-based preventive control validations, an external validation would best evaluate the validity of model for that use.

While no standardized methodology for predictive microbiology model validation has yet to be adopted by any scientific or regulatory organization, some validation metrics have been utilized in the literature. Ross *(205)* identified the lack of any standardized validation methodologies, and therefore developed two model-validation metrics coined accuracy (Eqn. 2.5) and bias factors (Eqn. 2.6)

$$parameter\ accuracy\ factor = 10^{\sum\left|\log\left(\frac{\beta_{predicted}}{\beta_{observed}}\right)\right|/n} \quad \text{Eqn. 2.5}$$

$$parameter\ bias\ factor = 10^{\sum\log\left(\frac{\beta_{predicted}}{\beta_{observed}}\right)/n} \quad \text{Eqn. 2.6}$$

where β represents the parameters being compared, and n is the number of observations used in the calculation. While originally developed for comparing the differences in growth time parameters, these metrics have been adapted for thermal inactivation validations where relevant parameters were estimated *(23, 79, 86)*. These metrics are limited to laboratory-scale model validations, where inactivation primary model parameters can be estimated. Additionally, because the formulation of these metrics only evaluates differences in parameters, no information about data variability or reproducibility can be captured. Other studies have adapted accuracy and bias factors to reflect microbial data *(72, 152, 231)*, where few have perhaps misused these factors as goodness-of-fit metrics (Eqns. 2.7 and 2.8) *(152, 231)*.

$$residual\ accuracy\ factor = 10^{\Sigma\left|\log\left(\frac{log\frac{N}{N_0}_{predicted}}{log\frac{N}{N_0}_{observed}}\right)\right|/n} \quad \text{Eqn. 2.7}$$

$$residual\ bias\ factor = 10^{\Sigma\log\left(\frac{log\frac{N}{N_0}_{predicted}}{log\frac{N}{N_0}_{observed}}\right)/n} \quad \text{Eqn. 2.8}$$

Several approaches have been published to characterize differences in data variability between prediction and observation. Common phrases incorporated included "fail-safe" or "fail-dangerous," where lethality was sufficiently underpredicted or overpredicted, respectively *(33, 37, 79, 86, 97, 163, 191, 230)*. Some studies use these terms to designate residuals as either positive or negative. Two studies *(79, 191)* reported that residuals overpredicting lethality by 0.5 log CFU/g or underpredicting lethality by 1.0 log CFU/g were fail-dangerous or fail-safe, respectively. However, these thresholds of variation are arbitrarily predetermined prior to experimentation and before the authors understand the variability inherent to the process or the microorganism. Terminology like fail-dangerous quickly communicates the outcome of a predictive model; however, any reported threshold should consider the error inherent to the pathogen, product, process, and scale before being used. Some studies qualitatively compare RMSE or mean residual bias *(79, 112, 121, 215)*; however, these studies also fail to provide sufficient context as to what would be acceptable in the given situations.

Overall, literature on predictive inactivation model validation is relatively scarce. When developing models, two key metrics for defining the robustness of inactivation models were confidence and prediction intervals *(66, 67)*; however, neither of these metrics are incorporated as validation metrics in any known predictive model reviews *(117, 166, 167)*. In 1999, McDonald and Sun *(167)* stated that "validation is often described as an ill-defined aspect of predictive food microbiology" and "to date no standard methods for model validations have been

published." After twenty years, standardized methodologies for model validations remain a critical gap in the field.

2.7 Preventive Control Validations

The driving force for industry-scale validation of *Salmonella* inactivation in low-moisture foods is the 2011 FSMA *(234)*, where validation was defined as the "element of verification focused on collecting and evaluating scientific and technical information to determine whether the food safety plan, when properly implemented, will effectively control the identified hazards." Besides providing a regulatory requirement for science-based preventive controls and validations, no standards or guidance for what constitutes sufficient scientific evidence was provided by the FDA. In a FSMA supporting document, Brackett et al. *(35)* summarized validation as "doing the right thing (using sound science) to control the hazard" and concedes that "industry lacks consensus and consistency in what control measures require validation and how the validations are conducted." This supporting document included approaches that have been successful, such as applying predictive models or surrogates, but did not go into further detail.

Other validation guidelines and supporting documents do not sufficiently characterize statistical criteria needed for predictive model-based process validations *(8, 9, 16, 98, 99)*. Validation guidelines often include validation-critical information, such as technology-dependent considerations, how to gather accurate temperature profiles, and how to monitor and collect a_w samples. However, specifics on interpreting predictive modeling results are scarce and often merely suggest including a skilled statisticians in the validation process *(8, 9)*. Even these highly specialized guidance documents for industry-scale process validations, with several specifically

for low-moisture foods *(8, 9, 16)*, do not sufficiently characterize statistically based validation methodologies.

2.8 Surrogate-based Validation

In the 2017 Hu and Gurtler review "Selection of Surrogate Bacteria for Use in Food Safety Challenge Studies: A Review" the desired characteristics of an appropriate surrogate were defined as *(110)*:

> "(i) nonpathogenic, (ii) behavior, inactivation characteristics, and kinetics that can be used to predict those of the target microorganism when exposed to similar or specific environmental conditions, (iii) simple to prepare, (iv) genetically stable, and (v) susceptibility to injury similar to that of the target pathogen."

Busta et al. *(39)* also listed criteria to consider that are relevant outside the scope of an surrogate-based validation of a preventive control (e.g. sanitation controls):

> "stable and consistent growth characteristics, easily prepared to yield high-density populations, ... population is constant until utilized, easily enumerated using rapid, sensitive, inexpensive detection systems, easily differentiated from other microbiota, attachment characteristics…mimic those of the target [pathogen], will not establish itself as a 'spoilage' organism"

Both reviews provide excellent overviews of important factors to consider when selecting a surrogate; however, neither address some important concerns of surrogate-based process validations, which are discussed exhaustively in chapter 6; please refer to that chapter for more information on surrogate validations.

Enterococcus faecium NRRL B-2354 is the most commonly utilized surrogate for *Salmonella* in low-moisture foods. Related *Enterococcus faecium* have been used in meat

fermentation, cheese making, and as a probiotic, while other strains have been identified as a BSL-2 microorganism *(65, 138)*. The strain specifically propagated as a surrogate (*Enterococcus faecium* NRRL B-2354) was determined not to pose a significant health risk if continued to be used in surrogate-based preventive control validations *(138)*. Since being considered acceptable for use as a surrogate for *Salmonella* PT 30 on almonds *(13, 120)*, numerous other studies have confirmed the suitability of *Enterococcus faecium* NRRL B-2354 as a surrogate for *Salmonella* in a large variety of low-moisture foods *(65)*.

2.9 Summary

Both predictive inactivation models and surrogates, which are essentially living inactivation "models" of the pathogen, require extensive research and due diligence before being considered viable for use in preventive control validations. Much of the foundational knowledge about models and surrogates originates from researchers developing the data and tools, but not necessarily with a direct connection to the industry using the tools for validating preventive controls. The food industry has limited capacity to independently develop or verify the appropriateness of predictive models or surrogates for use in validations, despite bearing the burden for ensuring sufficiency of pathogen control processes. Therefore, research meant to inform or be applied directly in preventive control validations must maintain high standards of thoroughness and utility to meet industry needs.

CHAPTER 3: REVIEW OF STATE-OF-PRACTICE FOR PREDICTIVE INACTIVATION MODELS

This review evaluated 100 studies published from 2000–2020 for common predictive inactivation modeling practices. The objective was to identify gaps in practice and propose standards to improve the utility of such models for validating preventive controls.

3.1 Review Parameters

3.1.1 Identification of Literature and Good Modeling Practices

To conduct an in-depth review and analysis of the state-of-practice of predictive microbial inactivation modeling, particularly in the field of food microbiology and safety, 100 studies were identified that reported predictive inactivation or survival (passive inactivation) models from 2000 to 2020. The 100 studies were selected from two searches, both using the Web of Science database (Clarivate, PA, USA). The first ~60 studies were identified using search criteria consisting of a pathogen of interest (*Salmonella*, *E. coli*, *Listeria*) and predictive thermal inactivation keywords (heat, inactivation model, thermal resistance, thermal inactivation) with the purpose of creating a collection representing a variety of products and microorganisms. The remaining set of ~40 studies were collected to expand the representation of the full collection to be more representative of general inactivation model use, by including years not well represented in the original study set, new processing technologies (e.g., radiation, extrusion, radiofrequency), more examples of model validation, and studies investigating pathogen survival. Overall, the final collection of studies was selected to be representative of inactivation models from 2000 to 2020, with a focus on thermal processing technologies and more recent literature (2016-2020).

With an objective to review the state-of-practice for predictive inactivation models, there were many considerations as to what constituted a good modeling practice *(66, 169, 205, 206,*

265). These considerations included specific recommendations for experimental methods for data collection, model regression, and model validation. Additionally, the Organization for Economic Co-operation and Development (OECD; Paris, France) proposed principles of quantitative structure-activity relationship models *(188)*. A modified version of the OECD principles is presented here as a set of good predictive microbiology modeling principles:

1. A defined endpoint (i.e., purpose of the predictive model)
2. Unambiguous methodologies and model
3. Appropriate measures of model estimates, uncertainty, and variability
4. Appropriate measures of goodness-of-fit, model selection, and predictive robustness
5. A defined domain of applicability for the predictive model
6. A mechanistic interpretation of model kinetics and observed differences, if possible

These six principles will be referred to as the Good Standards of Practice for Predictive Modeling (GSPPM). Studies were reviewed for information that informed adherence to each of the principles. A defined endpoint required studies to identify the purpose of the model (e.g., model describing *Salmonella* reduction in wheat flour); because all models reviewed were reported as part of the peer-reviewed literature, this should be a default principle. While a defined endpoint was considered either present or absent, information was extracted from each study that could inform adherence to the other principles. Because current predictive models are considered largely empirical and/or mechanistically informed *(146, 206)*, GSPPM principle 6 was not evaluated in this review.

GSPPM principle 2 requiring unambiguous methods focused on experimental methods known to impact resistance to the mode of inactivation prior to model regression *(57, 109, 149)*. While not exhaustive to every known impact, this included product(s), microorganism(s), use of

a single strain or cocktails, strain(s) grown on agar or broth immediately prior to inoculum formation, and experiment type (isothermal, dynamic thermal treatment, radiation, chemical). Adhering to this principle also required unambiguous model design and regression methods.

3.1.2 Predictive Model Regression and Completeness

Following the common classification scheme for predictive microbiology models, primary and secondary model components were identified *(146, 265)*. Brief descriptions of commonly reported models are included below. Each primary model describes the impact of time (t) on the log transformed survival ratio of the microorganism (S) under static conditions and with ε residual error. The log-linear model (Eqn. 3.1) represents a constant rate of inactivation, where D is the decimal reduction time, or the time needed to reduce the population by 90%.

$$S = \frac{-t}{D} + \varepsilon \qquad \text{Eqn. 3.1}$$

The Weibull model (Eqns. 3.2 & 3.3) has two forms that are commonly included in predictive microbiology literature *(160, 194, 240)*.

$$S = -bt^n + \varepsilon \qquad \text{Eqn. 3.2}$$

$$S = -\left(\frac{t}{\delta}\right)^p + \varepsilon \qquad \text{Eqn. 3.3}$$

The Eqn. 3.2 Weibull model form is designed after the cumulative distribution function for the namesake Weibull distribution. In this simplified form, b and n are scale and shape factors. van Boekel (2002) and Mafart et al. (2002) proposed a revised version of the Weibull model (Eqn. 3.3), where δ is defined as the time to first decimal reduction, and p is the shape parameter that modifies the nominal survival ratio (t / δ). Further discussion of Weibull models will refer to them as Eqn. 3.2 or 3.3 forms.

Another category of potential primary model forms includes those with multiple phases, such as the Gompertz, Geeraerd, Cerf (biphasic), and Baranyi models. Several of these models are repurposed growth models that are utilized for their ability to describe shoulders, linear, and tailing phases.

Models developed from linear regressions (such as the multiple linear model in Eqn. 3.4) were utilized as primary and secondary models. As a purely empirical model form, the independent variables (X_i) were scaled by coefficient parameters (β_i). Linear regressions included simple linear, polynomial, and response surface models (RSM).

$$S = \beta_0 + \sum \beta_i X_i + \varepsilon \qquad \text{Eqn. 3.4}$$

Secondary models, or models that describe the impact of environmental factors on primary model parameters, are typically closely associated with specific primary model forms. The Bigelow-type secondary model is commonly used with the loglinear primary model (Eqn. 3.5) but has also been used with Eqn. 3.3 Weibull models.

$$D = D_{ref} \times 10^{\frac{T_{ref}-T}{z}} \qquad \text{Eqn. 3.5}$$

The Bigelow-type secondary model shown here describes a *D*-value as a function of temperature (*T*), using a reference parameter (D_{ref}) and a *z*-value, which is the change in temperature needed for a 10-fold change in the *D*-value. The reference temperature (T_{ref}) is the temperature where the primary model *D* equals D_{ref}, and may be included as a parameter optimized to minimize parameter covariance but is usually pre-determined. Other studies have expanded on this type of secondary model by adding *z*-values for other factors such as pH, moisture, and water activity *(91, 119, 173, 213)*. Models represented as the combination of primary and secondary model forms were referred to as combined models.

Information on model regression methodologies and subsequent estimates of model parameters, uncertainty, and variability was collected to inform the completeness and quality of the models (GSPPM principles 2 and 3). Regression methodologies were examined to determine whether they involved linear or nonlinear regression, linearization of nonlinear models for parameter estimation, and, when secondary models were involved, whether a one-step or two-step approach was used. Except for linear models (Eqn. 3.4), all inactivation models that have been discussed are nonlinear; therefore, a nonlinear regression method was considered appropriate. When it was not clear whether nonlinear, linear, or linear approximation regression strategies were used, the software utilized in the model regression (if reported) was examined. If the software could perform the appropriate regression strategy, then it was assumed that was the case.

For a predictive model to be considered complete (GSPPM principle 3), the presence of parameter estimates, parameter error estimates, and regression/model error (e.g., Mean Square Error (MSE) or Root MSE (RMSE)) was checked. RMSE is a measure of the expected difference between observed data and their predicted values *(168)*, and in cases of normally distributed residuals RMSE represents the standard deviation of residuals. For models to be able to describe the potential spread of data, RMSE was needed in the same units as the microbial reductions (i.e., log reductions or log survival ratio) for the final predictive model. This would be for the combined model in cases where a secondary model is used. RMSE for secondary models would inform end users of residual parameter errors, which would impact the variability of the mean prediction if incorporated into a predictive application. Failure to report primary and secondary model parameters and error estimates was not considered inappropriate (and not annotated) for two circumstances. First, if multiple model forms were ranked via a model

selection process, then model estimates were not sought for those with demonstratively poorer fit. Second, when it was clear that the combined model that solely utilizes the secondary model parameters was the objective of the study, the presence of primary model estimates was not needed.

Practices for evaluating the appropriateness of the model fit were investigated by examining model selection methods (when applicable) and metrics for goodness-of-fit (GSPPM principle 4). Model selection methods were only applicable when studies were considering multiple forms of primary or secondary models, or when model parameters could be reduced through significance testing (i.e., linear models). Any goodness-of-fit metrics explicitly used in the studies were documented. The coefficient of determination (R^2) did not need to be explicitly qualified to be recorded as a goodness-of-fit metric. Predictive robustness, or the ability of a predictive model to maintain accurate predictions despite the quality of or reasonable adjustments made to the model inputs, may be demonstrated with either an internal or external model validation. This is not to be confused with robustness of the model fit, which relates to the ability of a model regression to consistently and accurately converge to a set of model parameters and is impacted by data collection, regression methods, and the model form. Internal validation (AKA cross-validation) refers to validating the model against a dataset generated using the same experimental parameters as the model development dataset. External validation refers to validating the model against a dataset generated using dissimilar experimental parameters (e.g., similar product, different process scale). When applicable, information on model validation design and evaluation metrics was collected.

A defined applicability domain (GSPPM principle 5) was required to qualify the suitability for a model to be used in a predictive capacity. This required the implication that a

model is suitable for predictive applications when the model was developed with relevant parameters (e.g., describes the impact of temperature and moisture for pathogen inactivation in a baked good). For most applications that require dynamic inputs, this would require the use of a secondary or combined model; however, in applications where the mode of inactivation can be delivered at a constant rate (i.e., radiation or some chemical treatments) then a primary model may be sufficient. Additionally, studies were required to explicitly qualify the potential uses or limitations for the model; therefore, unqualified predictive models were defined as those that did not benefit from the author's perspectives or expertise on the utility of the model. Consideration of model completeness (GSPPM principle 1-3) and predictivity (GSPPM principle 4-5) resulted in three final classifications for models: (1) predictive, (2) unqualified predictive, and (3) non-predictive models.

3.2 Results and Discussion

3.2.1 General Study Characteristics

Of the 100 studies collected, the most-represented characteristics included studies from 2016-2020 (n=40), and those encompassing low-moisture foods (LMF; n=56), *Salmonella* (n=80), pathogen cocktails (n=53), a broth-based culturing method immediately before inoculum formation (n=55), and a heat treatment (n=94) (Table 3.1). The full list of the 100 reviewed studies and a breakdown of general study and modeling characteristics can be found in the Appendix A and B. From the two-phase selection process, each year was represented by at least two studies. Additionally, the second iteration of study selection focused on improving representation of non-*Salmonella* studies and "other" product categories, which included eggs, fresh produce, and high-moisture manufactured foods *(29, 55, 70, 87-89, 174, 238)*. While studies developing inactivation models were most often isothermal heat treatments describing

Salmonella inactivation with the log-linear model for each block of time included in this review, there were some noticeable shifts over time. Over the years included in this review, there was a shift from more meat-based studies to LMF studies, which likely also contributed to the relative increase in agar-based cultivation methods and inclusion of "other" microorganisms, which were often non-pathogenic surrogate organisms.

Table 3.1 General study demographics for the 100 reviewed inactivation modeling studies

Years	Products	Microorganisms	Culturing Method	Experiment Type**	Primary Models
2000-2005 (n=18)	Meat (n=14)	*Salmonella* (n=17)	Broth (n=18)	Isothermal (n=16)	Log-linear (n=12)
	LMF (n=0)	*E. coli* (n=3)	Agar (n=0)	Dynamic Temperature (n=4)	Weibull (Eqn. 2 / Eqn. 3) (n=2/0)
	Lab (n=4)	*L. monocytogenes* (n=4)	Cocktail (n=13)	Radiation/Light (n=0)	Multi-Phased (n=1)
	Other (n=1)	Other* (n=3)	Single Strain* (n=10)	Chemical (n=0)	Linear (n=2)
				Passive Survival (n=0)	Novel/Other (n=4)
2006-2010 (n=22)	Meat (n=7)	*Salmonella* (n=15)	Broth (n=19)	Isothermal (n=17)	Log-linear (n=12)
	LMF (n=4)	*E. coli* (n=5)	Agar (n=3)	Dynamic Temperature (n=6)	Weibull (Eqn. 2 / Eqn. 3) (n=4/4)
	Lab (n=6)	*L. monocytogenes* (n=6)	Cocktail (n=9)	Radiation/Light (n=2)	Multi-Phased (n=3)
	Other (n=7)	Other (n=3)	Single Strain (n=13)	Chemical (n=2)	Linear (n=2)
				Passive Survival (n=0)	Novel/Other (n=2)
2011-2015 (n=19)	Meat (n=2)	*Salmonella* (n=15)	Broth (n=12)	Isothermal (n=11)	Log-linear (n=10)
	LMF (n=14)	*E. coli* (n=4)	Agar (n=7)	Dynamic Temperature (n=5)	Weibull (Eqn. 2 / Eqn. 3) (n=3/6)
	Lab (n=3)	*L. monocytogenes* (n=3)	Cocktail (n=11)	Radiation/Light (n=1)	Multi-Phased (n=5)
	Other (n=2)	Other (n=6)	Single Strain (n=12)	Chemical (n=1)	Linear (n=3)
				Passive Survival (n=3)	Novel/Other (n=0)
2016-2020 (n=41)	Meat (n=2)	*Salmonella* (n=33)	Broth (n=7)	Isothermal (n=34)	Log-linear (n=36)
	LMF (n=38)	*E. coli* (n=5)	Agar (n=35)	Dynamic Temperature (n=8)	Weibull (Eqn. 2 / Eqn. 3) (n=0/15)
	Lab (n=2)	*L. monocytogenes* (n=4)	Cocktail (n=20)	Radiation/Light (n=1)	Multi-Phased (n=1)
	Other (n=0)	Other (n=12)	Single Strain (n=25)	Chemical (n=0)	Linear (n=4)
				Passive Survival (n=1)	Novel/Other (n=0)
2000-2020 (n=100)	Meat (n=25)	*Salmonella* (n=80)	Broth (n=56)	Isothermal (n=78)	Log-linear (n=70)
	LMF (n=56)	*E. coli* (n=17)	Agar (n=45)	Dynamic Temperature (n=23)	Weibull (Eqn. 2 / Eqn. 3) (n=9/25)
	Lab (n=15)	*L. monocytogenes* (n=17)	Cocktail (n=53)	Radiation/Light (n=4)	Multi-Phased (n=10)
	Other (n=10)	Other (n=24)	Single Strain (n=60)	Chemical (n=3)	Linear (n=11)
				Passive Survival (n=4)	Novel/Other (n=6)

*Most "other" microorganisms were considered to be surrogates and were exclusively used as single strains within the 100 reviewed studies

**Experiment type listed refers to portion of the study used for model development or to which the model was applied

The inclusion of different model types was correlated with other features of the studies. Studies that included only the Gompertz model all originated from the same research group *(172-174)*. Other than those three studies, studies that included three or four primary model forms accounted for most of the uses of multiphase models (Baranyi, Gompertz, Geeraerd, Cerf) and also were most of the studies modeling microbial survival *(34, 72, 137)*. The most common usage of primary models was log-linear only (n=46), log-linear and Weibull (n=20), Eqn. 3.2 Weibull only (n=6), linear/RSM only (n=6), and Eqn. 3.3 Weibull only (n=5). The Eqn. 3.3 Weibull model was widely introduced to predictive microbiology in 2002 *(160, 240)*; however, these studies or others using this form between 2000-2005 were not captured in the selection process for this review. Over time, Eqn. 3.3 Weibull became the more utilized version in studies.

Most studies reviewed included a secondary model (n=66). The most common secondary model was a Bigelow-type model (n=52), which was commonly paired with log-linear primary models (n=47). Linear equations were the second most common secondary model (n=12) and were used in combination with various primary model forms.

All studies were direct about why an inactivation model was utilized (GSPPM principle 1). When considering the experimental procedures generating data, all categories of information searched for were observable (GSPPM principle 2).

3.2.2 Model Regression

The utilization of regression methodologies had a significant impact on whether studies reported the necessary information for complete and predictive models. To demonstrate an issue that was found to be pervasive in the reviewed literature, we will first discuss model regression practices and their impact on the quality, or even presence, of model estimates (parameters, uncertainty, error).

Except for models based on linear equations, all predictive models within the 100 studies were nonlinear and should incorporate nonlinear regression techniques to optimize parameters and yield accurate error estimates. When examining studies with only the log-linear model without secondary models (n=11), a majority (n=6) estimated the *D*-value through a linearization of the model (i.e., used the inverse of the slope). For studies with more than one primary model, with non-log-linear models, or with actual linear models, indication was provided that the preferred regression methodology was used in each case. This is partially based on the assumption that a predictive modeling tool was used appropriately. Linear approximations and nonlinear regressions may yield similar *D*-value estimates *(108)*; however, the estimates of standard error are often lost because the error estimates for the slope (inverse of D) are not translated to error estimates for the reported *D*-value.

Most studies with a secondary model utilized a two-step regression process (n=54), a process that pairs easily with isothermal studies and is easier to perform than a one-step regression analysis. The immediate impact of using a two-step regression analysis is a sub-optimal estimate of parameter and model errors when compared to a one-step regression analysis *(108, 162)*. This is partially because a two-step regression uses the original microbial population dataset to estimate the primary model parameters, which then comprise the secondary dataset for secondary model regression. Studies that utilized a one-step regression of the combined model form (n=11) simultaneously estimate model parameters and errors, which is generally a necessity for studies with dynamic conditions *(97, 120, 174, 230)*. The model resulting from a one-step regression therefore has more biologically meaningful interpretations of parameter and error estimates, which can be lost during a two-step regression. The loss of information is further compounded when considering that studies that used a two-step regression (n=44) often used a

linearization of the log-linear and/or Bigelow-type models at some point in the parameter estimation process (n=37).

While most studies were sufficient with the descriptions of regression methods (GSPPM principle 2), a portion of the studies were not transparent with regression methods (n=9). Another reporting issue resulting from regression methods revealed a connection between regression methods and failure to report complete models (parameter and error estimates; GSPPM principle 3). Two studies failed to report Weibull model estimates *(103, 104)*, instead choosing to report a metric derived from the model (i.e., time to achieve a predetermined level of inactivation). Additionally, there was a unique phenomenon where studies with log-linear and Bigelow-type models (AKA *D*/*z* models; n=47) that utilized two-step regressions with linearized estimates (n=36) would fail to report a common model parameter (n=34). The Bigelow-type model (Eqn. 3.5) used in these studies requires at least two parameter estimates: the D_{ref} and a *z*-value(s). While these studies would report the *z*-value, they would fail to report the corresponding D_{ref}. If the *D*/*z* model was only accounting for time and temperature effects, a key model parameter to describe the complete model was not reported. Seven other instances of missing secondary model parameters occurred, each associated with a two-step regression. Of the 100 studies reviewed, nearly half were missing a primary model parameter (n=2) or a secondary model parameter (n=41), which prevents any predictive utility of such modeling results.

3.2.3 Model Completeness

Reporting of error terms was also largely absent from the reviewed literature (GSPPM principle 3). Of the 100 studies reviewed, only 37 studies reported a measure of overall model error (MSE or RMSE), meaning most predictive models were missing an associated principal measure of survivor data variability. Studies with metrics that sufficiently quantified parameter

uncertainty were more frequent (n=58). Most parameter-based error metrics were standard errors (n=30), standard deviations (n=19), or a product of one of these two metrics (e.g., relative error, confidence interval, standard half-width). Only 19 studies reported models with both model- and parameter-based estimates of error. As was demonstrated with reporting model parameters, two-step or linear approximation regression methods generally had a detrimental effect on the quantity and quality of error information reported. Unsurprisingly, all studies that failed to report key model parameters (n=43) also did not report that parameter's error estimate. Of the 34 studies failing to report D_{ref}, 13 studies also failed to report an error estimate for the *z*-value(s). Four other studies that used a two-step regression or insufficiently described the regression process failed to report secondary model error estimates *(22, 53, 189, 200)*. Ten studies failed to report any type of model and parameter error.

Most studies reported at least one metric for model goodness-of-fit (GSPPM principle 4; n=86), the most common metric being the standard or adjusted R^2, which was automatically accepted as a goodness-of-fit metric regardless of contextualization. Numerous studies (n=31) included explicit statements of model goodness-of-fit, with the most common metric being RMSE. Ten studies failed to report any measure of goodness-of-fit or use any form of model selection.

Model selection activities and metrics for goodness-of-fit are meant to document the appropriateness of the inactivation model fit in describing the original data (GSPPM principle 4), whereas a model validation demonstrates the model's robustness for prediction of future outcomes. All three (model selection metrics, goodness-of-fit measurements, model validations) were used as model selection criteria at different points in the reviewed literature. A minority of studies incorporated model selection details (n=34), and it was estimated that model selection

would not be applicable to ~54 studies. Most instances of model selection were on primary models (n=25), either as part of a study without secondary models (n=13) or performed prior to secondary modeling (n=12). Most primary model selection criteria used were generally not appropriate for differentiating nonlinear models (e.g., R^2), were biased in favor of additional parameters (e.g., RMSE), or were not appropriate for nested models (e.g., Akaike Information Criterion) *(176)*. Appropriate primary model selection methods included significance testing *(30, 126, 180, 246, 247, 249)* or reporting indicators for a poor model fit (e.g., failure to converge to parameters) *(62, 91)*. Of the studies with model selection criteria applied to secondary models (n=14), another type of model selection included was based on model validations *(120, 215)*, where the models were applied to a dataset separate from model development. Bialka et al. (2007) utilized a cross-validation approach in which the third replication of data was used as the validation dataset, and the subsequent RMSE and R^2 were used to evaluate the linear and Eqn. 3.3 Weibull model fits. This represents a different approach to model selection, where model performance is determined by accuracy of prediction rather than accuracy of description (model fit).

3.2.4 Model Predictivity

Most instances of model validations from the reviewed literature were internal validations and may be considered extensions of verifying model goodness-of-fit (n=8). This included studies incorporating a cross-validation approach *(29, 120)*, a training-calibration-validation approach *(215, 230)*, comparing model parameters to mid-point treatment conditions *(200)*, or comparing a model developed on isothermal data to a dataset generated with identical experimental parameters except for a dynamic temperature profile *(97, 112, 227)*. Because these internal validations involve comparisons against datasets generated with nearly identical

characteristics, there is an expectation that a model with an appropriate structure would be able to describe the validation dataset. Therefore, the resulting comparisons inform model goodness-of-fit more than they inform predictivity. However, if the internal model validation reveals poor model performance, then it is likely that model is not appropriate to describe the phenomenon within the local and predictive applications. The metrics used by internal validations included RMSE (n=7), R^2 (n=3), accuracy/bias factors (n=1), fail-safe/dangerous concepts (n=1).

External validations are better for evaluating an inactivation model's applicability domain and general predictivity because they test the model's performance in true prediction of independent data, where there is less expectation of accuracy. Therefore, external validation can inform where an inactivation model performs well or is unfit to predict. The four studies including external validations each utilized the validation for different purposes or to varying levels of success. Xu et al. (2018) examined the potential of using freeze-dried *Enterococcus faecium* as a surrogate in radio-frequency heat treatments *(270)*. While that study contains the components of an external validation (i.e., isothermally derived model applied to a scaled-up process), because it was ultimately validating the use of *E. faecium* as a *Salmonella* surrogate, the model reported by Xu et al. was not intended to be used as a predictive model elsewhere. Similarly, Mattick et al. (2001) used an external validation for a proof-of-concept *(163)*. In that study, inactivation models developed with a series of water activity modified broths were validated against "real" food, to demonstrate the imperative for LMF research to be performed with actual LMFs. Therefore, Mattick et al. used external validation to demonstrate poor predictivity in the tested cases. The external validation reported by Murphy et al. (2004) resulted in more predicted and observed pathogen reductions in turkey and beef patties than could be quantified using standard plate methods *(183)*, which eliminates the possibility of a quantitative

evaluation of model performance. Santillana Farakos et al. (2013) validated a *Salmonella* persistence model developed using whey protein against various LMFs *(72)*. This is an example of a successful validation of a predictive model because, using metrics like acceptable prediction zones and accuracy/bias factors, they were able to demonstrate better model performance with other non-fat LMFs (wheat flour, non-fat dry milk) than with low-fat LMFs (peanut meal, cocoa powder). Therefore, Santillana Farakos et al. were able to better characterize the model's applicability domain. Overall, internal and external validations are mostly absent in literature and remain a subject that warrants further review.

Both implicit and explicit descriptions of a predictive model's applicability domain were sought for predictive models (GSPPM principle 5). Implicit communication of model applicability was achieved by including parameters that describe the effect of all relevant factors. Then explicit qualification of the model's boundaries or pitfalls would communicate to the reader, and potential end-user of the model, where the authors saw the model being applied successfully. Over half of the studies reviewed reported non-predictive models (n=58). While it may seem like an oxymoron to classify predictive microbiology models as non-predictive, it became clear that many study authors never intended for a given model to be applied elsewhere. Instead, the authors used these models to describe the impact of treatments, often relying on the model parameters for statistical comparisons. For example, Velasquez et al. (2010) studied the impact of intact and nonintact pork muscle structure on the thermal resistance of *Salmonella* at five different temperatures *(244)*. While this study had the necessary foundation to estimate a model that accounts for time and temperature (e.g., Eqn. 3.5), the study did not do so. Rather, inclusion of the log-linear model was intended to estimate thermal resistance, which was then statistically compared at each temperature level. This then allowed the authors to contextualize

their estimates of thermal resistance with other studies reporting thermal resistance as a *D*-value. Several other studies reported a model that would seem predictive normally, but because these studies were designed to evaluate the appropriateness of a non-pathogenic surrogate organism for process validation (e.g., *Enterococcus faecium* NRRL-B2354), the models were intended as tools for comparisons and not for subsequent use predicting inactivation *(127, 152-154, 231, 247)*. Other indicators that predictive models were intended to be non-predictive included using non-food products *(7, 22, 53, 89, 118, 153, 161, 163, 164, 182)*, investigating the impact of experimental or modeling methodologies *(109, 118, 127, 161, 172, 173, 226)*, and using predictive models to approximate universal time-temperature combinations for a process *(1, 102)*. In most of these cases, the design of the experiment and model implicitly limits any utility of the model as a predictive tool.

When the design of the experiment and model implicitly allows for the model to have an applicability domain, often defined by the boundaries of the conditions tested, then explicit statements were needed to reinforce the applicability of the model. Of the 42 studies that reported models that were potentially developed for predictive applications, 10 studies reported models that were classified as unqualified predictive models. To illustrate the need for clear statements of applicability, we will examine the study by Murphy et al. (2002). In that study, there were direct comparisons of *D*/*z* models between multiple meat products, notably chicken patties, chicken tenders, and ground chicken breast from a prior study *(180, 182)*. While the authors made comparisons between model parameters for these three types of products, there was no indication when or if the models should be applied for predictive purposes. The product identities could suggest some applicability domain; however, in applications where a chicken product does not exactly match those used in the study, there was no guidance from the authors

about which model to use, if any. It is possible that some of the unqualified predictive models were meant to be non-predictive but did not include sufficient guidance to inform readers of the distinction.

Interestingly, a minority of studies utilizing predictive models reported models that were interpreted as intended for predictive applications (n=32). While inclusion of an internal or external validation should be a good indication of an inactivation model's predictive capacity, there were other uses for validations within the reviewed literature. A model validation could be used to reinforce the non-predictive characteristics of a model *(163)* or as an assessment of model goodness-of-fit specific to that study's experiment *(29)*. If most studies using inactivation models never intended for those models to be used in a predictive capacity, this may help explain why model parameter and error estimates are frequently underreported. Studies that are not predictive may not need to have the same complete representations of error needed for model predictions. For nonpredictive applications of models, estimating a D_{ref} may not provide more utility than estimating *D*-values. While parameter error may still be necessary for making statistical comparisons between parameters, information specific to the combined model (RMSE, model selection, goodness-of-fit) does not provide the same utility.

A predictive model fit for predictive applications would adhere to the GSPPM principles, which would result in a model being complete (GSPPM principles 1-3) and predictive (GSPPM principles 4-5). There were pervasive issues found within the 100 studies reviewed that would negatively impact predictive models, including incomplete descriptions of regression methodologies (n=9), use of regression practices that complicate estimating error terms such as two-step regressions (n=54) or linearization of nonlinear models (n=44), missing parameter estimates (n=43), missing estimates of model error (n=63), and missing descriptors for parameter

error (n=42). If all predictive models were meant to be predictive, then this would reflect poorly on the standard practice used in predictive microbiology. However, most studies are using these models in nonpredictive capacities (n=58), where limitations to model predictive quality may not be as important. When considering all the required GSPPM principles, only 13 studies were complete (GSPPM principles 1-3), from which only 5 studies were complete and predictive (GSPPM principles 1-5) *(62, 87, 108, 126, 249)*. This is not to say that only 5 predictive microbiology studies between 2000-2020 were reporting high quality predictive models; if this collection of 100 studies is representative of literature during that time, then the implication is that about 5% of studies are reporting predictive models suitable for robust application scenarios, such as a model-based validation of a preventive control.

3.2.5 Summary

Predictive microbiology is a critical tool for understanding and predicting inactivation of foodborne pathogens. Given the recent history of practice for predictive microbiology, it may be best to acknowledge that studies are incorporating descriptive (or non-predictive) and predictive models. Descriptive microbiology may be the use of modeling for deriving statistically significant parameters and comparisons, in contrast to predictive microbiology meeting the criteria for reliable prediction of future outcomes. The evidence of underreporting model parameter and error estimates is still a poor reflection of the current state-of-practice, so separating descriptive and predictive microbiology should not lessen the negative impact of these issues. Good modeling practices should be universal in this field, and the inclusion of descriptive microbiology would just acknowledge that those models are not intended to be used to predict future microbial inactivation outcomes. The hazard of uniformly calling all microbial modeling

"predictive" microbiology is the implication that all reported models are meant to be predictive, enabling the use of a model where it was never intended to be used.

Publications before and after 2000 have expressed the need for complete reporting of predictive model methods, parameter estimates, and error estimates *(38, 66, 241, 243, 265)*, and multiple studies have focused on practical regression techniques *(108, 162, 164, 242)*. However, even studies with the clear predictive inactivation purpose failed to report D_{ref} in several cases *(45, 70, 129-131)*. One reason only 5 of the 32 potentially predictive models were considered complete and predictive was because 17 studies failed to report an appropriate measure of RMSE for the population. Because RMSE of the microbial population describes the likely variability of future data from predicted values, this is a critical estimate of error to report. Even in non-predictive capacities, it is a valuable metric for model goodness-of-fit, as various sources of uncertainty may inflate RMSE *(27)*. Furthermore, no studies performing an internal or external validation reported a model that was considered complete and predictive. As discussed before, inclusion of an internal validation did not necessarily automatically communicate predictivity because it could be used to inform model selection or goodness-of-fit *(22, 29, 97, 112, 120, 200, 215, 227, 230)*, or even be used to indicate non-predictivity *(163)*. The two studies that utilized external validations to evaluate model robustness and predictivity both failed to report population-level RMSE for the combined models *(72, 183)*. Santillana Farakos et. al (2013) reported population-level RMSEs for 5 primary models across 90 treatment combinations as part of the model selection process; however, once the final combined model was reported, all error estimates were solely for the secondary model. There is no doubt that the model reported by Santillana Farakos et al. is predictive and robust, as demonstrated by their use of an external

validation. Their model, like many other incomplete predictive models, was missing a component that communicates expected levels of variability of reduction outcomes.

3.3 Conclusions

A set of Good Standard Practice of Predictive Microbiology was established that consisted of six principles that were desired for predictive models. These principles focused on unambiguity of methods, completeness of model estimation, and expectation of model predictivity. These principles were the basis for evaluating the quality of modeling practices from 2000 to 2020.

A primary finding was that the most common use of predictive inactivation models has nothing to do with predicting inactivation. Because models like the log-linear and Weibull models are commonly accepted for describing inactivation kinetics, studies are perhaps inappropriately applying predictive microbiology terminology to describe the impact of treatments tested without the intention of reporting predictive models. While a standard language for reporting inactivation kinetics likely benefits this field, the assumption that studies using models are necessarily reporting predictive models is potentially hazardous to underinformed users of predictive models. Therefore, it is important to acknowledge when these models are being used for descriptive microbiology vs. predictive microbiology.

Overall, this review highlights the substantial standard practice improvements needed for microbial inactivation models. Each parameter or error estimate absent from inactivation models undermines the potential utility of that model, predictive or non-predictive. By the very nature of predictive models, it should be assumed that someone else may utilize or apply a reported model and, therefore, should be equipped with a complete representation of that model's predictive capacity. Conversely, inactivation models may be properly utilized in nonpredictive research

capacities and should be communicated as nonpredictive to prevent misuse. Future studies should consider the Good Standard Practices for Predictive Modeling principles discussed in this review and should consider incorporating internal/external validations as a robust means to challenge an inactivation model's goodness-of-fit or predictivity, respectively.

CHAPTER 4: REPRODUCIBILITY OF *SALMONELLA* THERMAL RESISTANCE MEASUREMENTS VIA ISOTHERMAL INACTIVATION EXPERIMENTS

This work was published was published in the *J. Food Prot.* in 2020, with the following reference information:

Hildebrandt, I. M., B. P. Marks, N. M. Anderson, and E. M. Grasso-Kelley. 2020. Reproducibility of *Salmonella* thermal resistance measurements via multilaboratory isothermal inactivation experiments. *J. Food Prot.* 83:609-614.

The objective of this study was to quantify the reproducibility of *Salmonella* isothermal resistance results, via an inter-laboratory comparison among six laboratories. Developing standardized methodologies for reproducible estimation of thermal resistance parameters was essential for enabling cross-study comparisons and synthesis.

4.1 Materials and Methodologies

4.1.1 Experimental design overview.

Reproducibility of *Salmonella* thermal resistance was determined by parallel isothermal inactivation experiments performed at the following institutions: Illinois Institute of Technology (IIT), U.S. Food and Drug Administration Center for Food Safety and Nutrition (FDA CFSAN), Michigan State University (MSU – 2 laboratories), University of Nebraska-Lincoln (UNL), and Washington State University (WSU). *Salmonella*-inoculated oat flour samples were prepared by IIT and distributed to each location with strict instructions for the isothermal treatment and data collection. Survivor data were analyzed by a single individual. Additionally, the MSU laboratory group performed a paired comparison using identical equipment but different laboratory personnel, and several batches included a non-shipped control (Control) that was analyzed by the

FDA CFSAN group to examine potential effects of shipping. Each laboratory was randomly assigned a unique identifying number (1-6), used for the duration of this study.

4.1.2 Salmonella Selection

Salmonella enterica serovar Agona 447967 (associated with a puffed cereal outbreak; FDA Arkansas Regional Lab, Jefferson, AR) was stored as a frozen stock at -80ºC. The stock culture was streaked on tryptic soy agar with 0.6% yeast extract (TSAYE, BD Difco, Franklin Lakes, NJ), incubated for 24 h at 37ºC, and subsequently kept at 4ºC for up to 6 months, with working cultures transferred monthly to fresh TSAYE.

4.1.3 Inoculum Preparation

An isolated colony was aseptically transferred to sterile 10 mL tryptic soy broth supplemented with 0.6% yeast extract (TSBYE; BD Difco) and incubated for 24 h at 37ºC. An aliquot (100 µL) of culture was spread on TSAYE and incubated for 24 h at 37ºC to obtain lawn cultures. After incubation, the lawn cultures were harvested by adding 1 mL buffered peptone water (BPW; BD Difco) to the plate, agitating the cells into suspension using a sterile spreader, and aseptically transferring the inoculum to a sterile conical tube. Cultures grown on 15 TSAYE plates were combined to yield sufficient inoculum volume. Harvested cultures were used for sample inoculation within 1 h.

4.1.4 Sample Preparation and Inoculation

Oat flour (ConAgra Foods, Omaha, NE) with an indigenous fat content of approximately 8.5% was acquired in a single batch that was used for the entire study. To reduce influence from background microflora, the oat flour was heat treated via autoclaving in 1 kg batches, as needed, within 24 h prior to inoculation. Oat flour was aseptically transferred to a sterile mixing bowl (900 g) and placed into a mixer (Model N50A, Hobart, Troy, OH), with the remaining 100 g oat

flour aseptically transferred to a sterile Whirl-Pak bag (24 oz, Nasco, Fisher Scientific, Pittsburgh, PA). The flour in the Whirl-Pak bag was inoculated with 10 mL of inoculum and then hand massaged for at least 3 min to incorporate the inoculum uniformly throughout the flour. The inoculated flour was slowly added to the 900 g of oat flour as it mixed at the lowest mixing setting and continued to mix for 30 min after the inoculated flour was completely added. All mixing was performed in an operating biosafety hood with the sash closed whenever possible. The 1 kg inoculated oat flour was then transferred to a sterile aluminum tray (22.9 x 33.0 cm, ~1 cm product depth), and equilibrated to approximately 0.45 a_w by holding at 45% relative humidity (RH) in a humidity-controlled chamber *(213)* for 5 d.

Homogeneity was assessed by taking 1 g samples (n = 5 to 10) randomly from 1 kg of inoculated, equilibrated flour. Samples were diluted with BPW in 7 oz Whirl-Pak bags, hand massaged for approximately 30 s, and then serially diluted using BPW. Aliquots were spread plated on duplicate TSAYE plates, incubated for 24 h at 37ºC, and survivors counted. Only homogenous batches, defined as having standard deviations less than 0.33 log CFU/g, were distributed for subsequent isothermal inactivation analysis.

4.1.5 Isothermal Inactivation.

Equilibrated, homogenously inoculated oat flour (50 g) was packaged into airtight plastic containers (Fisherbrand LDPE Wide-Mouth Bottles; Thermo Fisher Scientific, Waltham, MA), then packaged into a sealable metal container (401 Diameter Welded Sanitary Style Can; Container Supply Company, Inc., Garden Grove, CA), then into an insulated box and shipped overnight to the individual laboratories. Eight 1 kg inoculated batches were produced (designated as batches A-H). For each batch, at least 100 g of inoculated product was shipped to each participating laboratory. Serving as a control, portions of several batches were prepared to ship,

but not shipped, and then analyzed by the FDA CFSAN group. Because IIT and FDA CFSAN occupied separate laboratories in the same building from which the samples were distributed, samples designated for these laboratories were shipped to MSU and then returned to IIT/FDA CFSAN to achieve similar handling history as the samples processed at other laboratories.

Each laboratory was provided with a checklist of requirements to ensure treatment uniformity across laboratories. This checklist included requirements about equipment calibration and record keeping. Limited information about the expected isothermal inactivation rate was provided to each laboratory (e.g., the approximate initial population levels and *D*-value range of 20-35 min was provided), allowing each group to determine its own sampling time points. In addition to following the required actions, each group was required to document that each action was performed. Each laboratory was required to process ≥ 3 batches within 3 weeks of batch inoculation.

Prior to isothermal treatment, a series of confirmation measurements were conducted to verify that test parameters were consistent across the participating groups. Three a_w measurements were taken once the inoculated oat flour was opened prior to packing into aluminum test cells *(54)* and after the test cells were packed. Test cells were packed to maximum capacity (~0.8 g). Laboratory 3 used similar stainless-steel test cells with a larger sample capacity (~7 g oat flour) and analyzed 3 g samples. Sample manipulation was conducted within a relative humidity-controlled chamber at 45% RH to minimize rapid moisture changes that can occur when powders are exposed to ambient conditions in a laboratory environment *(213)*. The average a_w of these three measurements had to be 0.45 ± 0.025 to proceed to subsequent isothermal inactivation treatment. Additionally, while samples were within the humidity-controlled chamber, three 1 g samples of inoculated oat flour were taken, and *Salmonella* were

enumerated as described previously to verify that *Salmonella* population levels did not change during transport, storage, and handling.

Two different temperature measurements were performed to confirm that the target temperature was achieved for each isothermal treatment: a T-type thermocouple was used prior to and during treatment to verify that the heating medium was at the correct temperature, and an aluminum test cell fitted with a T-type thermocouple and packed with inoculated oat flour was used to measure come-up-time (CUT; the time for the product to be within ± 0.5 °C of the target treatment temperature) and to verify that the product reached and maintained the treatment temperature.

The isothermal inactivation treatment consisted of at least 6 time points with 3 subsamples at each time point treated at 80°C. Either a circulating water- or oil-bath was used to treat the samples, depending on the laboratory equipment. Regardless of the heating medium, the CUT was measured for each batch, and the initial samples were collected 3 min (which was greater than all measured CUTs) after the treatment began to ensure that the treatment was at isothermal conditions for all sampling time points. Once taken out of isothermal treatment, the test cells were immediately submerged into an ice-water bath to drop the sample temperature below inactivation temperatures (< 30 s). Surviving *Salmonella* was then enumerated as described above for the homogeneity testing.

4.1.6 Model Regression and Statistics

Each replication of survivor population was converted to the logarithmic scale (log N) and then transformed to the survivor ratio (log N/N_0) by subtracting the average of the time 0 counts (log N_0, the samples pulled after the 3 min CUT) from each data point. Each data series was categorized using laboratory and batch source information for regression analyses.

Log-linear inactivation kinetics were estimated using MATLAB nonlinear regression tool *nlinfit* (MathWorks, Natick, MA) and Eqn. 3.1.

$$log\left(\frac{N}{N_0}\right) = -\frac{t}{D} \quad \text{Eqn. 3.1}$$

where *D* is the decimal reduction time (min), N and N_0 are the *Salmonella* population at time t (min) and 0, respectively. For each model fit, the root mean squared error (RMSE) and adjusted correlation coefficient (R^2_{adj}) were calculated. Resultant model residuals were visually inspected to verify no systematic bias. The Weibull model *(160)* parameters were also estimated as described for the log-linear model; however, because 70% of the experimental replications resulted in ≤ 3 log reductions of *Salmonella*, the resultant Weibull model overfit the data. Therefore, log-linear model estimates were used as the metric describing thermal resistance for all subsequent comparisons.

Potential differences between *D*-values that resulted from processing laboratory or batch were determined via pairwise comparisons of Welch's t-test (Eqn. 3.2) using the parameter (*D*) and standard error (SE) estimates (type I error = 0.05). Using the same comparisons, potential differences in model residual variances were evaluated using an F-test. Analysis of covariance (ANCOVA) tests were performed to determine whether sample shipping, isothermal heating medium, and measured a_w impacted the thermal resistance (type I error = 0.05), because *D*-values were not estimated for these subsets of data.

$$t = \frac{|D_1 - D_2|}{\sqrt{SE_1^2 + SE_2^2}} \quad \text{Eqn. 3.2}$$

4.2 Results and Discussion

4.2.1 Inoculation Homogeneity

The homogeneity results for each batch of inoculated oat flour yielded a mean population between 8.13 and 8.38 log CFU/g, with population standard deviations from 0.14 to 0.31 log CFU/g. The low standard deviations indicated that the populations were homogeneously distributed in all the batches. There were no significant differences in initial inoculation levels between batches (type I error = 0.05). The mean post-shipment *Salmonella* population measured by each location was within ± 0.5 log CFU/g of the mean pre-shipping population levels.

4.2.2 Sample Water Activity

Pre-shipping a_w measurements ranged from 0.43 to 0.47, all within the ± 0.025 limit. A limit of ± 0.025 of the target a_w was selected because this value was larger than the accuracy of common laboratory-scale a_w meters, allowed tolerance for some potential conditioning offset or a_w drift during shipment, and still required strict conditioning/handling controls to maintain. Because some pre-treatment a_w measurements were outside of the acceptable a_w range, several isothermal treatments either used the second 50 g container included with each batch or postponed experimentation until a new batch of samples was available. All isothermal treatments utilized samples with a_w measurements within the acceptable range.

4.2.3 Model Fit

Visual inspections of model residuals resulting from the log-linear model fit to data grouped by laboratory supported the use of the log-linear model with 6/7 datasets showing no signs of non-linearity. Inspection of residuals for laboratory 3 suggested a possible biphasic trend; however, because a majority of the datasets resulted in ≤ 3 log reductions of *Salmonella*,

the estimation of nonlinear or biphasic models was problematic. Therefore, all comparisons of *Salmonella* thermal resistance reproducibility were based on results from the log-linear model.

4.2.4 Thermal Resistance

Differences among batch-dependent *Salmonella* thermal resistances resulted in two groupings – the first comprised of batches A-C and F-H, and the second compromised of batches D-E (Table 4.1, Figure 4.1). The significantly lower *D*-values ($P < 0.05$) for batches D and E were consistent regardless of laboratory, which suggested that the inoculation methodology was the source of the differences, but no noted deviations in procedures were noted or recorded. Batches were sent out in three phases (A-C, D-E, and F-H), allowing for the possibility that the lower *D*-value with the second-phase batches could have been an anomaly. After an internal

Table 4.1 Log-linear model parameters for each batch of *Salmonella*-inoculated oat flour.

Batch	a_w* (unitless)	$D_{80°C}$** (min)	*D*-error (min)	RMSE*** (log CFU/g)	R^2_{adj} (unitless)
A	0.44±0.013	28.85[a]	1.09	0.54[a,b]	0.58
B	0.46±0.009	30.87[a]	0.83	0.44[a]	0.73
C	0.44[+]	38.04[a]	3.48	0.57[a,b,c]	0.50
D	0.46±0.01	15.21[b]	0.53	0.87[c]	0.79
E	0.46±0.017	16.98[b]	0.99	0.78[c]	0.82
F	0.45±0.003	30.06[a]	0.70	0.53[a,b]	0.86
G	0.46±0.007	29.42[a]	1.01	0.66[b,c]	0.81
H	0.47±0.009	28.38[a]	0.92	0.71[c]	0.74

* a_w values are represented as means ± standard deviation.
** Parameter estimates sharing a common superscript letter are not significantly different via t-test.
*** RMSE values sharing a common superscript letter did not have significantly different residual variance via F-test.
+ only one laboratory collected data using batch C.

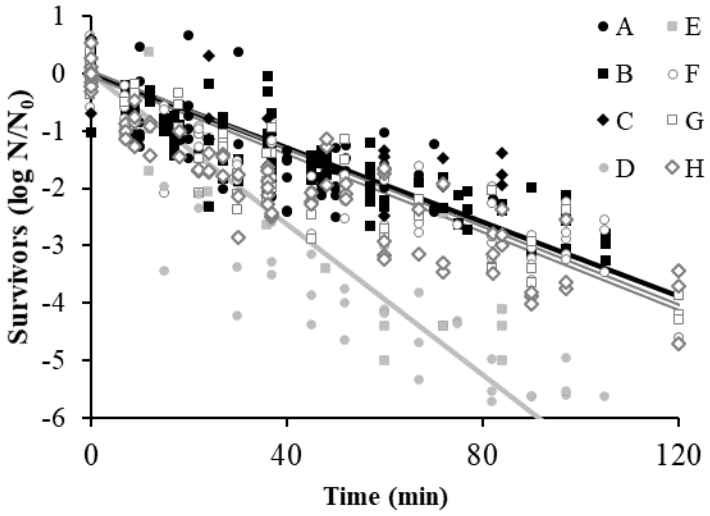


Figure 4.1 Isothermal (80°C) *Salmonella* survivors and log-linear model fits for each *Salmonella*-inoculated batch of oat flour. Batches A, B, C are represented with black shapes and line; batches D, E are represented with gray shapes and line; batches F, G, H are represented with hollow gray shapes and line.

review of procedures, the exact source of the difference could not be identified; however, a potential source of the discrepancy was identified as the use of a culture beyond the pre-determined shelf-life. The standard operating practice for this experiment required that the cultures never exceed five passages (or monthly transfers to fresh media) removed from the original frozen stock culture. As a result, the *Salmonella* culture used for batches D and E may have accumulated genetic drift that weakened the resultant thermal resistance. These results could not be replicated because the culture potentially used for batches D and E was disposed of immediately upon discovering discrepancy in *D*-values. Other non-inoculation sources potentially contributing to the observed difference in *D*-values between batches (such as incorrect treatment a_w or temperature, shipping, laboratory-bias, or human error) were dismissed due to unlikelihood of the difference caused by these other sources resulting in the magnitude

and systematic nature of the observed error. Because of the differences seen between batches, only data collected from batches A-C and F-H were used for subsequent analyses.

Table 4.2 *Salmonella* $D_{80°C}$-values generated using identical materials/methods for each lab

Laboratory	Batches Processed	$D_{80°C}$* (min)	*D*-error (min)	RMSE** (log CFU/g)	R^2_{adj} (unitless)
Control	A,B,F,G	29.26[a]	1.04	0.50[b]	0.70
1	B,C,H	32.58[a]	1.58	0.59[b,c]	0.60
2	A,B,F	30.79[a]	1.22	0.59[b,c]	0.68
3	F,G,H	30.31[a]	0.77	0.72[c]	0.82
4	A,B,F	29.68[a]	1.03	0.46[a,b]	0.76
5	A,B,H	28.70[a]	1.09	0.56[b,c]	0.66
6	A,B,G,H	19.96[b]	0.62	0.36[a]	0.79
All+	-	30.21	0.41	0.57	0.80

* Parameter estimates sharing a common superscript letter are not significantly different via t-test.
** RMSE values sharing a common letter did not have significantly different residual variance via F-test.
+ did not included data from laboratory 6.

There were no significant differences between the non-shipped control and the participating laboratories, except for laboratory 6 (Table 4.2, Figure 4.2). The *D*-value from this group was 19.96 min, while *D*-values from other laboratories were 28.70 - 32.58 min. The *D*-values for each replication generated by laboratory 6 were consistently less than the *D*-values generated by all other groups, yielding *D*-values of 19.41, 24.94, 18.41, and 18.33 min for batches A, B, G, and H, respectively. Therefore, a post-test review of procedures was conducted to confirm that all procedures were strictly followed. The conclusion of the review was that the treatment temperature was not adequately monitored, and the actual temperature of the oil bath was likely 0.5-2.0°C higher than the target isothermal temperature. While this study did not investigate the effect of temperature on *D*-values, *Smith* et al. *(213)* reported a *z*-value for *Salmonella enterica* Enteritidis Phage Type 30 in wheat flour of 15.2°C. Assuming an average

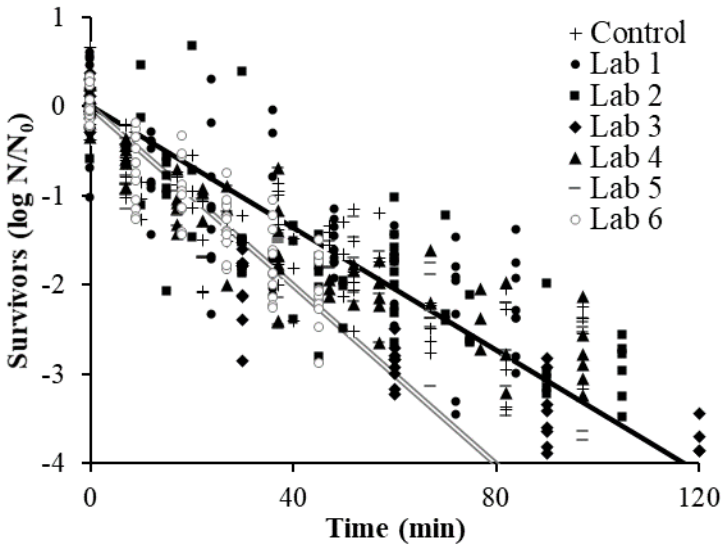


Figure 4.2 Isothermal (80°C) *Salmonella* survivors for each laboratory, with statistically similar thermal resistances sharing a similar line and gray-scale color.

D-value of approximately 30 min at 80°C and using the 15.2°C *z*-value, the *D*-value at 82°C would be approximately 22 min, which would be within the range of *D*-values observed at laboratory 6.

Differences in measured a_w (0.427 - 0.475) and storage time during the 3-week experimental window did not appear to impact *Salmonella* thermal resistance (results not shown).

4.2.5 Experimental Reproducibility

This study demonstrated that if strict experimental controls are maintained, then repeatable thermal resistances were achievable. It also demonstrated that slight deviations in procedures may yield significant differences in measured thermal resistance. Meta-analyses for pathogen inactivation in foods have indicated that cross-study variation is an issue *(79, 239)*, but could not determine whether such variation was caused by differences in methodologies or variability inherent to isothermal inactivation studies. For example, van Asselt and Zweitering

(239) examined *Salmonella D*-values from 20 studies with a mean and upper 95% percentile $D_{70^{o}C}$ of 0.15 and 3.89 min, respectively. Despite the studies investigating different products, *Salmonella* serovars, and a_w, no trends in the data could account for the greater than 20-fold difference in *D*-values collected. Previous research has suggested that differences in measured thermal inactivation parameters may be the result of different *Salmonella* serovars incorporated amongst studies; however, Hildebrandt et al. *(108)* reported significantly different thermal inactivation kinetics using identical *Salmonella* cocktails and food materials but different isothermal inactivation methodologies. Additionally, future research should consider that the shape of the inactivation curve, and not just the resultant model parameters, could be influenced by experimental methodologies.

Results from the present study indicated that results from identical methodologies, without procedure deviations, yielded *D*-values that were ± 10% of the mean value. However, small procedural deviations resulted in *D*-values ~50% of the previously mentioned mean value. While the actual difference caused by various isothermal methods is unknown, the results of this study indicate that the ability to synthesize general trends from a group of studies using even slightly different methods may be limited.

With few exceptions, the *D*-values generated by the different laboratories were not significantly different; however, differences in laboratory personnel appear to have affected resulting standard error of parameter estimates and model error (RMSE; Table 4.2). The relative standard error of the *D*-values was similar, with errors that were 3.5, 4.8, 4.0, 2.6, 3.5, 3.8, and 3.1% of the parameter for the control and laboratories 1-6, respectively. Laboratory 3 achieved the smallest relative *D*-value error (2.6%); laboratory 3 also achieved the highest residual variance, with a RMSE of 0.72 log CFU/g, though the residual variance was not significantly

different than those generated by laboratories 1, 2, and 5. Despite the procedural deviation with temperature control, laboratory 6 achieved the smallest residual variance, with a RMSE of 0.36 log CFU/g.

Despite yielding similar *D*-values, the model error varied significantly between laboratories. Additionally, if the isothermal experiment did not contribute to model error, then RMSEs should be comparable to the standard deviations estimated in the homogeneity tests. All RMSEs were larger than the initial population standard deviations, indicating that the error increases during thermal treatment data collection. While the methods incorporated in this study yielded reproducible *D*-values, lab-dependent human/equipment error influenced the quality of data collection, as expected. Previous research reported that RMSEs from *Salmonella* inactivation models generated using similar methodologies varied between two laboratories *(108, 109)*; however, no analyses that were performed in that study concluded whether residual variance was significantly different. Overall, estimates of RMSE for *Salmonella* inactivation models are partially attributable to human, equipment, and scale errors, which may be another source of error that can be minimized through detailed inactivation procedures and personnel training.

4.3 Conclusions

Overall, this study quantified the reproducibility of *Salmonella D*-values to be within a ± 10% of the mean value, and through procedural failures this study was able to demonstrate that small deviations in methodologies may yield differences up to ~50% of the mean *D*-value. Additionally, heating medium, sample distribution, personnel, and small differences of measured a_w (± 0.025 of target a_w) did not impact the *Salmonella* thermal resistance; however, such factors still may impact data collection error during the isothermal experiment. Future research should

carefully evaluate isothermal inactivation studies to identify methodologies that may bias thermal resistance or promote reproducibility. This would help promote uniformity across studies and laboratories, enabling more impactful results from each study and from the broader and growing body of work in this area.

4.4 Acknowledgements

The work was supported by the U.S. Department of Agriculture, National Institute of Food and Agriculture, award no. 2015-68003-23415. Additionally, this work was supported by the U.S. Food and Drug Administration collaborative grant number 5U01FD003801 and by appointments to the Research Participation Program at the Center for Food Safety and Applied Nutrition administered by the Oak Ridge Institute for Science and Education (ORISE) through an interagency agreement between the U.S. Department of Energy and the U.S. Food and Drug Administration. This work would not have been possible without the laboratory staff at each location who participated in this study. We acknowledge and thank Shannon Pickens at IIT, Dr. Jeyamkondan Subbiah and Soon Kiat Lau at UNL, Dr. Juming Tang, Dr. Meijun Zhu, and Jie Xu at WSU, and Nicole Hall, Pichamon Limcharoenchat, and Francisco Garces-Vega at MSU. Additional thanks go to Dr. Susanne Keller with FDA CFSAN for contributing her expertise on microbiology methodologies during the study design.

CHAPTER 5: DEMONSTRATION OF INAPPROPRIATE VALIDATION METHOD FOR A CRACKER BAKING PROCESS USING PREDICTIVE MODELING

This study was submitted to the *J. Food Prot.* and is under review. This study was performed in response to a series of journal articles that promoted a flawed approach to model-based validations of low-moisture foods. In addition to highlighting incorrect techniques for developing models for use in preventive controls, this study includes discussion on the importance of model validations.

5.1 Methods and Materials

This study included three parts: (1) development of a *Salmonella* inactivation model based on isothermal, single-moisture data from a dough product, (2) collection of *Salmonella* survival data in a baking process, and (3) evaluation of inactivation model performance against the *Salmonella* survivors in the baked product, testing both a time/temperature-only model and a candidate time/temperature/moisture model. Crackers were selected as the baked product to test, due to the simple formulation and consistent product dimensions. Additional isothermal inactivation tests were conducted to compare *Salmonella* inactivation kinetics in raw vs. post-baked product, to demonstrate the magnitude of differences in thermal resistance over the baking process.

The isothermal, single-moisture inactivation experiments for model development and data collection for the baked cracker dataset were designed to match studies that disseminated predictive inactivation models for baked goods *(45-48, 140, 141, 237)*. All other analyses (validation of the isothermal, single-moisture model; isothermal inactivation of a post-baked product; and a time/temperature/moisture model regression on the validation dataset) were performed to evaluate the appropriateness of the core study design rigorously and quantitatively.

5.1.1 *Salmonella* Inoculation and Cracker Formulation

Salmonella enterica serovars Agona 447967 (associated with cereals), Reading Moff 180418 (FDA CFSAN, Bedford Park, IL), Tennessee K4643 (associated w/ peanut butter), Montevideo 488275 (associated w/ black and red pepper), and Mbandaka 698538 (associated w/ tahini) were maintained as frozen stock cultures at -80°C. The stock cultures were resuscitated with two subsequent 24 h incubations at 37°C in tryptic soy broth supplemented with 0.6% yeast extract (Difco, BD, Franklin Lakes, NJ). Then each culture was transferred to a plate (150 x 15 mm) of tryptic soy agar supplemented with 0.6% yeast extract (Difco, BD) and incubated for 24 h at 37°C to form a lawn culture. Each lawn culture was harvested with 2.5 ml of buffered peptone water (BPW; Difco, BD) using a sterile L-shaped spreader to agitate the culture into suspension and then combined in equal volumes to form the inoculum. Each replication of the treatments repeated this process to yield separate biological replications.

A simple cracker formulation was used as the model food for the baking process, which consisted of 160 g (50%) whole wheat flour, 90 ml (27%) water, 51.2 g (16%) vegetable shortening, 18.8 g (6%) sugar, and 2.8 g (1%) salt. Each dry ingredient was purchased at a local supermarket and stored at room temperature. Water was de-ionized and sterilized prior to dough formation. Whole wheat flour (16 g) in a sterile plastic bag was inoculated with 1.6 ml of the *Salmonella* cocktail, hand-mixed until the inoculum was visibly incorporated (i.e., liquid no longer visibly observable), then combined with 144 g of whole wheat flour, hand-mixed for 10 min, and stored in the sealed bag at room temperature for at least 24 h.

On the day of experiments, the cracker dough was formed by combining each ingredient in a sterile plastic bag and hand-massaging the dough until each component was well mixed. The cracker dough was not proofed and was utilized within 2 h of mixing. Water activity (a_w) of the

dough at 25°C was measured using an AquaLab 4TE a_w meter (Decagon Devices, Pullman, WA). Each batch of inoculated cracker dough was divided and used in: (1) isothermal/iso-moisture experiments to estimate *Salmonella* inactivation kinetics or (2) cracker baking experiments in a bench-top convection oven. Each batch yielded paired results of data used to generate an inactivation model and data to test the model.

5.1.2 Isothermal/Iso-Moisture Inactivation in Raw Dough

For the isothermal/iso-moisture treatments, the dough samples (~1 g, ~4 mm thick) were packed into aluminum test cells *(54)* then heat-treated in a circulating water bath at 56, 60, or 63°C. The initial "time zero" samples were pulled after isothermal conditions were achieved within the samples (come-up-times were 102-180 s), which was determined via a test cell fitted with a T-type thermocouple at sample center reaching within 0.5°C of the target temperature. Thereafter, triplicate subsamples (for each replicate) were pulled at ≥6 uniform time intervals. Samples were immediately cooled in an ice-water bath for ~30 s, aseptically removed from the test cells, diluted with BPW, homogenized for 3 min (400 ml IUL Masticator Silver, IUL S.A., Barcelona, Spain), serially diluted with BPW, and plated on tryptic soy agar supplemented with 0.6% yeast extract, 0.05% ammonium ferric citrate, and 0.03% sodium thiosulfate (mTSA), a differential medium on which *Salmonella* survivors were enumerated after 48 h incubation at 37°C.

5.1.3 Cracker Baking Experiments

The cracker baking tests were conducted in parallel to the isothermal treatments described above. To form the crackers, 100 g of inoculated cracker dough was rolled out and cut into squares (25.4 x 25.4 x 2 mm). Cracker squares were baked on an aluminum baking sheet in a benchtop forced-convection oven (Smart Oven Pro, Breville, Sydney, Australia) at 177°C for

30, 60, 90, 120, 150, 180, 210, 240, 270, 300, 330, and 360 s (with a separate oven run for each of those sampling intervals, so that the oven door was closed during the entire baking time for any given sample). Each treatment consisted of four cracker squares, three for *Salmonella* survivor enumeration and one for temperature, moisture content, and a_w measurement. During the baking process, product temperature was recorded at 2 s intervals, using a K-type thermocouple inserted at the geometric center of the cracker. At the predetermined times, samples for enumeration were immediately transferred to chilled BPW, homogenized for 3 min, serially diluted with BPW, and plated on mTSA, with survivors enumerated after 48 h incubation at 37°C. The cracker used to monitor temperature was transferred to an air-tight container, allowed to cool to room temperature, prior to measuring a_w using the AquaLab 4TE a_w. After a_w measurement, sample moisture content was determined by AOAC method 950.46B *(19)*.

5.1.4 Isothermal/Iso-moisture Inactivation in Post-Bake Cracker Crumbs

To quantify the differences in *Salmonella* thermal resistance in raw dough vs. baked crackers, due to moisture differences, an additional set of isothermal/iso-moisture inactivation experiments was conducted using post-baked crackers. Uninoculated cracker dough was prepared, formed into crackers, and baked as described above, to yield a final a_w similar to that of the inoculated crackers baked for 360 s (~100 g at 0.4-0.5 a_w). These baked crackers were then crushed into a crumb consistency via stomacher for 3 min. The crumb then was inoculated with 1 ml *Salmonella* cocktail, hand-mixed until 3 min past when the inoculum was visibly incorporated, and placed within a humidity-controlled chamber *(213)* set at 45% relative humidity for 24-72 h. Inoculated and re-equilibrated cracker crumb samples (0.452 ± 0.001 a_w) were packed into the aforementioned aluminum test cells *(54)* and then treated in a circulating water bath at 90°C. Once the isothermal conditions were met within the samples (as described

above for the dough samples), triplicate subsamples were pulled at ≥7 uniform time intervals. Samples were immediately cooled in an ice-water bath for ~30 s, aseptically removed from the test cells, diluted with BPW, homogenized for 3 min, serially diluted with BPW, plated on mTSA, and survivors enumerated after 48 h incubation at 37°C.

5.1.5 Model Regression and Validation

Each isothermal/iso-moisture survivor value was converted to the logarithmic scale (log *N*) and then transformed to the survivor ratio (log N/N_0) by subtracting the average of the time 0 counts (log N_0) from each data point.

Isothermal log-linear inactivation kinetics (*D*-values) were estimated from the isothermal/iso-moisture data using MATLAB nonlinear regression tool *nlinfit* (MathWorks, Natick, MA) and Eqn. 5.1.

$$log\left(\frac{N}{N_0}\right) = -\frac{t}{D_T} \qquad \text{Eqn. 5.1}$$

Where D_T is the decimal reduction time (min) at the isothermal treatment temperature *T* (°C) and specific moisture content, and *N* and N_0 are the *Salmonella* populations at time *t* (min) and 0, respectively. The appropriateness of the log-linear primary model was verified by visually inspecting the distribution of data about the prediction lines and by evaluating the normality of the residual distribution at each temperature. Secondary model parameters for the cracker dough were estimated using a Bigelow-type secondary model (Eqn. 5.2) combined into Eqn. 5.1, and one-step/global regression (with MATLAB nonlinear regression tool *nlinfit*) applied to the pooled survivor data from all times, temperatures, and replicates from the specific dough tests (raw or baked, separately).

$$D_T = D_{ref} \times 10^{\left(\frac{T_{ref}-T}{z}\right)} \quad \text{Eqn. 5.2}$$

Where D_{ref} is the decimal reduction time (min) at the reference temperature T_{ref} (60°C in this study), and z is the change in temperature (°C) needed for a decimal change in D_T. Root mean squared error (RMSE) was calculated to quantify the goodness of each model fit.

After estimating the global model parameters from the isothermal/iso-moisture inactivation data, the predicted *Salmonella* inactivation during the cracker baking experiments was estimated using a trapezoidal integration of the raw cracker dough inactivation model (Eqn. 5.1 and 5.2) over the cracker temperature profile starting at the 30 s sample. Although it was known that the cracker moisture was changing significantly during baking, and the model was based on iso-moisture data (from the dough samples), this model validation was nonetheless carried out to achieve the explicit objective of this study – to demonstrate the systematic errors that can occur when such an approach (as has been previously published and promulgated) is followed. Multiple studies have reported pathogen inactivation models based on dough-based isothermal experiments *(45-48, 140, 141, 237)*, despite also reporting product drying during processing. The RMSE of the applied prediction model and a modified Acceptable Prediction Zone (mAPZ) were evaluated as model validation metrics. The mAPZ metric modifies the original APZ metric *(192)* to account for observed variability during model development using the criteria outlined in Eqn. 5.3.

$$R_i = log\left(\frac{N_i}{N_0}\right)_{observed} - log\left(\frac{N_i}{N_0}\right)_{predicted}$$

$$-1.96 * RMSE \leq R_i \leq 1.96 * RMSE \quad Acceptable\ Prediction\ Zone$$

$$\phi * -1.96 * RMSE \leq R_i < -1.96 * RMSE \quad Lean\ Fail\ Safe$$

$$1.96 * RMSE < R_i \leq \phi * 1.96 * RMSE \quad Lean\ Fail\ Dangerous$$

$$R_i < \phi * -1.96 * RMSE \quad Fail\ Safe$$

$$R_i > \phi * 1.96 * RMSE \quad Fail\ Dangerous$$

Eqn. 5.3

Where R_i (log CFU/g) is the model residual for sample *i*, *RMSE* is the model error (RMSE in log CFU/g) observed during model development, and ϕ (unitless) is the scale-up factor to adjust for the expected increase in variability when collecting data in a less controlled environment. With this definition, the mAPZ can discriminate between samples that would occur within the 95% prediction interval for the original model (acceptable prediction zone), the anticipated 95% prediction interval for the scaled-up process (anticipated prediction zone that includes lean fail safe to lean fail-dangerous samples), or samples outside the accepted or anticipated prediction zones that are either fail-safe or fail-dangerous. The scale-up factor ϕ can be set to account for increased prediction uncertainty and variability associated with larger-scale processes *(106, 119)*. Because the mAPZ metric is novel to this paper, and the anticipated increase in variability from isothermal inactivation to an oven process was unknown, ϕ was conservatively set to 2, based on prior studies that included both lab- and pilot-scale inactivation data *(121)*.

As a final means to evaluate the appropriateness of using an isothermal, single-moisture model applied to a dynamic temperature/moisture case, the cracker baking validation dataset was used to estimate inactivation model parameters for a model that included a moisture term (Eqn. 5.4) *(42)*.

$$D_T = D_{ref} \times 10^{\left(\frac{T_{ref}-T}{z_T}+\frac{MC_{ref}-MC}{z_M}\right)} \quad \text{Eqn. 5.4}$$

Where the parameters are as described previously, with the following additions: z_T is the change in temperature (°C) needed for a decimal change in D_T specific to the dynamic profile model, MC_{ref} is the reference dry basis (db) moisture content (MC), which was set to the approximately maximum moisture content measured (0.5 g water/g dry), and z_M is the change in moisture content needed for a decimal change in D_T. Parameters were estimated using MATLAB nonlinear regression tool *nlinfit*, where the implicit form of the primary model was integrated over cracker temperature and moisture profiles beginning with the 30 s samples. In preliminary work, both a_w MC were evaluated as candidate variables for the water effect in Eqn. 5.4, and moisture content was selected because it resulted in a better overall fit. Because the validation dataset was generated specifically to test the isothermal, single-moisture model, the data collected were not sufficient to create a valid stand-alone time/temperature/moisture model, so that this subsequent analysis was conducted only for illustration of one appropriate approach for including dynamic moisture effects.

5.2 Results and Discussion

5.2.1 Isothermal/Single-Moisture Inactivation Kinetics

The initial *Salmonella* populations in the raw dough and post-bake cracker crumb samples were 8.1 ± 0.15 and 8.3 ± 0.15 log CFU/g, respectively. Isothermal treatment of *Salmonella* in cracker dough resulted in inactivation outcomes (Figure 5.1) that supported the use of the log-linear model to describe *Salmonella* inactivation kinetics (Table 5.1). The average a_w of the raw cracker dough was 0.956 ± 0.002, and the average moisture content (db) was 0.50 ± 0.01. Estimated raw cracker dough *Salmonella* $D_{60°C}$ and *z* values were 4.6 min and 4.9°C, respectively (Table 5.1). The cracker dough in this experiment yielded results similar to other

studies reporting *Salmonella* inactivation in dough products. Jin et al. reported $D_{60°C}$ and *z* for *Salmonella* in a matrix loosely based on cracker dough (0.98 a_w) of approximately 1.23 min and 6.03°C, respectively *(126)*, with the lower $D_{60°C}$ expected at the higher a_w. Inactivation kinetics reported for *Salmonella* in hamburger bun dough (0.97 a_w) were also similar, with $D_{61°C}$ and *z* of 3.1 min and 6.6°C, respectively *(45)*. Another study tested *Salmonella* in a slightly lower a_w muffin dough (0.92 a_w) and reported, as expected, a higher $D_{61°C}$ of 16.5 min, with a *z* of 10.4°C *(141)*.

Salmonella inactivation in post-bake cracker crumbs (0.45 a_w) resulted in a $D_{90°C}$ of 2.97 min (standard error of 0.11 min) and RMSE of 0.71 log CFU/g. While the raw cracker dough-based *Salmonella* inactivation parameters from this study were comparable to those for other high-moisture products, the $D_{90°C}$ in cracker crumble (3.0 min) was ~6 orders of magnitude larger than the extrapolated $D_{90°C}$ from the raw cracker dough-based model (0.0000035 min). However, the results from the post-bake cracker crumbs were similar to the estimated 1.85 min $D_{90°C}$ for *Salmonella* in a "cracker" matrix (0.50 a_w) reported by Jin et al. *(126)*. The large

Table 5.1 *Salmonella* $D_{60°C}$ and *z* estimates (w/ standard errors) for each experimental replication and the combined dataset from isothermal and single-moisture experiments in raw cracker dough (a_w ~0.96, moisture content = 0.50 ± 0.01 dry basis).

Batch	$D_{60°C}$ (min)	*z* (°C)	RMSE (log CFU/g)
1	4.24 (0.11)	4.54 (0.07)	0.42
2	5.12 (0.10)	5.27 (0.07)	0.27
3	4.50 (0.09)	4.98 (0.07)	0.37
Combined	4.60 (0.06)	4.90 (0.05)	0.41

Salmonella thermal resistance difference between pre- and post-bake cracker products can be explained by the difference in a_w. As a result, the *Salmonella* inactivation model based on isothermal/iso-moisture data from the raw cracker dough is inappropriate for predicting *Salmonella* inactivation at the end of the process; however, this information alone is insufficient to determine when the cracker dough model becomes inappropriate to use in baking validation experiments. Theoretically, a cracker dough-based model might be adequate to use if the cracker moisture is sufficiently high for sufficient time during baking to achieve the target *Salmonella* inactivation prior to the thermal resistance increasing significantly. However, the use of model parameters developed using the wet product dough (which is when the product contains the most moisture) would result in the least conservative estimate when applied to a process where heating and drying are immediate and inseparable. Therefore, independent validation of such a model in processes approximating commercial baking conditions is critical.

5.2.2 Cracker Baking Results

By the end of the baking treatment (Figure 5.2), cracker moisture content was reduced to 0.06 ± 0.018 db, which was close to the moisture content of 0.04 required to meet the standard of identity for whole wheat crackers *(233)*. The dynamic temperature and moisture profiles of the crackers illustrate a rapid increase of temperature until approximately 120 s, where the temperature plateaued at ~100°C, while the moisture content and aw steadily declined during the entire baking process. This baking treatments yielded 5.95 ± 0.76 log reductions of Salmonella after 360 s of bake time, making the mean reduction > 5 log ($P < 0.05$; Figure 5.3).

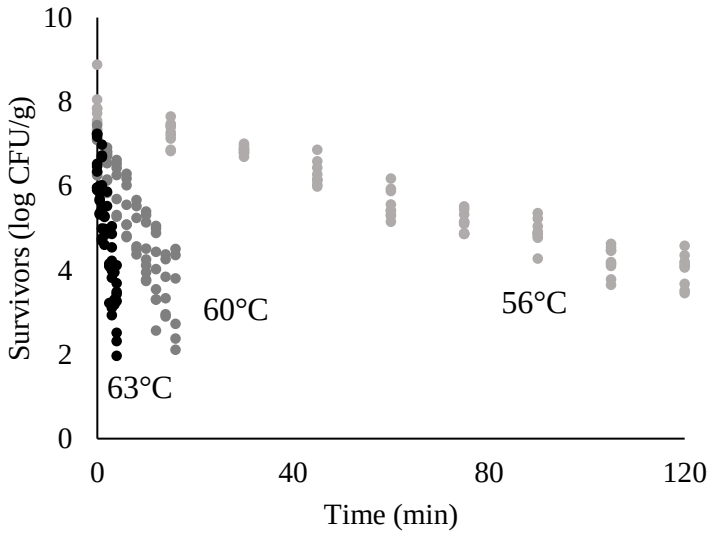


Figure 5.1. *Salmonella* survivors in raw cracker dough (a_w ~0.96, moisture content = 0.50 ± 0.01 dry basis) during isothermal and single-moisture treatments.

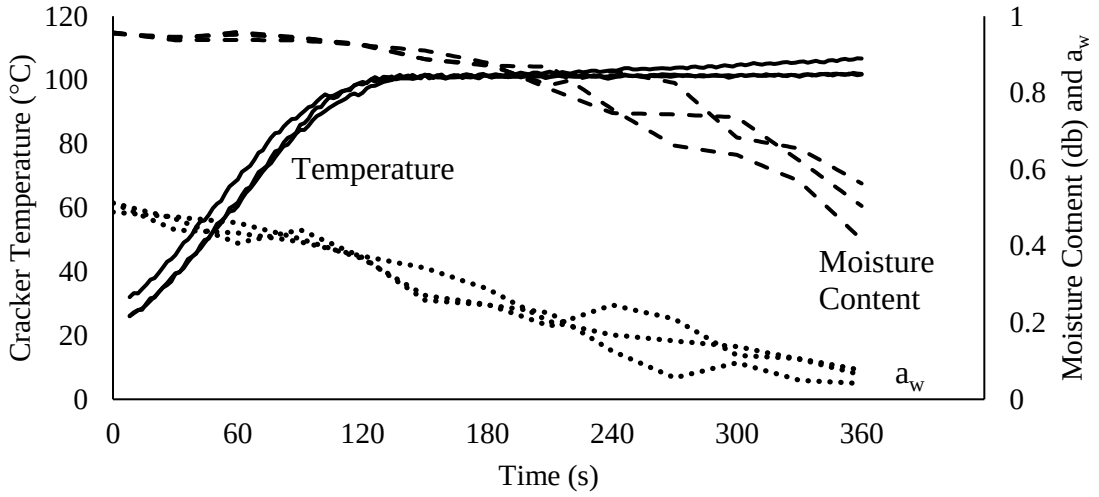


Figure 5.2. Transient cracker core temperature, moisture content (dry basis), and a_w (measured at 25°C) during triplicate 360 s cracker baking experiments.

5.2.3 Model Validation

The application of the *Salmonella* inactivation model for cracker dough to the observed temperature profiles of the cracker baking process, without considering dynamic a_w, rapidly

yielded model failure (Figure 5.3). While most of the observed *Salmonella* inactivation did accumulate during the first 90 s (~4 log reductions), during which the product moisture remained relatively high (>0.94 a_w, 0.42 db), the model predicted over 90 log reductions by 70 s. As a result, the RMSEs of prediction by 60, 90, and 360 s were 5.3, >40,000, and >2,000,000,000 log CFU/g, respectively. The validation metric mAPZ, which was inclusive of the mean predicted inactivation value ± 1.67 log CFU/g, identified 6 of 9 residuals that were neither fail-safe nor fail-dangerous by 60 s; however, after 60 s, all residuals were fail-dangerous. Prior studies that reported both *Salmonella* inactivation kinetics in dough and survival curves during baking processes did not report such model failures *(45, 46, 140, 141)*; however, those prior studies did not directly validate the dough-based predictive model against the results during baking, which is a critically important step that should be completed before applying such a model in actual process validations.

The time/temperature/moisture model (Eqn. 5.4) fit to the validation dataset resulted in hyper-specific fit to the validation dataset (Figure 5.3). With D_{ref}, z_T, and z_M values of 0.15 min, 2.00 x 10^8 °C, and 0.21 db, respectively, the model fit was sensitive to moisture content but insensitive to temperature. To illustrate the problematic nature of these parameter estimates, this estimates are not valid and should not be used beyond the illustration in this study. Although the baking data were insufficient to estimate robust time/temperature/moisture model parameters, the results do demonstrate that a moisture-inclusive pathogen inactivation model can account for the dynamic conditions occurring during baking, reinforcing that moisture is a critical factor that predictive inactivation models must consider before being applied to baking processes. Further experiments would be necessary to generate sufficient data to yield valid and robust time/temperature/moisture model parameters.

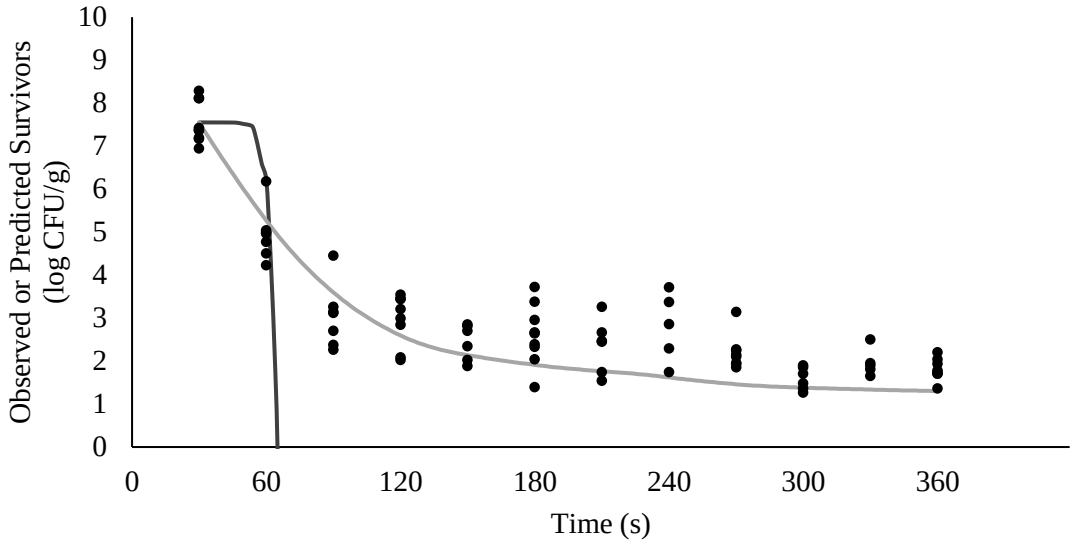


Figure 5.3. Predicted, fit and observed *Salmonella* survival curves during cracker baking. Predicted inactivation from the isothermal, single-moisture experiments (black line) and the dynamic temperature/moisture model fit (gray line) were integrated over the measured cracker profiles (temperature, or temperature and moisture content, respectively).

There is broad consensus in the literature that moisture impacts *Salmonella* inactivation kinetics, with decreasing moisture significantly increasing thermal resistance *(20, 72, 104, 120, 126, 153, 212, 213, 252).* Studies have reported that changes in a_w nearly instantaneously change *Salmonella* thermal resistance *(213),* and that the relative impact of temperature changes with a_w *(20, 126, 157)*.

Several recent studies have evaluated multiple inactivation modeling approaches to account for dynamic product moisture *(42, 126, 157, 212, 238)* and process humidity *(42, 120)* during dynamic, open-air processing of intermediate and low-moisture foods. One such study by Suehr et al. (2021) investigated the reduction of *Salmonella* in a low-moisture peanut-based cookie dough (a_w 0.25 ± 0.06) and reported a grossly fail-dangerous model prediction when

applying an isothermal and iso-moisture model to a dynamic baking profile that resulted in a final a_w of 0.19 ± 0.06. Suehr et al. was able to improve the model fit by performing a model regression using surface temperature and process humidity. The inclusion of such inputs for this study's cracker baking may have improved model performance. There is not yet a universal consensus on the best approach to account for moisture effects when predicting *Salmonella* inactivation in baked products *(2)*. The present study was not designed to provide a solution to that question (which were addressed in several prior manuscripts cited above), but rather to quantitatively and explicitly demonstrate that temperature-only inactivation models (which are currently distributed for use in the industry) are clearly insufficient for predicting *Salmonella* inactivation in these types of processes.

This study does not indicate that the *Salmonella* inactivation kinetics estimated in this study or others are inaccurate, but rather points to three key causes of model application failures that should be carefully considered before applying such models for validating pathogen reduction in baking or other low-moisture processes.

1. *Failure to estimate relevant factors impacting* Salmonella *inactivation kinetics*. The *Salmonella* inactivation model utilized in this study intentionally failed to quantify inactivation kinetics for conditions relevant to the process (e.g., surface temperature, a_w, oven humidity). As stated earlier, if sufficient *Salmonella* lethality accumulated while the cracker still was sufficiently high moisture (close to the raw dough), then the dough-based model may have been appropriate. However, a 5-log reduction of *Salmonella* was achieved experimentally only after the model predicted >100-log reductions. A more appropriate approach would be to characterize *Salmonella* inactivation kinetics at product/process conditions that encompass the realistic conditions experienced by

industrial-scale processes. This has been reported using point-specific inactivation kinetics at multiple discrete moisture content levels *(157)*, and using models accounting for continuously changing product moisture, a_w, or humidity *(42, 120, 126, 212, 223)*.

2. *Failure to measure critical factors with sufficient resolution*. The product factors measured in this study were core temperature in real-time and a_w of a post-bake cracker. The measurement of a_w may be inherently flawed when characterizing intermediate/low-moisture products sampled during dynamic processes, because a_w represents a bulk measurement (i.e., the cracker) of an equilibrium property, where baked products may have significant moisture gradients between the product surface and core. If the location of greatest pathogen survival (i.e., the "least lethal spot") is at the location of lowest a_w (the surface of the cracker), then any measurements including more than the surface may yield higher values than the "true" surface a_w. This likely is why the observed *Salmonella* survivor curve in the present study tailed after ~90 s, reflecting a significant decrease in the inactivation rate, despite the recorded full-cracker a_w of 0.94. Therefore, critical factors should be specific to the location of greatest concern within the product, which is not necessarily the cold spot. An example identifying/recording such measurements can be found in Jeong et al. *(119)*, where surface temperature and moisture (accounted for by process dew point) were measured and included in the inactivation model.
3. *Failure to quantify model adequacy with quantitative metrics*. Assuming that an adequate model framework was adopted, and the correct measurements taken, there are no standard metrics for validating model robustness in process validation applications. Two previously disseminated sets of metrics for evaluating inactivation model validations include prediction bias/accuracy factors (B_f/A_f) and the APZ *(192, 205)*. Both sets of

metrics were originally developed for application to validation of growth model parameters and were later adapted for evaluating inactivation model residuals *(79, 190)*. While B_f, A_f, and APZ metrics include designations such as "fail-safe" or "fail-dangerous," the determination of what constituted a model failure was partially arbitrary and did not consider the expected level of variability for the application. For example, consider a scenario where the inactivation of a pathogen in a product had an inherent variability of 0.8 log CFU/g RMSE. A perfect model would result in the percentage of residuals within the APZ of 63%, which, when applying the APZ validation criteria (requiring >70% residuals within the APZ), would result in a failed model validation *(190)*. To overcome this issue, the proposed mAPZ accounts for variability of inactivation outcomes during model development (RMSE) and the potential impact of scale-up (ϕ) during a validation test.

Overall, based on the results in this study (and other baking studies), baking processes like the one tested here are most likely sufficient to achieve a target level of pathogen control (>5 log reductions). However, the present study was designed to explicitly demonstrate the potential hazards of applying a predictive model based on data from an isothermal and a single-moisture experiment to validate a dynamic baking process, and to warn against the potential misuse of such models to erroneously conclude that an "under-processed" baked product is microbiologically safe. The previously disseminated application approach of such models was proven ineffective at predicting *Salmonella* inactivation, yielding significantly fail-dangerous predictions ~70 s into a 360 s baking process. Future studies with model validations in low-moisture processes should consider the relevance of inactivation model development conditions,

the spatial resolution needed for adequately characterizing the location(s) of concern for inactivation model application, and the tools/metrics needed for validating inactivation model performance in real-world applications.

5.3 Acknowledgements

The work was supported by the U.S. Department of Agriculture, National Institute of Food and Agriculture, award 2015-68003-23415.

CHAPTER 6: IMPROVING THE UTILITY OF SURROGATES INTENDED FOR FOODBORNE PATHOGEN PREVENTIVE CONTROL VALIDATIONS

This opinion paper has been accepted for publication in the journal *Current Opinion in Food Science*, and focuses on the gap between characterizing surrogates and successfully incorporating that information into preventive control validation.

6.1 Premise

Beyond the concepts of "good bacteria" and "bad bacteria" that are often associated with food, there is another group of microorganisms that can be considered "useful bacteria." Different classifications of utility microorganisms were described by Busta et al. *(39)*, including surrogate microorganisms, which are the topic of this review. A surrogate microorganism may be used in the place of a pathogen to evaluate the effectiveness of an intervention during a pilot- or industrial-scale process study.

The Food Safety Modernization Act (FSMA) preventive controls rules *(234)* require establishing food safety systems that incorporate hazard analysis and preventive controls (hazard reduction) in addition to several oversight functions. If there is a known or reasonably foreseeable risk of a pathogen-contaminated food product reaching consumers and causing illnesses, then there must be preventive control(s) that reduce/minimize that risk, either through prevention or population inactivation/reduction. Using a valid surrogate to estimate potential pathogen reductions has been an FDA-accepted validation methodology for over two decades *(114)* when other methods are not feasible *(43)*. However, there is a need for improved methods to optimally design and analyze process validation studies that utilize surrogate organisms. As of this review, the most recent guidance from the FDA does not yet include how to validate processes for pathogen control *(235)*.

There are several reviews that define and discuss the characteristics of good surrogates for pathogen control studies, including a general review (2003) by Busta et al. *(39)*, a review focused on surrogates for use in pathogen challenge studies (2017) by Hu and Gurtler *(110)*, and a review providing guidance on validations with important information on using surrogates (2021) by Ceylan et al. *(43)*. However, the state of practice for surrogates is still evolving, with several gaps remaining that limit the utility of surrogate studies. Furthermore, the perspectives on surrogate applications is broadening as studies begin to examine indirect food applications, such as dry sanitation of food processing equipment *(49)*. Specifically, we will evaluate the current state of practice and recent developments for surrogate microorganisms intended for use in preventive control validations, based on a systematic analysis of recent (59 studies from 2018-2023) literature, with particular focus on factors that are limiting the translation of surrogate utility from research to industry-implementation, such as inconsistent benchmarking of a surrogate's resistance against the pathogen and the lack of a standard statistical framework for qualifying and applying surrogates for process validation.

6.3 General Surrogate Criteria

The qualitative characteristics of a "good" surrogate have not changed substantially since the Busta et al. *(39)* review, and have been adapted by the surrogate-focused reviews published since *(43, 110)*. We will particularly emphasize two criteria for a "good" surrogate:

1. The resistance characteristics of the surrogate to the mode(s) of inactivation during a process control should allow for a predictable relation to the mean and variability of the corresponding pathogen's reduction.
2. The specific environment and mode(s) of reduction, or conditions reasonably similar, should have been used as supportive evidence for the surrogate's appropriateness.

Most other requirements *(39, 43, 110)* focus on the practical considerations of surrogate use, such as it must be non-pathogenic, easy to use, and does not cause problematic side-effects, such as promoting spoilage. The first requirement highlighted above is the requirement that the surrogate must be predictive of the pathogen's population, which is the ultimate purpose of a surrogate. The latter requirement was often stated but not included on previous lists; however, it is a key consideration for those utilizing surrogates for process validations, as they must rely on the existing body of research as supporting scientific evidence for the appropriateness of the surrogate and the interpretation of any validation study results.

Recent reviews of surrogate-related research focused on reporting comparisons of surrogates to pathogens of interest for a variety of food matrices *(2, 65)*. Most recent surrogate research has been dominated by work with low-moisture products and a specific *Enterococcus faecium* strain with clonal copies that can be sourced from collections with the USDA ARS (NRRL B-2354) and the American Type Culture Collection (ATCC 8459). The Almond Board of California established a precedent when *E. faecium* NRRL B-2354 was accepted as a *Salmonella* surrogate for almonds with thermal processing technologies *(13)*.

The continued proliferation of low-moisture food surrogate studies may be a response to the industry's need for a robust validation strategy that can overcome the complexities of dynamic heat and mass transfer processes, which limit the potential for real-time measurement of critical variables needed for the application of predictive microbial models for process validation. The continued focus on *E. faecium* NRRL B-2354 may be attributed to the fact that it has been shown to meet most of the characteristics of an appropriate surrogate in various environments where *Salmonella* is identified as the pathogen of concern, although not in all cases, depending on specific products/technologies *(4, 65)* or how the criteria defining an

appropriate surrogate are interpreted. Additionally, the perceived cost barrier of identifying any new/alternative surrogate likely limits such efforts, making the robustness of *E. faecium* NRRL B-2354 as an acceptable surrogate in a variety of low-moisture foods an important ongoing result.

6.4 Systematic Review of the State-of-practice

To rigorously evaluate the state-of practice for evaluating, qualifying, and utilizing surrogates for validating pathogen reductions, we sourced studies from the Web of Science (Clarivate, PA, USA) database by searching combinations of "pathogen," "surrogate," and the names of common pathogens/surrogates (e.g., *Salmonella*, *Listeria innocua*, *E. faecium*). This search yielded 59 studies (Appendix D) that evaluated surrogate and pathogen reduction resistance for the purpose of qualifying a surrogate as appropriate (published in 2018-2023). We analyzed the results of those studies to quantitatively characterize the various ways that surrogate appropriateness was defined and, more importantly, quantified, as the basis for our subsequent analysis, discussion, and proposed improvements in the field. It should be initially noted that these studies included a diverse range of pathogens (different species, different strains, individual-strains vs. cocktails, etc.), multiple candidate surrogates, a variety of food matrices and processes, and diverse approaches to characterizing the pathogen-surrogate relationship (Appendix D), all of which can affect interpretation and utility of the reported information, as discussed in more detail below.

Most of the surrogate studies analyzed focused on *E. faecium*, *Salmonella*, and low-moisture foods. However, a few studies examined the appropriateness of *E. faecium* NRRL B-2354 / ATCC 8459 as a potential surrogate in high-moisture foods *(52, 124, 147, 210, 221, 229)*, although there is little history of using *E. faecium* as a surrogate in such products *(18, 158, 199)*.

E. faecium was also evaluated as a potential surrogate for other pathogens, such as *Listeria monocytogenes (24, 25, 64, 147, 177, 201, 211, 273)* and *Escherichia coli* pathotypes *(24, 63, 105, 177, 208, 273)*, although most of these studies were in conjunction with *Salmonella enterica*, with a range of different individual serovars or cocktails of multiple serovars, which can affect the comparisons. While nearly all recent studies used clonal copies of *E. faecium* NRRL B-2354 or ATCC 8459, one study used a strain packaged as part of a lactic acid bacteria collection *(135)*. Given the surrogate-product-process specificity required for preventive control validations, data from this study cannot be merged with the broader body of knowledge developed around NRRL B-2354 / ATCC 8459.

Since 2020, a few papers have examined potential pathogen inactivation surrogates that were not *E. faecium*. These included *Klebsiella aerogenes* (formerly known as *Enterobacter aerogenes*), *Pantoea dispersa*, *Deinococcus radiodurans*, *Geobacillus stearothermophilus*, meat starter cultures, and lactic acid bacteria *(3, 63, 105, 133-135, 208, 211)*. It was common for these studies not to examine the full criteria for an acceptable surrogate; some were examined just for their known high thermal resistance or their association with the environment studied. These studies often focused on comparing inactivation resistance, and not all the potential surrogates listed were reported as appropriate for consideration in the scenarios tested.

Beyond *E. faecium* and the less common surrogates, numerous studies included avirulent forms of pathogens or closely related nonpathogenic bacteria as potential surrogates. This includes avirulent *Salmonella*, nonpathogenic *Escherichia coli*, and *Listeria innocua (25, 63, 69, 101, 105, 135, 208, 209, 211, 268, 273)*. It is our opinion that pathogen-adjacent microorganisms should not be used as surrogates for in-plant, commercial process validations, as the inclusion of microorganisms of the same species or genus as pathogens may lead to problematic side effects,

such as contributing to false positives in future environmental monitoring *(110)*. These microorganisms, or any that have detrimental effects if introduced to a processing environment, would be considered nuisance microorganisms. Surrogates that are avirulent to humans, but not nonpathogenic for all forms of life (e.g., *Pantoea dispersa* has been reported as a potential opportunistic pathogen to plants and humans *(272)*), may also be considered inappropriate for this reason, depending on the context.

The current state of surrogate readiness as process validation tools, in general, has focused on verifying that the likely surrogate is appropriate for specific products and processes, rather than identification of new surrogates. However, there are two unresolved critical issues in the current state of surrogate literature, which are discussed in the following two sections.

6.5 "Greater or Equivalent Resistance"

The relative mean effect of the pathogen vs. surrogate resistances may be expressed as a reduction ratio (log reductions of pathogen per log reduction of surrogate), which results in reduction ratios ≥1 translating to a surrogate of greater or equivalent resistance. Additionally, the relative population variability resulting from inherent/experimental artifacts or the treatment can be expressed as a variability ratio (standard deviation of pathogen survivors divided by the standard deviation of surrogate survivors). No known studies explicitly define the acceptable ranges of reduction or variability ratios, or any similar metric.

The typical objective of the 59 studies analyzed was to determine the appropriateness of the reduction ratio for use in future surrogate-based validations; however, there was no common or uniform definition of the ideal reduction relationship (Figure 6.1). Of the studies that included a pre-hoc description (in the introduction or methods sections) of the appropriate surrogate-

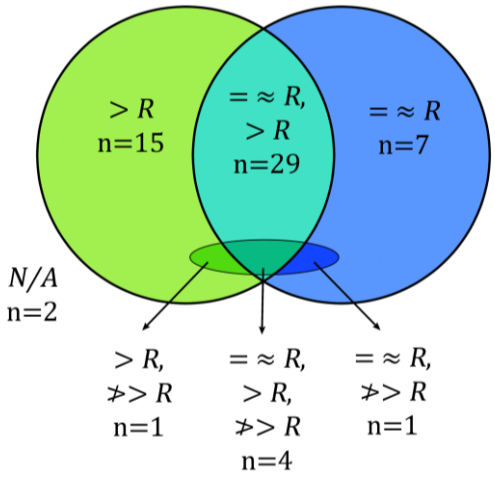


Figure 6.1. Number of studies in the analyzed literature that stated that an appropriate surrogate would have an equivalent/similar (=≈R), greater (>R), and not substantially greater (⊁>R) resistance to the mode of reduction when compared to the associated pathogen. These values represent the totality of the studies and are not separated by pre- and post-hoc statements.

pathogen relationship (21/59), the only criterion applied was that the surrogate had greater or equivalent resistance (reduction ratio ≥1), without addressing potential upper limitations for the reduction ratio. Most studies reported post-hoc descriptions when contextualizing their results, explicitly or implicitly indicating that the surrogate should be of equivalent (7/59), of greater (21/59), or of equivalent/greater (24/59) resistance. Two studies did not state the desirable pathogen-surrogate relationship or only used ambiguous language *(178, 267)*.

Few studies warned of overly resistant surrogates *(3, 4, 60, 105, 155, 273)*, always in combination with other post-hoc descriptions of appropriate reduction ratio characteristics (Figure 6.1). Of the studies raising concern about over resistance, it can be inferred that a reduction ratio >2.5 may be considered excessively resistant. Henz et al. *(105)* stated that *E.*

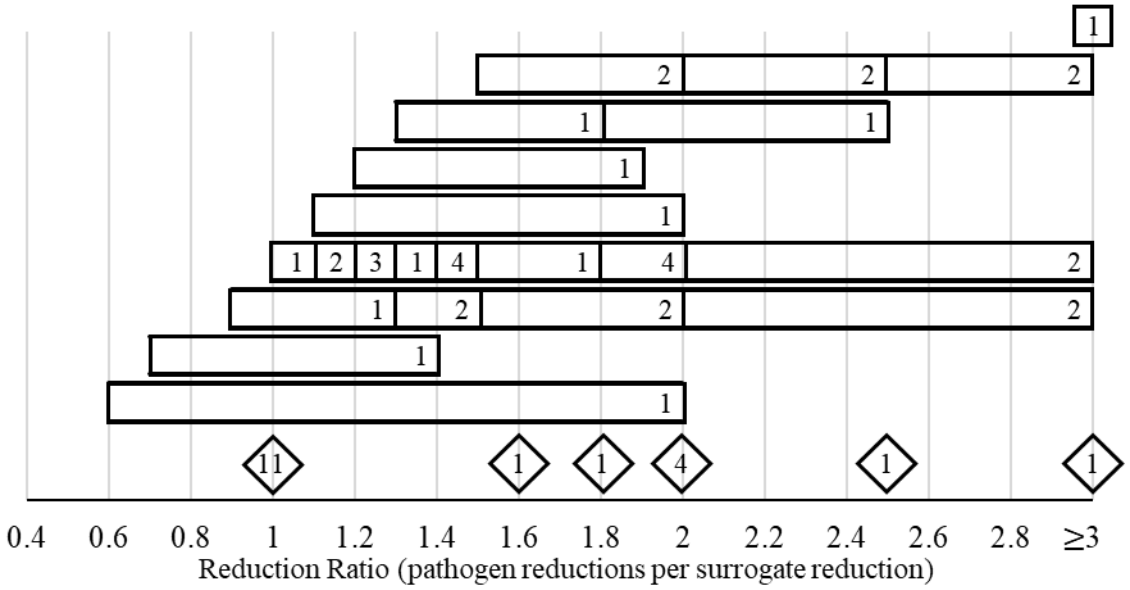


Figure 6.2. Frequency of reduction ratios in the analyzed literature, as reported or identified by the authors of this review. The number of studies with a single reduction ratio value are listed in diamonds. Studies containing a range of reduction ratios are grouped by the common lower limit, with the frequency located near the upper limit section divider or cap. Two studies contained a single reduction value or a range of reduction values >3.

faecium was too resistant as a *Salmonella* surrogate (reduction ratio 2.5 – 3.0) in pumpkin and flax seeds when treated with electron beam radiation. Zhou et al. *(273)* also described *E. faecium* as too resistant as a *Salmonella* surrogate (reduction ratio 2 – 2.5) when treating basil leaves with hot air. Both studies suggested different nonpathogenic *E. coli* as a better surrogate (reduction ratio ~1.1). Two studies expressed concern about overly resistant surrogates without disqualifying the surrogates researched *(3, 4)*. If this upper limit (≤2.5 reduction ratio) was applied to other studies, then surrogates reported in some studies as appropriate would become disqualified in part, or in the entirety, of the treatment conditions studied *(4, 50, 100, 178)*. Most studies reporting an appropriate surrogate outcome yielded reduction ratios between 0.9 – 2 for

at least a portion of the conditions tested (Figure 6.2). Studies that resulted in ranges of reduction ratios (38/59) often included more than one treatment level (i.e., temperature, water activity, concentration), which may imply that a surrogate is not universally appropriate for the product and process examined. It could be that if different or more treatment levels were evaluated, then a different conclusion would have been reached regarding the surrogate's appropriateness; however, given the current ambiguity on reduction ratio limits, the conclusions on surrogate appropriateness would still be subjective.

The most common method to quantify the reduction ratio was via comparison of predictive lethality model parameters (33/59; Figure 6.3); because more than one treatment level was often included, this typically consisted of comparing log-linear primary and Bigelow-type/response-surface secondary model parameters. The next most common quantification method involved comparing the mean effects of a static treatment (26/59) at different treatment levels. Although most parameter or treatment comparisons were statistical, having performed tests like the Student's t-test, F-test, or ANOVA (or at least included minimal statistical information about the comparisons; 47/59), some studies only performed arithmetic or visual comparisons to qualify surrogate appropriateness *(15, 51, 60, 63, 69, 101, 134, 135, 231, 255, 260, 263)*. The comparisons of model parameters, treatment outcomes, or reduction ratios (9/59) all represent comparisons of mean effects. No known prior studies statistically evaluated the potential differences in the variability of reduction, although some studies represented combined mean and variability effects in industry-relevant terms *(3, 177, 247)*.

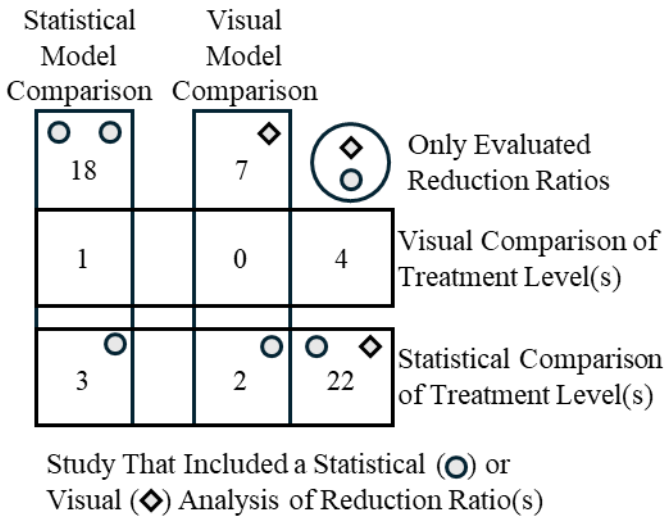


Figure 6.3. Distribution of analysis methods used to compare or quantify the surrogate-pathogen reduction relationship in the analyzed literature.

The current state of surrogate research is largely focused on verifying *E. faecium* NRRL B-2354 / ATCC 8459 appropriateness for various pathogens, products, and technologies; however, the benchmarks used to qualify the appropriateness of the surrogate's resistance are variable and may be insufficient to fully inform preventive control validations. A surrogate that is of "equivalent or greater resistance" compared to the pathogen, without any characterization of a variability relationship, has unknown statistical reliability or power when applied in actual process validations. Ultimately, the impact of any given surrogate-pathogen relationship depends on preventive control validation practices.

6.6 Surrogates as a Tool for Validation Beyond Just Research

In the current state of research, the objective of many academic surrogate studies is to determine only one aspect of surrogate appropriateness through comparing the relative resistances of the surrogate and pathogen. The objective of surrogate-based process validations in industry is to determine whether the process meets the performance criterion. The design of

academic studies and the communication of the subsequent results are not always aligned with the objectives of industry-scale application, resulting in information that may not completely address the utility or reliability of surrogates. One such deficit was addressed in the previous section, with studies commonly underreporting information needed for a statistically robust validation.

Another contributing factor to the disconnect between surrogate evaluation and surrogate application may be from ambiguity present in resources that support surrogate-based validations. As stated previously, there are numerous resources that provide listed, generally-accepted criteria for determining the appropriateness of a surrogate, but these resources are also ambiguous about how to quantitatively interpret surrogate-based validations or challenge studies *(39, 43, 110, 114)*. Before the US FSMA Preventive Controls for Human Foods Rule, the US FDA endorsed using surrogates for a variety of inactivation technologies *(114)*; however, that publication acknowledged that surrogate research was relatively new and reiterated the need for statistically sound methodologies. Recent reviews have significantly updated what is needed to execute a surrogate-based validation; however, none have defined the interpretation of such results beyond stating the importance of using "appropriate" statistical analyses and methods *(39, 43, 110)*. This ambiguity forces the industry to assume responsibility over acceptance of the validation results. Without an accepted standard around which to structure surrogate studies or their results, there will always be some disconnect between surrogate research and surrogate-based validations.

With that said, many recent surrogate studies still fall short of what is known to be necessary for industry-based application. Surrogate-based validations need to statistically support sufficient surrogate lethality (typically ≥4 log reductions), which then can be translated to achieving the target food safety outcome for the pathogen. A few comparison studies reported

pathogen inactivation that did not exceed an average 4 log *(203, 221, 255)*, despite reporting that the surrogate investigated would be valid for application. If the processor's food safety objective is to demonstrate that sufficient samples exceed the target reduction (e.g., 95% samples received ≥4 log reductions for the pathogen), then the processor will need to translate the mean lethality achieved, which can be done using reduction ratios, and the variability of the achieved lethalities. However, no known surrogate studies have statistically compared pathogen and surrogate population variabilities or developed a variability ratio, which is needed to interpret the validation study, in terms of uncertainty and statistical power.

Just a few studies communicated both mean differences and variability in industrially useful terms *(3, 177, 247)*. Moussavi et al. *(177)* demonstrated that a mean lethality of the surrogate greater than 4-log translated to ≥95% of samples likely achieving a 4-log pathogen reduction by using prediction intervals on the observed reduction ratios. Prediction intervals describe the likely spread of data around a mean estimate, which can be for a static treatment or for a predictive model. Generally, prediction intervals can be represented with Eqn. 6.1

$$\mu \pm t_c S\sqrt{1 + {}^{1}/_{n}} \qquad \text{Eqn. 6.1}$$

where μ represents the mean estimate, t_c is the reference Student's t-value for the desired percentile, S is the measure of variability (such as standard deviation or root mean square error), and n is the degrees of freedom. Following a similar approach, Verma et al. *(247)* and Acuff et al. *(3)* used prediction intervals to identify the conditions where surrogates would result in conservative estimates of the pathogen inactivation for at least 95% of samples collected. Murdoch et al. *(178)* included a surrogate-comparison component, but then also included a detailed approach to a low energy electron beam validation. Although that study considered processing variability, *E. faecium* survivor variability, and emphasized "worst case scenarios,"

they did not consider the inherent *Salmonella* variability in response to the treatment in the predication for the target *E. faecium* lethality, and later reported that the 8.33 reduction ratio also translated to an impact on variability, a potentially erroneous assumption. Several recent studies reported strategies for translating average surrogate reductions to average pathogen reductions using linear regressions of the observed reduction ratios *(21, 268)*; however, no information was provided on translating the expected levels of variability, and therefore, the inherent uncertainty in the predictions.

Without providing the information for translating the mean and variability, researchers may anticipate the industry approach to using surrogates is to assume a 1:1 relationship to the pathogen, which does not allow for ideal interpretation of validation results (Figure 6.4). Within Figure 6.4, two surrogate-pathogen relationships were examined using a reduction ratio of 1.2 (pathogen reduction as blue line) and 2.0 (pathogen reduction as red line), both with variability ratios of 1. Possible interpretations of the translation methods (Figure 6.4, A-C) are provided given a food safety criterion of ≥95% of the pathogen's population exceeding the target lethality; black, blue, and red arrows point to the processing time needed to achieve ≥5 log reductions for the surrogate, the pathogen given a 1.2 reduction ratio, and the pathogen given a 2.0 reduction ratio, respectively.

A 1:1 pathogen-surrogate relationship would imply that both microorganisms share identical resistance *and variability* characteristics across all relevant treatment conditions. This is a convenient framework for analyzing data, as it makes the translation from surrogate results to potential pathogen result trivial, if the assumptions are appropriate. Incorporating a sufficiently conservative surrogate may overcome the uncertainty from not knowing the variability

relationship. For the convenience of the 1:1 relationship to work, one of two conditions must be true.

1. The pathogen and the surrogate must exhibit a true 1:1 correlation in response to the processing environment. If true, it would be essentially a rare coincidence that two unrelated microorganisms happen to respond in identical ways, especially considering that even closely related microorganisms have been known to respond in significantly different ways.
2. The process being validated must be excessively effective at reducing the pathogen, to make it acceptable to evaluate inactivation results from a conservative surrogate as if it were the pathogen of concern.

Processors cannot rely on the identification of a "perfect" surrogate, especially when most of the surrogate research is based on a monoculture (42/59), using only clonal copies *E. faecium* NRRL B-2354 / ATCC 8459. Therefore, processors using the 1:1 assumption must rely on an excessively effective process (e.g., using a 1:1 assumption on a surrogate with a reduction ratio of 2 to support the performance criterion of 4 log corresponds to a true pathogen reduction closer to 8 log). Considering this, the results of a failed validation (surrogate reductions fail to exceed pathogen's target reduction) do not necessarily mean that the processors are not meeting the performance criterion for the pathogen; it just means that they are not meeting the performance criterion if the 1:1 assumption is perfectly correct.

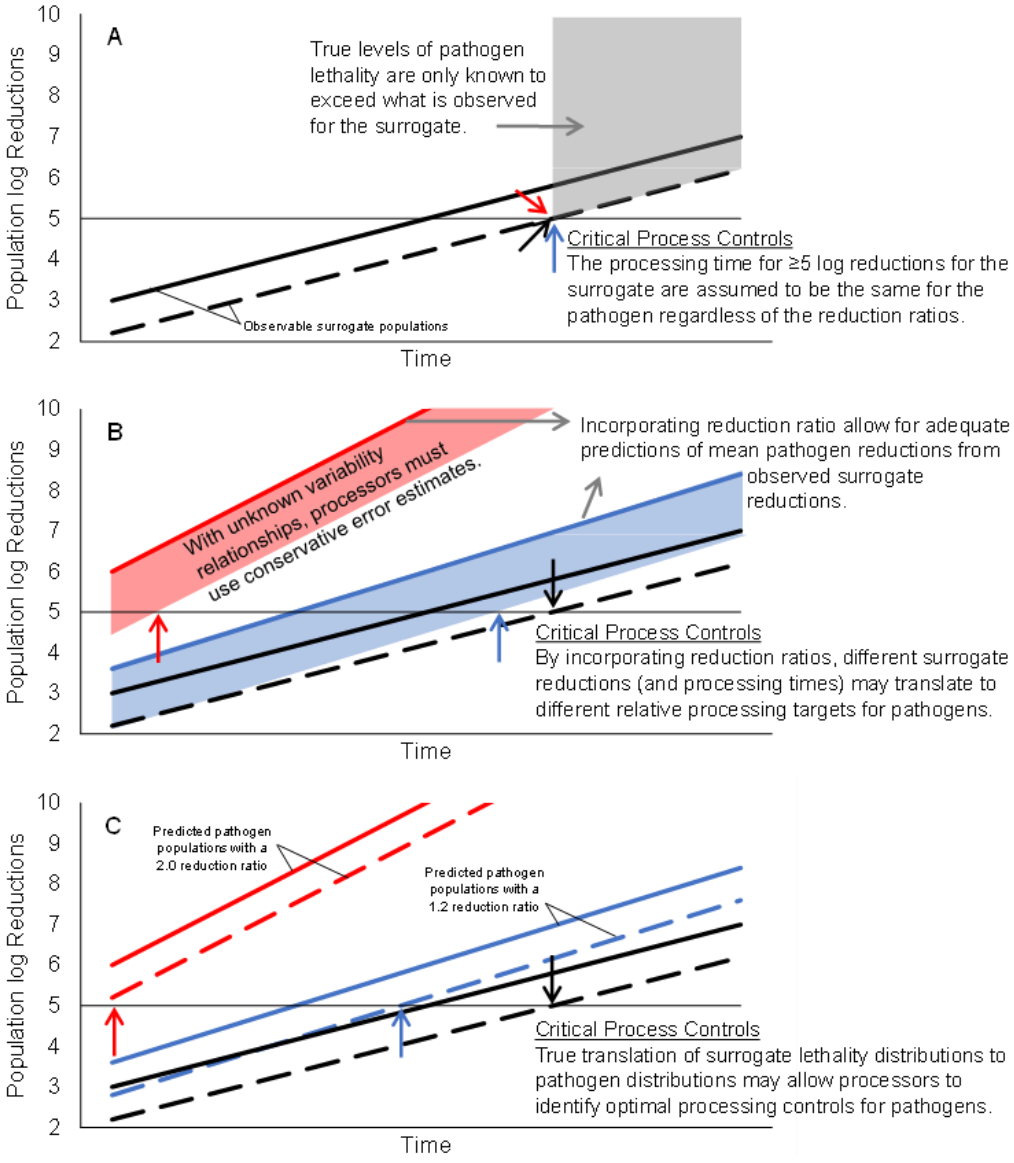


Figure 6.4. Illustrations of different possible approaches to translating surrogate data (black) into predicted pathogen mean (solid lines) and 95% lower prediction interval (dashed lines) outcomes, including: (A) the "1:1 assumption," (B) accounting only for the surrogate-pathogen reduction ratio, and (C) an ideal approach incorporating both reduction and variability ratios.

For processes that are excessively effective at reducing the pathogen, a convenience framework for analysis may be used. For processes that are marginally surpassing the

performance criterion, or even if they are just not sufficiently excessive, the assumption of a 1:1 relationship is not convenient. These processors may need to overprocess their product, include additional preventive controls, or consider changing their processing technology altogether. Most of the studies reviewed (37/59) included a reduction ratio ≥1.5 for at least part of the conditions studied (Figure 6.2), meaning that validations that are marginally surpassing the performance criterion for these surrogates are likely achieving excessive levels of pathogen inactivation (Figure 6.4A). If a process is meeting the food safety objective, then the means of analysis should be able to support this outcome with high accuracy. By only translating surrogate outcomes with the reduction ratio, processors can more accurately understand when the process is achieving the performance criterion for the pathogen (Figure 6.4B) but would need to use conservative estimates of variability to overcome that unknown. If enough is known about the reduction ratio, variability ratio, and the preventive control, then a statistical framework that accurately translates surrogate outcomes to pathogen estimates should be possible (Figure 6.4C). A statistical framework for validation analysis would provide the mathematical relationships to better define the boundaries of what defines an appropriate surrogate (e.g., the reduction ratio boundary of an overly resistant surrogate). Additionally, such a framework may allow for dissimilar target reductions for the pathogen and surrogate, a more robust methodology than what is proposed by some studies *(21, 178, 268)*. More flexibility in the application of surrogates would increase the utility of existing surrogate studies and reduce the pressure to identify a "perfect" surrogate for every product and technology combination.

6.7 Conclusions and Research Needs

Although the purpose of most surrogate research is to advance the evaluation and utilization of potential surrogates for industry preventive controls, a pervasive disconnect exists

between surrogate studies and surrogate-based validations. Guidelines developed by academic, industry, and government sources remain excellent resources for most of the details needed to execute a validation study *(39, 43, 110, 114)*; however, insufficient details are provided regarding how to interpret the results from a validation study beyond emphasizing the need for "appropriate" statistical analyses. Additionally, there are no standards of communication for surrogate research studies to translate their results to industry application. Studies that communicated a reduction ratio without a known grading standard often yields a qualification, and not a quantification, of a surrogate's appropriateness. Most studies fail to quantitatively evaluate the surrogate-pathogen variability relationship, which is needed for a complete statistical evaluation of validation results.

The disconnect between surrogate research, surrogate-based validations, and the various guidelines for surrogate use can be bridged with the development of a robust statistical framework for interpreting surrogate validations. A framework that translates surrogate results to potential pathogen results, while maintaining statistical integrity, would enable processors to incorporate statistically appropriate methods while enabling use of surrogates that do not comply with the ideal 1:1 relationship. Furthermore, researchers would be able to consider the inputs needed for utilizing such a framework when quantifying surrogate appropriateness and communicating results.

Ultimately, the most needed innovation in the surrogate field is a quantitative and standard statistical framework for interpreting surrogate data, to improve both research into and utility of surrogates as robust validation tools. Just as previous guidelines outlining surrogate appropriateness originated from academic, industry, and/or government efforts, similar

collaborations are needed to test a standard statistical framework rigorously and scientifically before wider adoption within the industry or incorporation into regulatory frameworks.

6.8 Acknowledgements

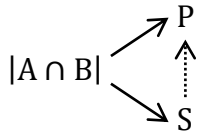
This work was supported by the Agriculture and Food Research Initiative, Sustainable Agricultural Systems Program grant no. 2020-68012-31822 from the USDA National Institute of Food and Agriculture.

CHAPTER 7: STATISTICAL FRAMEWORK FOR SURROGATE-BASED VALIDATIONS OF PREVENTIVE CONTROLS BASED ON PREDICTED PATHOGEN REDUCTIONS

This study is the direct follow-up to the review presented in the previous chapter. The chapter presented here is not the complete version of the manuscript to be submitted to a peer-reviewed journal, to eliminate redundancy between chapters. When reviewing the state-of-practice for surrogate-based validations of preventive controls (Chapter 6), two major issues were identified: (1) the lack of statistics-based metrics and determination of performance criteria, and (2) how the target reduction for controls are often based on pathogen reduction goals but are evaluated using surrogate-based data. The statistical framework proposed in this study seeks to resolve both issues. Additionally, Chapter 8 presents a case study utilizing the framework proposed in this chapter.

7.1 Background

Nonpathogenic surrogate microorganisms should mimic the behavior of the target pathogen when validating preventive controls.



For any appropriate surrogate application, product (A) and treatment (B) both impact the pathogen (P) and surrogate (S) where the observed behavior of the surrogate infers the behavior of the pathogen. While the surrogate-pathogen relationship can be investigated for a variety of products and treatments, the causal effect of both elements ($|A \cap B|$) need to be explicitly investigated for the surrogate-pathogen relationship to be considered acceptable.

The Food Safety Modernization Act (FSMA) Preventive Controls Rules *(234)* require that processors establish food safety systems that incorporate hazard analysis and preventive controls (hazard reduction). Within this requirement, processors are responsible for the risk management decisions, which includes defining the preventive control performance criterion to fit within the food safety objective and setting the acceptable level of risk. Because processors are required to define their performance criterion, guidance on surrogate-based validations may be hesitant to prescribe specific statistical tests beyond suggesting "appropriate statistical analyses" *(39, 43, 110)*. There are two indications that reduction distribution-based analyses (i.e., large percentage of samples exceeding the target lethality) may be the preferred metric for the performance criterion. First, while not statistical, metrics based on "minimum reduction" have the explicit purpose of verifying that a large proportion of reduction outcomes, represented by the surrogate dataset, exceed the target lethality *(13, 43)*. Ceylan et al. (2021) makes a case that performance criteria based on minimum reduction cases (MRC) may be superior to confidence interval-based tests because MRC estimates would be more conservative and therefore would ensure that a larger fraction of product would achieve the desired risk reduction.

The second indication that reduction distribution-based tests may be a preferred performance criterion metric can be identified when reviewing risk assessments. For risk assessments examining salmonellosis from U.S. consumption of nuts (almond, pistachio, walnut, pecan, peanut) the most common description for preventive control reductions was a uniform treatment *(41, 61, 74-77, 142, 143)*, meaning that it was assumed that 100% of the samples received the designated reduction. A few cases examined the impact of preventive control reduction variability, but ultimately reinforced standing regulations for minimum reduction targets in almonds *(61, 143)* or highlighted that a uniform 4-log reduction achieved the desired

food safety objective in pistachios *(142)*. Each of these studies indicated that process variability was unknown, which likely contributed to the assumption of 100% of the product receiving the target reduction from the preventive control in most analyses. When examining reductions as a continuous normal distribution, it is theoretically impossible for 100% of the product to achieve a non-zero target reduction. Therefore, it makes the most sense to interpret 100% of product achieving the target reduction as some number close to 100%, such as 95%, of sample achieving or exceeding the target reduction.

The statistical determination that a preventive control results in the target lethality for a percentage of a population can be performed using a tolerance interval, a method for estimating confidence bounds for a proportion of data in conformance with a specification *(170)*. While tolerance intervals often are associated with engineering specifications or pharmaceutical/clinical fields, tolerance intervals have been incorporated as a validation metric in analytical chemistry research *(81, 82)*. However, that example is not a direct analogue, because it focuses on direct measurements rather than indirect measurements using surrogates. The authors know of only one other example of tolerance intervals being applied to surrogate data. Wang et al. (2017) proposed using tolerance intervals to improve the utility of nuclear power plant surrogate safety goals; in that study, radiation-emitting events were used as a surrogate for adverse health effects on the surrounding population *(254)*. As a tool for engineering operation controls, tolerance intervals may be utilized in the monitoring of indicator organisms; but these analyses may stand alone and do not necessarily involve translation to the microorganisms of concern.

The implementation of tolerance intervals for the evaluation of surrogate data against a pathogen-based performance criterion may be an improvement over commonly used analysis methods. However, we will make the case that a statistical framework may be adopted that

would allow the estimation of pathogen-based outcomes to be tested against the pathogen-based performance criterion. This translation of surrogate results to estimated pathogen results will be the focus of the statistical framework development. However, there are other criteria that must be satisfied before considering a surrogate appropriate for a pathogen of concern *(39, 43, 110)*. For this reason, there was an assumption that all other required elements for a surrogate to be investigated and considered acceptable were satisfied, for this framework to be successfully utilized. Therefore, the utility of this statistical framework will then depend primarily on whether the right assumptions can be made for the statistical test and whether there is a reasonable model for translating surrogate information to pathogen information.

7.2 Tolerance Bounds

The purpose of a tolerance interval, or a tolerance bound when considered from 1-side, is to describe the probability of a value existing above or below a tolerance limit. Because we are interested in determining whether *q* percentage of reduction outcomes for the treated population exceeds a target lethality (*TL*), we can use the equation for estimating a tolerance bound (Eqns. 7.1 and 7.2)*(186)* and define a validation performance criterion using Eqn. 7.3.

$$X_q = \bar{X} - k_1\sigma \qquad \text{Eqn. 7.1}$$

$$k_1 = \frac{z_q + \sqrt{z_q^2 - ab}}{a}$$

$$a = 1 - \frac{z_\alpha^2}{2(n-1)} \qquad \text{Eqn. 7.2}$$

$$b = z_q^2 - \frac{z_\alpha^2}{n}$$

$$\Pr\,(X > TL) \geq q \qquad \text{Eqn. 7.3}$$

The population of reduction outcomes X (i.e., "log reductions") is normally distributed with a mean of $\bar{X}$, a standard deviation of σ, and a lower $100(1\text{-}q)^{th}$ percentile of X_q. The scale factor k_1 accounts for the uncertainty in the mean and spread of data and is a function of the number of samples (n) and the standard scores z_q and z_α for the q percentage of population described in the performance criterion and α confidence level, respectively. This estimation of the tolerance bound requires $n > 10$ samples and has n-1 degrees of freedom. Typically, in prior literature, the variables q and σ would be described as p and s, but the q and σ notation is used here to avoid confusion with pathogen and surrogate designations, respectively.

This description of the tolerance bound requires that samples and errors are normally distributed, independent, and uncorrelated *(170)*. Homogeneous naturally contaminated and inoculated samples are often assumed to follow a lognormal distribution of cells *(128)*, requiring any analysis of pathogen or surrogate populations to be log-transformed to assume a normal distribution. During a process validation test, the surrogate reductions can be assessed for normality; however, normality must be assumed for the potential pathogen population. With the equation for a tolerance bound (Eqn. 7.1), there are only two variables that are inherent to the population being studied, the mean and standard deviation. If we can assume the same level of uncertainty determined by the number of surrogate samples for the pathogen population, then translation of surrogate outcomes to pathogen outcomes would depend on the successful estimation of the mean and standard deviation for the pathogen.

7.3 Framework Foundation Part One: Translation of Mean Outcomes

With the goal of enabling evaluation of pathogen-based performance criteria using estimated pathogen outcomes, there needs to be a method for estimating pathogen mean reductions and standard deviations. Fortunately, it is standard practice for surrogate-pathogen

comparison studies to make surrogate utility determinations based on mean reduction comparisons (as was demonstrated in the previous chapter). While quantification methods can vary from study to study, the reduction ratio (*R*) can be used to describe the expected pathogen log reduction for every 1 surrogate log reduction (Eqn. 7.4).

$$\overline{X_P} = R\overline{X_S} \qquad \text{Eqn. 7.4}$$

The reduction ratio has been used to describe how conservative a surrogate is, with desirable ratios ≥1. Whether this reduction ratio *R* has any potential for unbiased translations of reduction outcomes for use in statistical tests depends on the quality of the reduction ratio estimate.

Suppose that the log reductions of the surrogate and the pathogen populations can be described with a predictive inactivation model, where the difference in inactivation rates between the two microorganisms is described by *R*. Using the loglinear inactivation model form as an example, the predictive models would look like those in Eqn. 7.5.

$$X_P = \int \frac{dt}{D_P} + \varepsilon_P$$

$$X_S = \int \frac{dt}{D_S} + \varepsilon_S \qquad \text{Eqn. 7.5}$$

$$\overline{X_P} = R\overline{X_S} = R\int \frac{dt}{D_S}$$

Assuming identical secondary models that describe the impact of intrinsic and extrinsic factors on the *D*-values for the surrogate (D_S) and pathogen (D_P), then the reduction ratio *R* can be used to accurately estimate the mean pathogen population for any level of surrogate reduction. Use of predictive models to quantify the appropriateness of a surrogate is a common practice in surrogate-pathogen comparison studies *(4, 5, 44, 64, 144, 262)*; where the reduction rate ratio can be estimated from a ratio of *D*-values at any system state (D_S / D_P).

The reduction ratio can be used to translate average surrogate reductions to average pathogen reductions if the reduction ratio is representative for the application. Ahmad et al. (2022) evaluated *E. faecium* NRRL-B2354 as a surrogate for *Salmonella* in a variety of foods (date paste, almond meal, wheat flour, peanut butter, nonfat dry milk powder, black pepper) and found the surrogate-pathogen relationship was different for each product *(4)*. More importantly, that study illustrated that the secondary model that described the impact of temperature on microbial thermal resistance was not uniform. For most products, the surrogate was more sensitive to changes in temperature, resulting in decreasing reduction ratios with increasing temperature. For any acceptable surrogate-pathogen-product combination with dissimilar secondary models (rate sensitivity to the modes of inactivation) there will be an applicability domain that results in acceptable reduction rate ratios from which a reduction ratio (Eqn. 7.4) can be estimated.

Any study that reports that a surrogate is appropriate for a pathogen-product-technology combination explicitly or implicitly defines an applicability domain. Studies that estimate robust predictive models may tolerate a reasonable degree of interpolation and extrapolation where the reduction ratio can be estimated for a preventive control validation. Studies without predictive models that focus on a single factor *(178, 256)* or a confluence of factors *(121, 177)* are more limited in applicability domains; however, if there is sufficient evidence that the experimental parameters are similar to the preventive control, then these studies may result in direct estimations of reduction ratios. Because the accuracy of the reduction ratio depends on the applicability domain of the supporting surrogate-pathogen research, by extension, the applicability of the statistical framework utilizing the tolerance bounds depends on the overlap between supporting evidence and preventive control treatment conditions. Any process treatment

conditions that would result in the relative pathogen/surrogate relationship substantially diverging during treatment would potentially invalidate use of the surrogate for predicting the expected level of pathogen inactivation.

In summary, applicability domain criteria needed to determine whether a surrogate is appropriate for a specific pathogen-product-process combination are the same criteria that would enable the use of a reduction ratio to translate mean surrogate reductions to predicted mean pathogen reductions. Therefore, prior supporting literature/evidence that resulted in robust reduction ratios estimates are already contributing to an essential component of the statistical framework developed thus far.

7.4 Framework Foundation Part Two: Translation of Outcome Spread

The second component needed to complete the foundation of the statistical framework is to determine whether surrogate reduction variability can predict the pathogen reduction variability. Estimation of pathogen variability becomes more complicated when including the scale of processing experienced during commercial-scale food safety preventive controls. Unlike the reduction ratio, quantifying the impact of reduction technologies on the relative variability between surrogate and pathogen outcomes is often not an objective of supporting research. Some research may incorporate reduction error in industry-relevant terms *(3, 177, 250)*, but no known research reports methodologies for predicting a microorganism's variability from a surrogate microorganism.

Following the example from the previous section, considering an ideal case where the inactivation models are known for both the surrogate and pathogen, it would be possible to translate surrogate reduction variability to pathogen reduction variability. Assuming that both

populations are in the same homogeneity region (space of treatment with minimal differences in treatment history), the translation of variability would become trivial, as represented in Eqn. 7.6.

$$X_P = \int \frac{1}{D_P} dt + \varepsilon_P$$

$$X_S = \int \frac{1}{D_S} dt + \varepsilon_S$$

$$\sigma = \sqrt{\frac{1}{n}\sum \varepsilon^2} \qquad \text{Eqn. 7.6}$$

$$X_S|_t = N(\overline{X_S}, \sigma_S)|_t$$

$$X_P|_t = N(\overline{X_P}, \sigma_P)|_t = N(R\overline{X_S}, \psi\sigma_S)|_t$$

In this scenario, the variability of the model is represented by a static value, meaning that at any given time (t), the spread of the reduction outcomes can be translated using a variability ratio (ψ; Eqn. 7.7).

$$\sigma_P^2 = \psi^2 \sigma_S^2$$

$$\sigma_P = \psi\sigma_S \qquad \text{Eqn. 7.7}$$

Here the preventive control reductions (X) and variability (σ) are specific to the validation scenario within the homogeneity region; therefore, the variability ratio ψ would also be specific to this scenario.

For the concept of a variability ratio to be valid, there would need to be several assumptions verified and practical concerns resolved. Surrogate-based validations are utilized because pathogen-based validations and direct estimation of pathogen reduction distributions are impractical, if not impossible, in most commercial processing systems. As a result, surrogate variability can be quantified, but pathogen variability, and therefore the variability ratio, cannot. This means that surrogate-pathogen comparison studies will need to be utilized to inform the

variability ratio. Conceptually, there are two likely formats that this could follow, one based on additive errors (Eqn. 7.8), or another based on predictive ratios (Eqn. 7.9).

$$\sigma_{P,int} = \psi\sigma_{S,int}$$

$$\sigma_{validation} = \sqrt{(\sigma_{int})^2 + (\sigma_{scale})^2} \qquad \text{Eqn. 7.8}$$

$$\sigma_P = \sqrt{\left(\psi\sigma_{S,int}\right)^2 + (\sigma_{scale})^2}$$

Here the variability reported in the surrogate-pathogen comparison study (σ_{int}) is representative of an intermediate estimate of variability indicative of population reductions, given the specific microorganism-product-technology combination. Therefore, the variability of outcomes in the preventive control validation would need to be adjusted with the variability imparted by the scale of the preventive control. This bottom-up approach to accounting for changes in surrogate/pathogen variability may conceptually work, but several issues render this approach invalid. While variances are generally recognized as additive, this would imply that any intermediate estimate of variability must be less than the variability experienced during industry-scale processing. Therefore, this interpretation of intermediate variability would require minimizing uncertainty of the estimate, which is another obstacle to successful surrogate utilization.

The translation concept proposed in Eqn. 7.9 follows more of a top-down approach to variability.

$$\sigma_{int}^2 = \sum \sigma_i^2$$

$$\sigma_{validation}^2 = \sum \sigma_i^2$$

$$\sigma_{P,int} = \psi_{int}\sigma_{S,int} \qquad \text{Eqn. 7.9}$$

$$\sigma_P = \psi\sigma_S$$

$$\psi_{int} \cong \psi$$

Here the sources of error within a homogeneity region that contribute to any estimate of error are the same (biological, sampling, handling, enumeration, reduction treatment, human, etc.), random, and normally distributed; however, the contribution of each component *i* to variability are unknown. The theory here is that the relative impact of the sources of error on the surrogate and pathogen may be characteristic of the microorganism-product-technology combination. Therefore, the ratio of the intermediate variabilities may be indicative of the validation variability ratio. The advantage of this approach is that it functions independently of "scale-dependent" variability, allowing for surrogate-pathogen comparison studies to have more variable outcomes than validation studies. Because variability ratios would require the estimation of the responses of two different microorganisms, the estimation of variability ratios are sensitive to sources of bias or uncertainty that may impact the microorganisms differently.

Previous discussions of reduction and variability ratios focused on homogeneity regions, which eliminated the necessity to consider spatial and replication errors; however, these will be considered for their impact on variability. These errors can change the distribution of microorganism responses and, if these sources of error are not sufficiently reduced or accounted for during the validation study, the statistical framework may be invalidated. To demonstrate the potential issue that spatial errors may have on accounting for variability, we will return to the theoretical kinetic inactivation descriptions. Previous discussion focused on homogenous

populations receiving a homogenous treatment; now we will examine the impacts of heterogeneous treatments between samples (n=1:*i*). For each individual sample X_i, the inactivation profile may be described by Eqn. 7.10.

$$X_i = \int \frac{dt}{D} + \varepsilon$$
$$D = f(intrinsic\ and\ extrinsic\ factors) \qquad \text{Eqn. 7.10}$$

The impact of the treatment, product, biological state, and other environmental conditions (i.e.., intrinsic and extrinsic factors) for both microorganisms will be described with the same secondary model form (*f*) but with similar, but not equal, parameters. As a result, the reduction ratio is not constant over the treatment, but the instantaneous resistance ratio (R') can be described with Eqn. 7.11.

$$R' = \frac{D_P}{D_S} \qquad \text{Eqn. 7.11}$$

Instead of considering the impact of heterogenous treatments on the randomness and normality of a single microorganism, we will consider the relative impact of spatial error between the two. Because different surrogate samples would receive different treatments, there would be a non-random spatial error resulting between two samples (Eqn. 7.12).

$$\varepsilon_{S,spatial,i-j} = \int \frac{dt}{D_S(f(i))} - \int \frac{dt}{D_S(f(j))} \qquad \text{Eqn. 7.12}$$

The issue of heterogeneity is further complicated considering that potential pathogen populations in identical locations may result in different magnitudes of spatial error. The difference in spatial errors is theoretically quantifiable (Eqn. 7.13); however, this would require knowledge of the inactivation history, because the instantaneous reduction ratio would be a function of time and impacted by the spatial differences.

$$\varepsilon_{P,spatial,i-j} = \int \frac{dt}{R'\big(f(i)\big) * D_S(f(i))} - \int \frac{dt}{R'\big(f(j)\big) * D_S(f(j))} \qquad \text{Eqn. 7.13}$$

This example focused on spatial error, but replication errors would result in the same issue. Not only would the heterogeneity of treatment potentially invalidate normality assumptions, but the incongruent impact of these errors would constitute a systematic bias that cannot be accounted for with reduction ratio (Eqn. 7.4) and variability ratio (Eqn. 7.5) translations. This has problematic implications for statistical frameworks for translating surrogate information to pathogen information for determination of process safety. Therefore, it is unlikely that a statistical framework based on large/diverse products or poorly reproducible processes can support accurate translations of surrogate information to pathogen.

There are several exceptions for surrogate-based validations that would bypass the issue of large/diverse products/processes and allow for a successful validation scenario as follows.

1. The process can be excessively effective at inactivating the surrogate, thus eliminating the need for accurate estimates of pathogen reductions. This would not necessarily allow for statistical evidence to support validation success based on the likely pathogen outcomes; however, if the surrogate is sufficiently conservative, then its statistically-evident reduction exceeding the pathogen's target lethality would logically infer that the pathogen reductions would be sufficient. Without a statistical framework for translating surrogate outcomes to likely pathogen outcomes, the necessary level of conservatism would be unknown.
2. A surrogate with similar inactivation kinetics (R=1 for the duration of the treatment), would help limit the unknown impact of spatial and temporal errors on surrogate and pathogen outcome distributions. With identical resistance characteristics, then the impact of these biases would be identical for both microorganisms. This would

potentially make use of the tolerance bounds imperfect estimates of the spread of potential reduction outcomes.

3. The product/process of concern can be small, uniform, and highly controlled (effectively maintaining a shared homogeneity region), which would then drive the expected levels of spatial and reproductive errors to zero. Without significant levels of these errors, then the inherent errors maintain their translative relationship. This condition, in combination with a supported reduction ratio, would allow for the translation between population reductions.
4. The most likely solution adapted by processors will be to identify the region of the product/process that yields the least surrogate and pathogen lethality, which will be designated as the worst-case (least-lethal) scenario for product safety. If a validation is based on a worst-case scenario, then it may be possible to render spatial and reproductive error negligible if the processor provides sufficient evidence that the sampling region is truly the worst-case scenario for reduction outcomes.

Worst-case scenarios (WCSs) are not a novel concept for preventive control validations; various resources and guidelines have suggested, or even required, validations based on WCSs *(2, 8, 11, 43, 98, 115)*. Previous iterations of validations based on WCSs meant achieving conservative estimates of reductions by purposely under-processing or by sampling only the samples most likely to accumulate the least inactivation (i.e., "cold spot" or "least lethal spot"). For applying this statistical framework for surrogates in preventive control validations, WCSs should consider the impact of process modifications on the average reduction and the variability of outcomes. A WCS modification should not necessarily be considered if it resulted in higher processing variability, even though higher variability would result in more conservative portions of lethality

outcomes. Instead, processors should identify whether the conditions resulting in the lower levels of reduction are reasonably expected processing parameters, and potentially focus on replicating those specific outcomes. A WCS should not be defined on how the process operates, but rather on achieving a representative worst-case processing outcome considering reasonable common-to-uncommon processing parameters.

7.5 Statistical Framework for Translating Surrogate to Pathogen Information

The advantage that this proposed statistical framework gives processors is the ability to establish preventive control performance criteria based on pathogen reductions instead of surrogate outcomes. With a solid foundation for translating the observable distribution of surrogate reductions to the likely distribution of pathogen reductions, there are several metrics to evaluate the efficacy of the preventive control, which can be chosen as part of determining the acceptable level of risk. We are proposing that using prediction intervals would be an appropriate statistical framework; therefore, the statistical tools discussed later in this section will be based on using tolerance bounds for determining preventive control performance criteria, power of validation design, and analyses of surrogate data. Consideration of common challenge study and validation practices will be used to inform the statistical framework and discussion.

With the objective of minimizing spatial and replication error, processing authorities would need to examine a singular condition for a validation. This condition, which would likely be considered the WCS for process operation, would become the minimum standard operation limit upon successful preventive control validation. The resultant surrogate mean and spread of reductions can be estimated using Eqn. 7.14

$$\overline{X_S} = \frac{1}{n}\sum_{i=1}^{n}(\overline{S_0} - S_i)$$

$$\sigma_S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_{S,i} - \overline{X_S}\right)^2} \qquad \text{Eqn. 7.14}$$

where the *n* posttreatment samples are normalized by subtracting the log surrogate survivors S_i from the average log initial population $\overline{S_0}$. These reductions are then used to calculate the mean ($\overline{X_S}$) and standard deviation (σ_S) of the surrogate reductions. Then the estimated mean ($\widehat{X_P}$) and standard deviation ($\widehat{\sigma_P}$) for pathogen reductions can be estimated (Eqn. 7.15).

$$\widehat{X_P} = R\overline{X_S}$$

$$\widehat{\sigma_P} = \psi\sigma_S \qquad \text{Eqn. 7.15}$$

Applying Eqn. 7.1 and 7.2 for the observable surrogate population, the tolerance bounds can be estimated for the surrogate population for α percent confidence that q percent of the surrogate reductions exceed the 100(1-α) percentile $X_{S,q}$. If processors choose to use this as the performance criterion, which would be useful if the surrogate is known to be conservative but the reduction/variability relationship is otherwise poorly known, then $X_{S,q} > TL$ would indicate that process exceeds the performance criterion. The same equation can be adapted to the predicted pathogen distribution using the translational equations (Eqn. 7.16), and the performance criterion can be likewise evaluated.

$$X_{P,q} = R\overline{X_S} - k_1\psi\sigma_S \qquad \text{Eqn. 7.16}$$

Direct translation of reduction/variability ratios and tolerance bound uncertainty (captured by k_1) would be appropriate if there is high confidence in the accuracy of the translation equations. However, several steps can be taken by processors to overcome uncertainty in the reduction and variability ratios. Confidence intervals could be determined for the ratio estimates, and

processors could use the combination that results in more conservative $X_{P,q}$ estimates (smaller R and larger ψ estimates). Alternatively or in conjunction, processors may replace standard score values z_q and z_α with Student's t-distribution values $t_{q,v}$ and $t_{\alpha,v}$, where v is the associated degrees of freedom for the surrogate estimate.

These equations would allow for a pathogen performance criterion of q percentage of the pathogen population ($X_{P,\alpha}$) achieving the target lethality for the pathogen (TL_P) to be evaluated. The same pathogen-based performance criteria ($X_{P,\alpha} \geq TL_P$) can be translated to an equivalent surrogate target lethality (TL_S) and performance criterion ($X_{S,\alpha} \geq TL_S$), which may allow for easier communication of performance criterion to process operators (Eqn. 7.17).

$$TL_S = R^{-1}(TL_P + k_1\psi\sigma_S) - k_1\sigma_S \qquad \text{Eqn. 7.17}$$

This may be especially helpful in the experimental design of challenge or validation studies if the expected levels of reduction and variability are sufficiently well known in advance.

7.6 Defining Surrogate Utility from Statistical Framework

A limitation of previous surrogate-based validation analysis methods is that there were no statistical rationales that could be used to grade a surrogate's appropriateness using the reduction ratio (R) or variability ratio (ψ). The previous chapter identified some perspectives on desirable or the limitations in reduction ratios that were available in recent literature, these limitations were based on logical concerns *(3, 4, 110, 273)*. The same literature did not report any desirable characteristics for variability ratios. With the proposed statistical framework, it is possible to determine the boundaries that define optimal surrogate-pathogen relationships, provided some practical considerations were established. While utilizing this proposed statistical framework enables for the possibility of dissimilar target reductions for the surrogate and pathogen, there is still the limitation that the surrogate reduction distribution must be accurately quantified.

Additionally, it must be considered that the closer the reduction and variability ratios are to a value of 1, the less that potential error from treatment heterogeneity may impact the ability to translate from surrogate to pathogen. Therefore, by considering surrogate reduction floors/ceilings and closeness of resistance/variability, different grades of surrogate utility were created.

Surrogate utility classifications are at least partially dependent on the performance criterion for the preventive control. For example, a preventive control with a performance criterion of 95% of product achieving 10 log reductions of the pathogen would not be supportable with a classic ideal surrogate with 1:1 translation. This is because validation studies typically cannot experimentally quantify 10 log reductions of a microorganism because of the limit of detection/quantification. In this case, a surrogate with a higher reduction ratio (e.g., R=2) would be more appropriate for this process. In contrast, if a processor was employing a hurdle approach to achieve the food safety objective and only needed to validate a pathogen performance criterion of 2 log reductions, then a less resistant surrogate (e.g., R<1) may be statistically valid.

The following sections were written as an illustrative case of a preventive control with a 4.5 log reduction performance criterion for the pathogen. A visual representation of the impact of reduction and variability ratios on surrogate utility classification can be seen in Figure 7.1. This figure is not an exact classification chart, as the ranges of classification will vary depending on the performance criterion, number of samples in the validation test set, and whether any conservative adjustments were applied to the predictive pathogen tolerance bounds. The performance criterion used was that 95% of the population must exceed the target lethality with

90% confidence. The following classifications include the mathematical expressions that were used to define the domains within Figure 7.1.

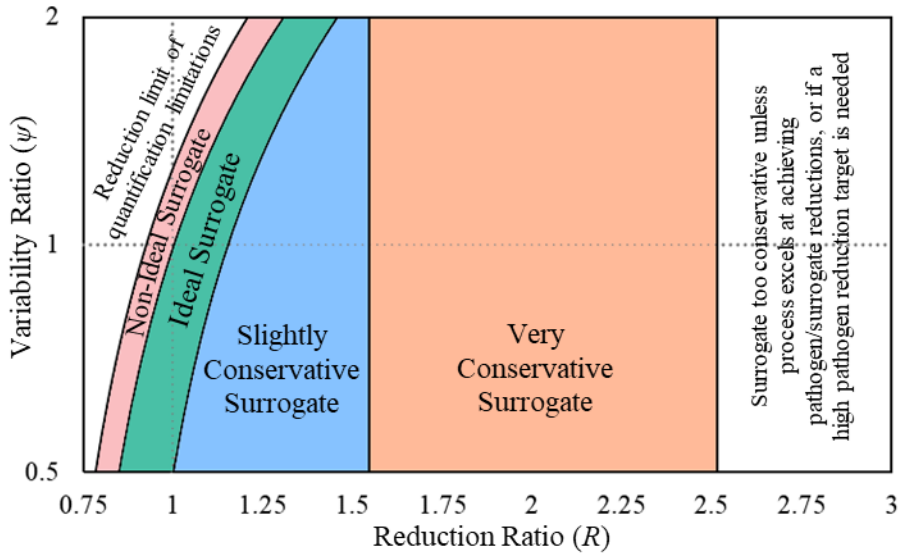


Figure 7.1. Surrogate utility classification chart based on reduction and variability ratios

7.6.1 Ideal Surrogate for Translation

For an "ideal" surrogate, surrogate resistance and variability characteristics are similar or slightly conservative compared to the pathogen's, allowing for a simple 1:1 translation of reduction information without the risk of excessive overprocessing (≤ 1 log reduction beyond the target lethality unaccounted for in translation; Eqn. 7.18). For processes on the margins of acceptance for the performance criterion, the translational equations may be used to assess the likely pathogen reduction distribution more accurately.

$$\left(TL_P \leq X_{P,q} \leq TL_P + 1 \middle| X_{S,q} \cong TL_P\right) \qquad \text{Eqn. 7.18}$$

7.6.2 Slightly Conservative Surrogate

Validations using this a "slightly conservative" surrogate would likely result in pathogen reductions substantially exceeding the pathogen performance criterion when the surrogate reductions meet the pathogen performance criterion ($X_{S,q} \cong TL_P$; Eqn. 7.19). Validations marginally exceeding the performance criterion based on pathogen outcomes would require use of the translational equations. Without greatly exceeding the performance criterion ($\overline{X_P}$ >10 log reductions), high performance validations with minor conservative surrogates may still yield surrogate reduction outcomes exceeding the pathogen performance criterion. Such an outcome would enable processors to form a logical conclusion of the process's efficacy based on the surrogate reductions satisfying the pathogen's performance criterion.

$$\left(TL_P + 1 + k_1\sigma_P \leq \overline{X_P} < 10 \middle| X_{S,q} \cong TL_P\right) \qquad \text{Eqn. 7.19}$$

7.6.3 Very Conservative Surrogate

A "very conservative" surrogate would be more conservative in nature than the slightly conservative surrogate, therefore excluding any likelihood that the surrogate reduction distribution would meet the pathogen's performance criterion ($X_{S,q} < TL_P$; Eqn. 7.20). Theoretically, there are no hard limitations to how conservative a reduction ratio can be, assuming that any level of pathogen reduction is achievable. The boundaries defining a very conservative surrogate must therefore be established based on what is reasonable to expect for a required pathogen reduction ceiling and the minimum requirements for needed to support a reasonable and quantifiable surrogate reduction floor. For this type of surrogate, it was assumed that the process was capable of exceptional levels of pathogen reduction ($\overline{X_P}$ =10 log reductions) and anything less than a 2-log reduction would not be acceptable for the surrogate ($X_{S,q} \geq 2$ log reductions).

$$\left(TL_P + 1 + k_1\sigma_P \leq \overline{X_P} \leq 10 \middle| 2 \leq X_{S,q} \leq TL_P\right) \quad \text{Eqn. 7.20}$$

7.6.4 Non-Conservative Surrogate

By strictly using the translation equations, it is possible to utilize a "non-conservative" surrogate that is slightly more susceptible to the mode(s) of reduction than that of the pathogen (Eqn. 7.21). The utility range of this type of surrogate is restricted by the fact that both mean and standard deviation of the reduction outcomes distribution need to be quantifiable, which is limited by the difference between the highest inoculation level attainable and the limit of quantification – both factors are highly dependent on limitations to the validation study design. A preventive control with a high pathogen performance criterion (e.g., 6 log reductions) would be effectively unable to utilize a non-conservative surrogate. For this type of surrogate, it was assumed that the highest quantifiable reduction of the surrogate was 7 logs.

$$\left(TL_P \leq X_{P,q} \middle| \overline{X_P} \leq \overline{X_S} \leq 7\right) \quad \text{Eqn. 7.21}$$

With these surrogate utility classifications, the concept that an "ideal surrogate" and pathogen of concern have equivalent resistance to the mode of inactivation (R=1) is a misconception (Figure 7.1). A surrogate with a R=1 would be classified as a non-conservative surrogate if the pathogen was inherently more variable for that product and process (ψ>1). A slightly more resistant surrogate (R≈1.1) would be more tolerant of differences in variability ($0.85 \leq \psi \leq 1.30$) and still be considered an ideal surrogate for translation, given the proposed framework. Prior to this framework, there has been no consensus on how much more resistant the surrogate could be while still being appropriate for use, which was discussed in the previous chapter. A few recent surrogate-pathogen comparison studies provided concern about an upper limit to a surrogate's conservatism or reduction ratio *(3, 4, 105, 273),* resulting in a range of reduction ratio of approximately 2-3 as being the upper limit for usefulness of a surrogate. While

the upper limit depends on several factors (e.g., pathogen performance criterion, inherent surrogate/pathogen variability), the upper limit for a slightly or very conservative surrogate given the parameters established for Figure 7.1 would indicate that the upper limit for R is closer to 1.5 and 2.5, respectively. The limitations that defined these upper limits for the proposed framework were also based on logical concerns, similar to other literature concerned about overly conservative surrogates *(3, 4, 105, 273)*. Unlike the previous studies, this statistical framework would enable statistical estimation of pathogen reductions up to $R \approx 2.5$. If processors were considering using a 1:1 translation, then they may be limited to slightly conservative surrogates, which caps at a $R \approx 1.5$ or achieving preventive control performance that exceeds 10 log reductions of the pathogen of concern.

7.7 Statistical Power of Prediction Interval Performance Criteria

The advantage of using tolerance bounds is that the confidence (type I error) of the test can be determined by the operator. This type I error rate is directly appliable to the observable surrogate population; type I and II error for the pathogen is unknown, as it is not quantifiable in typical surrogate-based validation tests. As was demonstrated with theoretical kinetic relationships between surrogates and pathogens, the ability to translate surrogate information to pathogen information depends on the accuracy of the translation equations. This would also extend to the accuracy of type I and II errors, as a pathogen identical to the surrogate would have identical error rates. Type II error, or the false negative rate β, and its counterpart statistical power ($1-\beta$), is not typically a selectable error rate because it implies knowledge that the statistical test should reject the null hypothesis, which is in this case that the target lethality is included in the q % of the population estimated. The alternative hypothesis is that the target lethality is not included in the q % of the population estimated. Prior knowledge of an

experimental design's statistical power would enable processors to optimize data collection practices to achieve an acceptable false negative rate. For example, a processor may reconsider data collection practices or improving the preventive control's performance if it was known that the likely statistical power of their original validation design was 50%, meaning that their process is just as likely to correctly be determined sufficient as it would be to incorrectly be determined insufficient.

Type II error for the tolerance interval can be empirically estimated by determining whether a population of outcomes with characteristics of the alternative hypothesis (i.e., should exceed the performance criterion) can successfully reject the null hypothesis. By utilizing bootstrapping methods on simulated validation experiments, where a pre-determined reduction distribution with a known percentage exceeding the target lethality (θ) can be tested against the performance criterion (q=95%), the false negative rate can be approximated for cases $\theta > q$. For statistical tests like tolerance bounds, bootstrapping methods to estimate the validation acceptance rate (AR, AKA as discovery rate) can be used to estimate type I error for cases $\theta < q$ (type I error=AR) and type II error for cases $\theta > q$ (type II error= 1-AR). For non-statistical tests used in surrogate-based validations, such as MRC (Eqn. 7.22),

$$MRC = \min(S_0) - \max(S_i) \qquad \text{Eqn. 7.22}$$

the AR does not correspond to type I or II error; however, ARs for tolerance bounds and MRC can be estimated for any θ to provide an "apples-to-apples" comparison of test robustness.

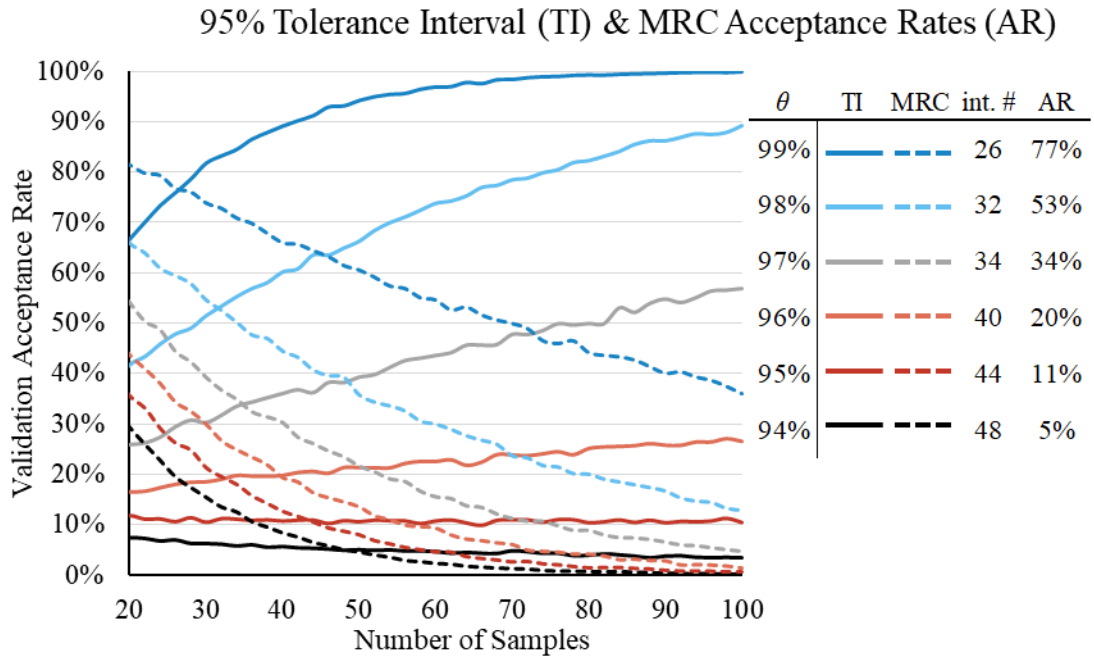


Figure 7.2. Acceptance rates (ARs) for tolerance interval- and MRC-based performance criterion as impacted by percentage of a population exceeding the target lethality (θ) and number of samples in a validation. AR crossover points are detailed using the intersection sample number (int. #) and the corresponding AR for each θ.

Therefore, ARs were bootstrapped using a Monte Carlo simulation (n=10,000), where normally reduction distributions with θ= 94-99% ≥ TL were randomly generated and evaluated with tolerance bound (q=95%, α=90%) and MRC based performance criteria to estimate AR (Figure 7.2). Datasets were generated with 10 initial population samples (N(0,0.1) log reductions) and 16-100 treated samples with a standard deviation of 0.51 log reductions. The mean lethality of treated samples was adjusted to achieved different θ levels. Simulated validation datasets with θ=0.94 represent a slightly underperforming preventive control when using the tolerance bound-based performance criterion, but were included to demonstrate the sensitivity of the tolerance bound test.

Both MRC and tolerance bound performance criterion are based on demonstrating that a high frequency of samples exceed the target lethality; however, because the tolerance bound test reduces uncertainty with additional samples, the accuracy of this assessment is improved with additional samples. The MRC performance criterion does not prescribe desirable reduction distribution characteristics; therefore, it cannot be said that sampling improves or worsens accuracy. It can be stated that any process with MRC-based performance criteria is more likely to reject the process with additional sampling.

Overall, the acceptance rate for a tolerance bound test and MRC-based performance criterion can be generally summarized with Eqns. 7.23 and 7.24, respectively, for reduction distributions with θ% population exceeding the target lethality.

$$Tolerance\ Bound\ Acceptance\ Rate = \begin{cases} \theta \leq q, & \leq 1-\alpha \\ \theta > q, & 1-\beta(n, \theta - q) \end{cases} \quad \text{Eqn. 7.23}$$

$$MRC\ Acceptance\ Rate = [\Pr((\min(S_0) - \max(S_i))) > TL]^n \sim (\theta)^n \quad \text{Eqn. 7.24}$$

The common trend influencing the acceptance rate for both tolerance bound- and MRC-based performance criterion is that the better performing the process (i.e., $\theta \rightarrow 100\%$) the higher likelihood of accepting the preventive control as effective. For each θ examined, there was an intersection of AR that divided regions that resulted in high AR for MRC-based analyses for fewer samples and higher AR for tolerance bound-based analyses with more samples (Figure 7.2). However, for $\theta > q$=95% and validation experiments with samples > 40, the MRC-based performance criterion resulted in lower AR than the tolerance bound-based performance criterion. This results in a dichotomy that favors processors with the objective of exceeding the TL for 95% of the surrogate population, because both tests can be utilized knowing the relative reliability of the test. For surrogate-based evaluations of the performance criterion, a validation consisting of 40 or more samples that was accepted using the MRC-based analysis would have a

high likelihood of describing a reduction distribution θ>95%; however, for a MRC-based failed validation, a processor may rely on tolerance-bound metrics to evaluate preventive control performance more accurately. For performance criterion based on likely pathogen outcomes, processors would need to rely on the statistical framework proposed in this study.

7.8 Conclusions

Current strategies for evaluating surrogate-based validations are biased for conservative outcomes, which do not encumber processes with "ideal" surrogates or those that sufficiently exceed the performance criterion. However, the inability of current surrogate-based validation evaluation strategies to estimate the probability of pathogen outcomes creates problems for marginally successful preventive controls that would have difficulty proving the pathogen-based performance criterion for a conservative surrogate. The statistical framework proposed in this study would improve the utility of all surrogate-based validations by improving the information provided by validation studies. This approach would then improve process authorities' ability to correctly identify sufficient operating parameters for optimal food production and risk reduction.

As a novel method of analysis for preventive control validations, there are some limitations to the applicability of this statistical framework. There are strict prerequisites for understanding the reduction and variability ratios defining the surrogate-pathogen relationship and the reproducibility of the process. Processors would need to verify that surrogate reductions are lognormally distributed with limited effects of spatial and replication error. Prior analysis strategies, such as the MRC method suggested by Ceylan et al. (2021), require that the surrogate be more resistant to the modes of inactivation and do not require lognormal distribution of reductions. However, such analysis methods are not statistics-based, which creates several issues. Empirical test performance criteria acceptance depends on the capacity to prove that the

surrogate exceeds the pathogen-based target lethality, and because such tests are not statistics-based, the rejection rate can be artificially inflated with additional sampling. Additionally, because empirical methods are only concerned with the determination that a surrogate is acceptable, the relative resistances and variabilities of the surrogate and pathogen are effectively ignored. With the translational equations included in the proposed statistical framework, processes may utilize both observed surrogate and estimated pathogen reduction distributions in their evaluation of process success. While the process of reducing spatial and replication errors may seem daunting, such practices are already commonplace when processors focus on WCSs. Finally, because this statistical framework establishes actual performance-based outcomes criteria (e.g., $\geq$ 95% of processed product exceeding the target lethality for the population) a failed validation can inform processors on the additional margins of safety needed. Validations based on simple pass/fail determination do not necessarily inform the processor as to the additional margins of safety needed to move a failed preventive control to a successful one.

This statistical framework has the potential to substantially improve the design and analysis of surrogate-based validations. However, with wider adoption and acceptance, this framework may have additional potential positives. By clearly identifying the types of information needed to communicate surrogate utility, the proposed framework may improve the translation of information between researchers, process authorities, risk assessors, and regulatory agencies. As with any adoption of a standard of practice, there may be reduced costs for surrogate-based validations because less resources would be needed to "reinvent the wheel" or determine how to develop validation studies and communicate the subsequent results.

CHAPTER 8: A CASE STUDY IN APPLYING A STATISTICAL TOLERANCE FRAMEWORK FOR EVALUATING SURROGATE-BASED ALMOND DRY ROASTING VALIDATIONS BASED ON SURROGATE- AND PATHOGEN-BASED CRITERIA

This chapter represents a component of ongoing work with the Almond Board of California (ABC) to improve surrogate-based process validation guidelines for almond processes. The ABC is responsible for almond process validations within California *(80)*, and the results included here were approved by the ABC for release, consistent with an active confidentiality agreement between ABC and Michigan State University (MSU).

The ABC has published validation guidelines for multiple technologies (e.g., oil-roasting, blanching, dry-roasting, moist-heat treatments) and maintains a list of approved process authorities to assist with process validations *(10-12)*. Additionally, there is a guideline specific to executing and analyzing a surrogate-based process validation *(13)*. Within these guidelines, the ABC provides the performance criterion and method for analyzing surrogate data. With a goal to "achieve the mandated 4-log reduction of *Salmonella* in California-grown almonds," the actual performance criterion is a 4-log reduction of the surrogate, *Enterococcus faecium* NRRL-B2354. Furthermore, ABC surrogate-based validation guidelines require a Minimum Reduction Case metric for performance criterion determination *(13, 43)*. Each process validation report is submitted to the ABC through their pasteurization program *(14)*, in which a Technical Expert Review Panel (TERP) reviews the validation report and, potentially through an iterative process, accepts the validation report on behalf of the ABC.

As discussed in the previous chapter, this approach, which incorporates surrogate-based performance criteria and a non-statistical basis for data analysis, is potentially detrimental to preventive controls that fail to achieve substantially more pathogen/surrogate reductions than

targeted for the performance criterion. Therefore, the objective of this case study was to utilize the tolerance interval statistical framework to update surrogate-based validation results for drying roasting of almonds to quantify likely pathogen-based outcomes.

8.1 Methods and Materials

The ABC maintains a database of all almond preventive control validations, active and inactive. In collaboration with the ABC, access to a secure server where active validation reports were maintained/catalogued was provided to Dr. Bradley Marks and Ian Hildebrandt from MSU in late 2020. The secure server maintained active records as of the date access was granted to MSU partners. Validation reports were assigned blind identification codes to prevent discussing identifiable information. Additionally, no company-identifiable information was collected during the review of validation reports. Twenty-six dry roasting validation reports were reviewed and any/all information on surrogate-based challenge testing (microbial results, process conditions, etc.) was extracted and catalogued using the identification code. Microbial data were converted to log reductions and stored by trial/replication. Microbial data that were censored (i.e., below the limit of detection) were annotated and not replaced with substitute values (e.g., the limit of quantification). Censored microbial data (e.g., survivors reported as <1.7 log CFU/g) disrupt typical statistical descriptions of reduction distribution, because the actual reductions received for this portion of the dataset is unknown (therefore considered censored), and replacement estimates may result in non-normal distributions and biased mean and standard deviation estimates. The following paragraphs discuss how uncensored and censored data were used.

Several checks were performed on each validation dataset to confirm the applicability of the statistical framework proposed in the previous chapter of this dissertation. Briefly, to apply this formulation of a tolerance interval, there is an expectation that the treatment is homogeneous

and the random errors are normally distributed. A homogeneous treatment was described having limited heterogeneity imposed on surrogate/pathogen reductions by spatial and replication errors, which may be achieved through targeting worst-case scenarios. The requirement for limited replication and spatial errors was assumed to be met because ABC dry roasting validation guidelines require that surrogate-based validations: (1) identify the coldest spot/path on the line, (2) identify the worst-case scenario parameters for each product, and (3) validate lethality for these worst case scenario parameters *(11)*. Additionally, uncensored validation replication datasets were tested to verify a potential normal distribution of reduction outcomes using the Anderson-Darling test (type I error = 5%). Because most experimental datasets included censored data, if the uncensored replications within the experiment were determined to be likely normally distributed, it was assumed to be true for the entire validation experiment. Uncensored replications were also used to identify the distribution of variability in reduction outcomes for use in potential Maximum Likelihood Estimation (MLE) of reduction distributions of highly censored datasets. Because validation datasets were already generated with worst-case scenarios, if the assumption of normally distributed reduction outcomes was supported, then tolerance intervals for the validation datasets were estimated using a 95% tolerance bound with 90% confidence (i.e., q=0.95 and α=0.90; Eqn. 8.1 and 8.2).

$$X_q = \bar{X} - k_1\sigma \qquad \text{Eqn. 8.1}$$

$$k_1 = \frac{z_q + \sqrt{z_q^2 - ab}}{a}$$

$$= 1 - \frac{z_\alpha^2}{2(n-1)} \qquad \text{Eqn. 8.2}$$

$$b = z_q^2 - \frac{z_\alpha^2}{n}$$

Briefly, the population of outcomes X (i.e., "log reductions") can be described with a mean of $\bar{X}$, a standard deviation of σ, and the q% lower tolerance bound X_q. The scale factor k_1 accounts for the uncertainty in the mean and spread of data and is a function of the number of samples (n) and the standard scores z_q and z_α for the q percentage of population described in the performance criterion and α confidence level, respectively. This estimation of the tolerance bound requires n > 10 samples and has n-1 degrees of freedom.

Tolerance intervals were estimated for surrogate validation experiments that satisfied the normality assumptions using a 95% tolerance bound with 90% confidence (i.e., q=0.95 and α=0.90). Validation experimental datasets that were minimally censored (≤10% censored) used the uncensored data to estimate the mean and standard deviation of reductions. Other validation datasets that were not near completely censored (≤90% censored) were characterized with MLE methods to estimate the likely mean and spread of the normally distributed censored datasets. If any validation dataset was estimated to have a variability larger than the largest observed variability from individual replications, the MLE was reperformed using a fixed variability of the 90th percentile of the observed variability from the uncensored replications.

Surrogate (*E. faecium* NRRL-B2354) tolerance intervals were translated to predicted *Salmonella enterica* PT 30 tolerance intervals using reduction (R) and variability (ψ) ratios extracted from a pilot-scale surrogate comparison study by Jeong et al. (2011). That study, which was also foundational for ABC dry-roasting surrogate guidance *(13)*, was considered a reasonable representation of the *E. faecium-Salmonella* PT 30 reduction relationship for this application, because it utilized identically specified materials (nonpareils 27/30 almonds), a lab-scale, moist-air impingement oven, and a range of conditions that encompass likely dry-roasting

conditions (121-204 °C, dry to 90% moisture by volume). The suitability of *E. faecium* NRRL-B2354 as a *Salmonella* PT 30 was subsequently reverified in a later publication *(119)*.

8.2 Results and Discussion

Of the 26 validation reports reviewed, 23 contained extractable surrogate reduction data. This yielded 176 treatment replications and 864 samples. While most replications contained censored data, 53 replications were uncensored and used for characterizing process variability (Figure 8.1). The 90th and 100th percentile of surrogate reduction variabilities were 0.81 and 1.03 log, respectively. From these uncensored replications, only 1 was determined to be likely not normally distributed. This was less than the type I error used for the Anderson-Darling test; therefore, it was assumed that all dry-roast reduction outcomes were likely normally distributed.

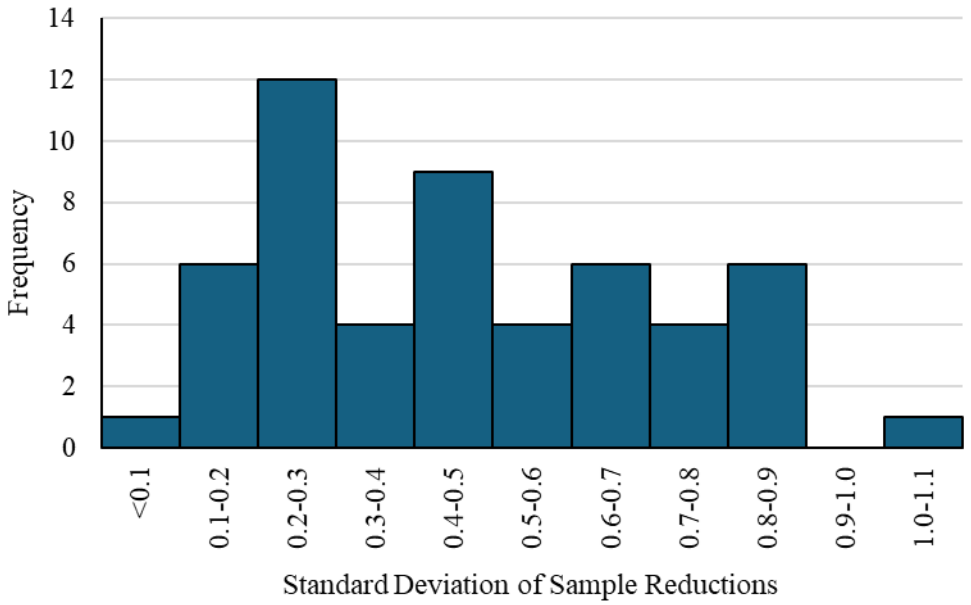


Figure 8.1 Distribution of variability among uncensored dry-roast validation replicates

Validation reports submitted by processors would need to demonstrate that all samples exceed a surrogate-based MRC calculation of 4-log before being accepted by the ABC. While each of the validation reports examined passed using ABC's MRC criterion, one validation

report resulted in a 95% tolerance bound that did not exceed a 4-log reduction of *E. faecium* (Table 8.1). Validation #3 also consisted of the fewest samples, meaning this dataset had the most uncertainty in estimating the mean and spread of data. Validation #3 surrogate tolerance bound failing to exceed the target 4-log reductions is not likely problematic, considering the projected *Salmonella* tolerance bound was 4.97 log reductions (Table 8.2), meaning that the process exceeded the target lethality.

Table 8.1 Summary of surrogate reductions and surrogate-based evaluations of the performance criterion

		Surrogate Dataset and Results				
Validation #	**Number Samples**	**% Censored**	**MRC***	**Mean Reduction**	**Std. Dev.****	**95% Tol. Bound**
1	24	0%	5.38	5.92	0.40	5.07
2	80	31%	5.19	7.72	0.81	6.19
3	9	0%	4.00	4.58	0.45	3.43
4	36	86%	6.23	7.97	0.83	6.30
5	36	25%	4.24	6.11	0.81	4.47
6	12	58%	4.30	6.18	0.81	4.24
7	58	29%	4.16	6.31	0.81	4.75
8	36	56%	4.61	6.70	0.81	5.07
9	12	0%	5.42	6.25	0.49	5.07
10	36	72%	5.67	7.85	0.83	6.17
11	44	34%	5.46	6.52	0.52	5.48
12	70	24%	5.49	6.47	0.59	5.34
13	16	0%	4.66	5.11	0.20	4.65
14	18	22%	4.82	6.24	0.81	4.44
15	68	18%	4.59	5.98	0.86	4.33
16	27	37%	5.03	6.83	0.80	5.16
17	80	29%	4.20	5.93	0.99	4.06
18	12	75%	6.48	6.97	0.31	6.24
19	48	4%	4.52	5.76	0.63	4.51
20	16	38%	5.20	6.49	0.81	4.66
21	44	34%	5.10	7.08	0.79	5.51
22	44	86%	5.54	8.01	0.81	6.40
23	56	82%	5.32	7.40	0.81	5.84

* MRC estimates are approximate values because data were extracted as average reductions, not survivor ratios

** MLE distrubtion estimates that exceeded 1.02 log std. dev. was re-estimated using 0.81 log std. dev.

Table 8.2 Predicted validation performance using projected *Salmonella* reductions

	Projected Pathogen Estimates		
Validation #	**Mean Reduction***	**Std. Dev.***	**95% Tol. Bound***
1	8.32	0.51	7.23
2	10.85	1.03	8.90
3	6.44	0.57	4.97
4	11.20	1.06	9.06
5	8.58	1.03	6.50
6	8.68	1.03	6.20
7	8.87	1.03	6.87
8	9.42	1.03	7.33
9	8.78	0.63	7.27
10	11.03	1.07	8.88
11	9.16	0.67	7.83
12	9.08	0.75	7.65
13	7.17	0.26	6.59
14	8.77	1.03	6.47
15	8.40	1.10	6.30
16	9.59	1.02	7.46
17	8.33	1.26	5.95
18	9.79	0.39	8.86
19	8.09	0.81	6.50
20	9.12	1.03	6.78
21	9.95	1.01	7.95
22	11.25	1.03	9.20
23	10.40	1.03	8.40

* Values were estimated using reduction and variability ratios applied to surrogate data

for the pathogen of concern. Overall, the high success rates of both surrogate- and pathogen-based evaluations of the performance criterion reinforce what the almond validations already demonstrated – that the almond dry roasting processes yielded sufficient pathogen control.

The analyzed collection of validation reports did not include processes that failed to achieve an MRC >4-log reduction for the surrogate during initial validation testing. Through evaluation of validation reports, it was evident that several validations were not reporting the first iteration of results, but had to modify the process to achieve the higher levels of reduction

ultimately submitted for consideration. MRC-based analyses do not have a reduction probability-based objective. In theory, the MRC-based analysis requires all samples to exceed the target lethality; therefore, each additional sample increases the probability that a single sample will cause a failure in the validation study. By incorporating the tolerance interval as the metric to evaluate the performance criterion, processors reduce uncertainty with each additional sample and improve the estimation of process performance. If an otherwise successful validation was rejected due to the randomness of MRC metrics, then using a tolerance interval may correct that false rejection.

The ABC and risk assessments for *Salmonella* on almonds generally agree that a 4-log reduction of *Salmonella* is a good target for preventive controls *(61, 77, 143)*. Because all dry-roasting almond validations analyzed in this study achieved this pathogen-based objective by assuming the same target for a surrogate-based performance criterion, most dry roasting validations easily exceed a 4-log reduction of *Salmonella* and may be achieving closer to 6-8 log reductions on average (Table 8.2). Unless extraordinary outbreak events were considered, a prior *Salmonella* risk assessment for almonds did not estimate any fewer illnesses-per-year for an industry-wide risk reduction exceeding 5 log reductions of *Salmonella (143)*, meaning that almonds receiving an 8 log reduction have the same practical food safety outcome (i.e., <1 case salmonellosis per year) as those achieving a 5 log reduction. Without an additional public health benefit, the potential additional operating costs for processors may be an undue burden. By utilizing the proposed surrogate-pathogen translation framework, processors could be better informed about the actual levels of risk reduction being achieved, and therefore make risk-informed decisions that impact the cost of conducting process validation studies, operating costs, and product quality.

8.3 Conclusions

This case study demonstrates the utility of a proposed framework for translating surrogate-based validation outcomes into predictions of pathogen outcomes. While it was known that these processes likely achieved the 4-log reduction of *Salmonella* PT 30 required by the ABC, this is the first time the actual levels of preventive control performance were statistically estimated. The tools demonstrated in this framework can help improve the utility of future surrogate-based validations and better enable processors to optimize processes while protecting both public health and the economic health of the processor.

CHAPTER 9: CONCLUSIONS

Given the overall goal of improving the utility of food safety research for improved application in predictive model- and surrogate-based preventive control validations for low-moisture foods, a successful outcome for this dissertation would represent a step in this direction.

As was demonstrated in Chapter 3, most predictive inactivation models are not appropriate for use in a predictive capacity. This is largely because research in the field typically does not differentiate between descriptive modeling and predictive modeling. More importantly, there is no adopted standard for what needs to be reported for a model to be utilized for predictive applications. Before there can be a discussion about how to best validate models for use in preventive control validations, basic practices like reporting every model parameter, parameter uncertainty, and regression error need to be standardized. Influenced by resources for improving model development and validation, a set of six principles were developed as the basis for Good Standard of Practice for Predictive Modeling (GSPPM), which should improve awareness of the key components needed to improve predictive inactivation modeling.

Chapters 4 and 5 addressed two immediate concerns for predictive modeling. One concern associated with predictive modeling was how reproducible and comparable results were across studies. A standard laboratory practice for estimating thermal resistance parameters was developed (Chapter 4) that yielded reproducibility of parameters within 10%. This set of standard laboratory practices was then made the default practice for several large research grants and was used across many research institutions. The other immediate concern was the dissemination of flawed, model-based process validation practices. The research in chapter 5 thoroughly demonstrated the “fail-dangerous” outcomes of a validation practice that utilized inactivation models that failed to appropriately account for the known impact of moisture on

Salmonella thermal resistance in low-moisture foods. This represents another step to discourage improper predictive microbiology practices, while presenting more robust practices.

Research on surrogates for application in preventive controls is at a better readiness-state than predictive models. Based on a detailed and quantitative analysis of the literature, most surrogate research has depended heavily on identifying the "ideal" surrogate for a "1:1 ratio of results" quantitatively evaluating what the surrogate-pathogen reduction relationship should be under ideal conditions. Even though there was clear evidence that the standards of analysis should be statistically sound, no evaluation metrics previously being recommended were based on rigorous statistical methods of quantifying the reliability of predicted pathogen outcomes. To advance the field of surrogate-based process validations, the structural requirements for an appropriate surrogate analysis were described (Chapter 6); a statistical framework that would improve the utility of surrogate *and* pathogen research was developed (Chapter 7); and the utility of the proposed analysis methods were demonstrated (Chapter 8) using real validation datasets from the almond industry. Implementation of statistical tolerance bounds to describe the likely distribution of all future outcomes would significantly improve surrogate-based validations. With enough improvement in the utilization of predictive models, tolerance bounds may have a promising application in predictive model-based validations.

CHAPTER 10: FUTURE WORK

Although this dissertation makes significant contributions to the field, the challenges related to validation of pathogen control processes are not resolved. Efforts to improve standards of practice take time and momentum.

The single best innovation that can drive improvement is the adoption of a robust standard of practice that sets the goal for the design and dissemination of research in this area. It is not a coincidence that *Enterococcus faecium* NRRL-B2354 is the most researched surrogate for *Salmonella* in low-moisture foods. The Almond Board of California took the initiative and generated a set of validation guidelines, one being the utilization of *E. faecium* as a validation surrogate, which became the standard for almonds, and likely many other low-moisture foods. The statistical framework for analyzing surrogate-based validation results has the potential for moving toward another industry-wide adoption; however, it is not a silver bullet. As was addressed in this dissertation, the current structure, and the translational equations for estimating pathogen-based outcomes, rely on many assumptions. The current formulation of the tolerance bound is the limit of that statistical test, formulations exist for other distributions or even non-parametric distributions. However, it is not known how applicable those formulations are to this field, considering the high likelihood of censored data. The immediate future for the surrogate-side of the work is to reach out to partners in research, the food industry, and in government to see how this statistical framework can be rigorously tested and perhaps ultimately adopted or endorsed by an appropriate scientific, industry, or regulatory body.

The future of improving predictive models for utilization in low-moisture food preventive controls is less obvious. Improving modeling and statistical literacies in our fields of research should improve the quality of research design, reporting, and application. However, it will take

more than improved practices to enable the use of predictive models in preventive control validations. One difficulty with low-moisture foods is that a "best" model form has not yet been convincingly developed, tested, and reported. To the best of my knowledge, no inactivation model has been proposed that seamlessly integrates into low-moisture processes and accurately predicts inactivation outcomes based on measurable, dynamic temperature and moisture profiles. With or without such a model, there always will be opportunities for creative empirical inactivation models that work for specific categories of products or processing technologies. This is where vigilance is needed, to ensure that predictive models are maximizing their utility by completely reporting estimated models and domains of applicability.

One aspect of predictive microbiology that has the most potential for improvement is the utilization of model validations. There are many instances where "validation" is used more as a keyword than as an informative analysis. Model validations are an underutilized tool for predictive inactivation models, and even non-predictive models. It is this author's opinion that several model validation metrics are so freely used without consideration for how the metric should be used that it has lost its sense of meaning. Future predictive modeling research should be focused on solving real problems but should also include rigorous validation approaches that test the validity of findings. There is an opportunity for a handbook of predictive microbiology model validations that clearly defines the advantages, limitations, and methodologies for various forms of internal and external validations.

To advance this field toward maximizing utility of research for use in preventive control validations, there are several immediate projects that could further improve knowledge and application:

1. *Challenging multiple model structures through internal and external validations.* Instead of using model selection procedures to reduce the models progressing to further development, this study will examine whether the model selection process was able to predict the robustness of different model forms tested. Additionally, this will challenge how non-linearity in model development datasets impact the predictivity of models that may over-fit the curve.
2. *Blind objective evaluation of model validation metric*. Model development involves a degree of "art of practice" in addition to fundamental scientific approaches; however, the metrics used to evaluate models should not be "an art" and should be objective. This study will apply different models with varying levels of success, and then use commonly utilized validation metrics to evaluate the model predictions. The validation results will then be provided to predictive modeling experts for determination of good or bad predictive models.
3. *Using predictive models as evidence to estimate surrogate-pathogen translation ratios for validation of preventive controls*. One issue identified with using predictive models as tools for quantifying the surrogate-pathogen relationship was that it was non-specific to a process, and evidence for developing translational equations may need to be specific. This study would utilize a surrogate with various classifications of quality (ideal vs. slightly conservative) and determine how different processing schedules may impact the effective resistance ratio. In theory, the effective ratio should be closest to the instantaneous reduction ratio associated with the conditions where most reductions were observed.

REFERENCES

1. Abd, S. J., K. L. McCarthy, and L. J. Harris. 2012. Impact of storage time and temperature on thermal inactivation of *Salmonella* Enteritidis PT30 on oil-roasted almonds. *J. Food Sci.* 77:M42-M47.

2. Acuff, J. C., J. S. Dickson, J. M. Farber, E. M. Grasso-Kelley, C. Hedberg, A. Lee, and M.-J. Zhu. 2023. Practice and progress: updates on outbreaks, advances in research, and processing technologies for low-moisture food safety. *J. Food Prot*. 86. 100018.

3. Acuff, J. C., J. Wu, C. Marik, K. Waterman, D. Gallagher, H. B. Huang, R. C. Williams, and M. A. Ponder. 2020. Thermal inactivation of *Salmonella*, Shiga toxin-producing *Escherichia coli*, *Listeria monocytogenes*, and a surrogate (*Pediococcus acidilactici*) on raisins, apricot halves, and macadamia nuts using vacuum-steam pasteurization. *Int. J. Food. Microbiol*. 333. 108814.

4. Ahmad, N. H., I. M. Hildebrandt, S. R. Pickens, S. Vasquez, Y. Q. Jin, S. X. Liu, L. A. Halik, H. C. Tsai, S. K. Lau, R. C. D'Souza, S. Kumar, J. Subbiah, H. Thippareddi, M. J. Zhu, J. M. Tang, N. M. Anderson, E. M. Grasso-Kelley, E. T. Ryser, and B. P. Marks. 2022. Interlaboratory evaluation of *Enterococcus faecium* NRRL B-2354 as a *Salmonella* surrogate for validating thermal treatment of multiple low-moisture foods. *J. Food Prot.*. 85:1538-1552.

5. Ahmad, N. H., B. P. Marks, and E. T. Ryser. 2022. Effect of lactose and milk protein on thermal resistance of *Enterococcus faecium* NRRL B-2354 and *Salmonella* in dairy powders. *J. Food Prot.* 85:1865-1874.

6. Ahmad, N. H., C. Oztabak, B. P. Marks, and E. T. Ryser. 2019. Effect of talc as a dry-inoculation carrier on thermal resistance of *Enterococcus faecium* NRRL B-2354 in almond meal. *J. Food Prot.* 82:1110-1115.

7. Aljarallah, K. M., and M. R. Adams. 2007. Mechanisms of heat inactivation in *Salmonella* serotype Typhimurium as affected by low water activity at different temperatures. *J. Appl. Microbiol.* 102:153-160.

8. Alliance for Innovation and Operational Excellence. 2012 Validating the reduction of *Salmonella* and other pathogens in heat processed low-moisture foods | OpX Leadership Network. Available at: https://www.opxleadershipnetwork.org/food-safety-plan-validation/download/validating-reduction-salmonella-and-other-pathogens-heat Accessed April 4, 2019.

9. Alliance for Innovation and Operational Excellence. 2019. Spotlight on baking. Available at: https://www.opxleadershipnetwork.org/food-safety-plan-validation/download/spotlight-baking Accessed April 8, 2019.

10. Almond Board of California. 2007. Guidelines for validation of blanching processes. Available at: https://www.almonds.com/sites/default/files/2020-05/blanching-validation-guidelines.pdf Accessed March 10, 2024

11. Almond Board of California. 2007. Guidelines for validation of dry roasting processes. Available at: https://www.almonds.com/sites/default/files/dry-roast-validation-guidelines.pdf Accessed March 10, 2024

12. Almond Board of California. 2007. Guidelines for validation of oil roasting processes. Available at: https://www.almonds.com/sites/default/files/2020-05/oil-roast-validation-guidelines.pdf Accessed March 10, 2024

13. Almond Board of California 2014. Guidelines for using *Enterococcus faecium* NRRL B-2354 as a surrogate microorganism in almond process validation. Available at: https://www.almonds.com/sites/default/files/guidelines_for_using_enterococcus_faecium_nrrl_b-2354_as_a_surrogate_microorganism_in_almond_process_validation.pdf Accessed March 10, 2024

14. Almond Board of California. 2023. Pasteurization Program. Available at: https://www.almonds.com/almond-industry/processors-and-suppliers/processing-safe-product/pasteurization-program Accessed March 10, 2024

15. Alshammari, J., N. Dhowlaghar, Y. C. Xie, J. Xu, J. M. Tang, S. Sablani, and M. J. Zhu. 2021. Survival of *Salmonella* and *Enterococcus faecium* in high fructose corn syrup and honey at room temperature (22°C). *Food Cont*. 123. 107765.

16. American Spice Trade Association. 2013. Validation of microbial reduction processes for spices. Available at: https://www.astaspice.org/food-safety/white-papers/ Accessed April 8, 2019.

17. Anderson, N. M., S. E. Keller, N. Mishra, S. Pickens, D. Gradl, T. Hartter, G. Rokey, C. Dohl, B. Plattner, S. Chirtel, and E. M. Grasso-Kelley. 2017. *Salmonella* inactivation during extrusion of an oat flour model food. *J. Food Sci.* 82:738-743.

18. Annous, B. A., and M. F. Kozempel. 1998. Influence of growth medium on thermal resistance of *Pediococcus sp* NRRL B-2354 (formerly *Micrococcus freudenreichii*) in liquid foods. *J. Food Prot.* 61:578-581.

19. AOAC International. 1996. Official methods of analysis. AOAC International, Gaithersburg, MD.

20. Archer, J., E. T. Jervis, J. Bird, and J. E. Gaze. 1998. Heat resistance of *Salmonella* Weltevreden in low-moisture environments. *J. Food Prot.* 61:969-973.

21. Arias-Rios, E. V., G. R. Acuff, A. Castillo, L. M. Lucia, S. E. Niebuhr, and J. S. Dickson. 2019. Identification of a surrogate to validate irradiation processing of selected spices. *Lwt-*

Food Sci. Tecnnol. 102:136-141.

22. Aryani, D. C., H. M. W. den Besten, W. C. Hazeleger, and M. H. Zwietering. 2015. Quantifying variability on thermal resistance of *Listeria monocytogenes. Int. J. Food Microbiol.* 193:130-138.

23. Aryani, D. C., M. H. Zwietering, and H. M. W. den Besten. 2016. The effect of different matrices on the growth kinetics and heat resistance of *Listeria monocytogenes* and *Lactobacillus plantarum. Int. J. Food. Microbiol*. 238:326-337.

24. Bai, X., M. Campagnoli, S. Butot, T. Putallaz, L. Michot, and S. Zuber. 2020. Inactivation by osmotic dehydration and air drying of *Salmonella*, Shiga toxin-producing *Escherichia coli*, *Listeria monocytogenes*, hepatitis A virus and selected surrogates on blueberries. *Int. J. Food. Microbiol*. 320. 108522.

25. Ballom, K., N. Dhowlaghar, H. C. Tsai, R. Yang, J. M. Tang, and M. J. Zhu. 2021. Radiofrequency pasteurization against *Salmonella* and *Listeria monocytogenes* in cocoa powder. *Lwt-Food Sci. Tecnnol.* 145. 11490.

26. Ballom, K. F., H.-C. Tsai, M. Taylor, J. Tang, and M.-J. Zhu. 2020. Stability of *Listeria monocytogenes* in non-fat dry milk powder during isothermal treatment and storage. *J. Food Microbiol.* 87. 103376.

27. Baranyi, J., and T. A. Roberts. 1995. Mathematics of predictive food microbiology. *Int. J. Food. Microbiol*. 26:199-218.

28. Bates, D. M., and D. G. Watts. 1988. Nonlinear regression analysis and its applications. Wiley, New York.

29. Bialka, K. L., A. Demirci, and V. M. Puri. 2008. Modeling the inactivation of *Escherichia coli* O157 : H7 and *Salmonella enterica* on raspberries and strawberries resulting from exposure to ozone or pulsed UV-light. *J. Food Eng.* 85:444-449.

30. Bianchini, A., J. Stratton, S. Weier, T. Hartter, B. Plattner, G. Rokey, G. Hertzel, L. Gompa, B. Martinez, and K. M. Eskridge. 2012. Validation of extrusion as a killing step for *Enterococcus faecium* in a balanced carbohydrate-protein meal by using a response surface design. *J. Food Prot.* 75:1646-1653.

31. Bianchini, A., J. Stratton, S. Weier, T. Hartter, B. Plattner, G. Rokey, G. Hertzel, L. Gompa, B. Martinez, and K. M. Eskridge. 2012. Validation of extrusion as a killing step for *Enterococcus faecium* in a balanced carbohydrate-protein meal by using a response surface design. *J. Food Prot.*75:1646-1653.

32. Bianchini, A., J. Stratton, S. Weier, T. Hartter, B. Plattner, G. Rokey, G. Hertzel, L. Gompa, B. Martinez, and K. M. Eskridge. 2014. Use of *Enterococcus faecium* as a surrogate for *Salmonella enterica* during extrusion of a balanced carbohydrate-protein meal. *J. Food Prot.*

77:75-82.

33. Blackburn, C. D., L. M. Curtis, L. Humpheson, C. Billon, and P. J. McClure. 1997. Development of thermal inactivation models for *Salmonella enteritidis* and *Escherichia coli* O157:H7 with temperature, pH and NaCl as controlling factors. *Int. J. Food. Microbio*. 38:31-44.

34. Blessington, T., E. J. Mitcham, and L. J. Harris. 2012. Survival of *Salmonella enterica*, *Escherichia coli* O157:H7, and *Listeria monocytogenes* on inoculated walnut kernels during storage. *J. Food Prot.* 75:245-254.

35. Brackett, R. E., W. Ocasio, K. Waters, J. Barach, and J. Wan. 2014. Validation and verification: a practical, industry-driven framework developed to support the requirements of the Food Safety Modernization Act (FSMA) of 2011. *Food Prot. Trends*. 34:410-425.

36. Brar, P. K., and M. D. Danyluk. 2019. Validation of *Enterococcus faecium* as a surrogate for *Salmonella* under different processing conditions for peanuts and pecans. *Food Microbiol.* 80:9-17.

37. Breslin, T. J., M. I. Tenorio-Bernal, B. P. Marks, A. M. Booren, E. T. Ryser, and N. O. Hall. 2014. Evaluation of *Salmonella* thermal inactivation model validity for slow cooking of whole-muscle meat roasts in a pilot-scale oven. *J. Food Prot.* 77:1897-1903.

38. Brul, S., S. v. Gerwen, and M. H. Zwietering. 2007. Modelling microorganisms in food. CRC Press; Woodhead Pub. Ltd., Boca Raton, FL. Cambridge, England.

39. Busta, F. F., T. V. Suslow, M. E. Parish, L. R. Beuchat, J. N. Farber, E. H. Garrett, and L. J. Harris. 2003. The use of indicators and surrogate microorganisms for the evaluation of pathogens in fresh and fresh-cut produce. *Comp. Rev. Food Sci. Food Saf.* 2:179-185.

40. Carlson, T. R., B. P. Marks, A. M. Booren, E. T. Ryser, and A. Orta-Ramirez. 2005. Effect of water activity on thermal inactivation of *Salmonella* in ground turkey. *J. Food Sci.* 70:M363-M366.

41. Casulli, K. E., S. Calhoun, and D. W. Schaffner. 2019. Modeling the risk of salmonellosis from consumption of peanuts in the United States. *J. Food Prot.* 82:579-588.

42. Casulli, K. E., K. D. Dolan, and B. P. Marks. 2021. Modeling the effects of product temperature, product moisture, and process humidity on thermal inactivation of *Salmonella* in pistachios during hot-air heating. *J. Food Prot.*. 84:47-57.

43. Ceylan, E., A. Amezquita, N. Anderson, R. Betts, L. Blayo, F. Garces-Vega, E. Gkogka, L. J. Harris, P. McClure, A. Winkler, and H. M. W. den Besten. 2021. Guidance on validation of lethal control measures for foodborne pathogens in foods. *Comp. Rev. Food Sci. Food Saf*. 20:2825-2881.

44. Ceylan, E., and D. A. Bautista. 2015. Evaluating *Pediococcus acidilactici* and *Enterococcus faecium* NRRL B-2354 as thermal surrogate microorganisms for *Salmonella* for in-plant validation studies of low-moisture pet food products. *J. Food Prot.* 78:934-939.

45. Channaiah, L. H., E. S. Holmgren, M. Michael, N. J. Sevart, D. Milke, C. L. Schwan, M. Krug, A. Wilder, R. K. Phebus, H. Thippareddi, and G. Milliken. 2016. Validation of baking to control *Salmonella* serovars in hamburger bun manufacturing, and evaluation of *Enterococcus faecium* ATCC 8459 and *Saccharomyces cerevisiae* as nonpathogenic surrogate indicators. *J. Food Prot.* 79:544-552.

46. Channaiah, L. H., M. Michael, J. Acuff, K. Lopez, D. Vega, G. Milliken, H. Thippareddi, and R. Phebus. 2018. Validation of a simulated commercial frying process to control *Salmonella* in donuts. *Foodborne Pathog. Dis.* 15:763-769.

47. Channaiah, L. H., M. Michael, J. C. Acuff, R. K. Phebus, H. Thippareddi, and G. Milliken. 2021. Thermal inactivation of *Salmonella* during hard and soft cookies baking process. *Food Microbiol*. 100. 103874.

48. Channaiah, L. H., M. Michael, J. C. Acuff, D. Vega, K. Lopez, R. K. Phebus, H. Thippareddi, and G. Milliken. 2021. Validation of simulated commercial manufacturing of flour tortillas to control *Salmonella* contamination. *J. Food Saf*. 41. e12879.

49. Chen, L., and B. S. Abigail. 2023. Surface inoculation method impacts microbial reduction and transfer of *Salmonella* Enteritidis PT 30 and potential surrogates during dry sanitation. *Int. J. Food Microbiol.* 406:110405.

50. Chen, L., J. Jung, B. D. Chaves, D. Jones, M. Negahban, Y. Y. Zhao, and J. Subbiah. 2021. Challenges of dry hazelnut shell surface for radio frequency pasteurization of inshell hazelnuts. *Food Cont*. 125. 107948.

51. Chen, L., X. Y. Wei, B. D. Chaves, D. Jones, M. A. Ponder, and J. Subbiah. 2021. Inactivation of *Salmonella enterica* and *Enterococcus faecium* NRRL B2354 on cumin seeds using gaseous ethylene oxide. *Food Microbiol*. 94. 103656.

52. Chen, Z., and J. H. Meng. 2021. Persistence of *Salmonella enterica* and *Enterococcus faecium* NRRL B-2354 on baby spinach subjected to temperature abuse after exposure to sub-lethal stresses. *Foods*. 10. 2141.

53. Chiewchan, N., W. Pakdee, and S. Devahastin. 2007. Effect of water activity on thermal resistance of *Salmonella krefeld* in liquid medium and on rawhide surface. *Int. J. Food Microbiol.* 114:43-49.

54. Chung, H. J., S. L. Birla, and J. Tang. 2008. Performance evaluation of aluminum test cell designed for determining the heat resistance of bacterial spores in foods. *Food Sci. Tecnnol.* 41:1351-1359.

55. Chung, H. J., S. J. Wang, and J. M. Tang. 2007. Influence of heat transfer with tube methods on measured thermal inactivation parameters for *Escherichia coli. J. Food Prot.* 70:851-859.

56. Coe, C., T. Boltz, R. Stearns, P. Foster, R. L. Taylor, J. Moritz, J. Jaczynski, A. Freshour, and C. L. Shen. 2022. Thermal inactivation of *Salmonella* Typhimurium and surrogate *Enterococcus faecium* in mash broiler feed in a laboratory scale circulated thermal bath. *Poult. Sci*. 101. 101976.

57. Condron, R., C. Farrokh, K. Jordan, P. McClure, T. Ross, and O. Cerf. 2015. Guidelines for experimental design protocol and validation procedure for the measurement of heat resistance of microorganisms in milk. *Int. J. Food Microbiol.* 192:20-25.

58. Cox, J. M., and A. Pavic. 2014. *Salmonella* Introduction. In R.K. Robinson, C.A. Batt, and P.D. Patel (ed.), Encyclopedia of food microbiology, vol. 3. Academic Press.

59. Cunha, L. M., F. A. R. Oliveira, and J. C. Oliveira. 1998. Optimal experimental design for estimating the kinetic parameters of processes described by the Weibull probability distribution function. *J. Food Eng*. 37:175-191.

60. Dag, D., R. K. Singh, J. R. Chen, A. Mishra, and F. B. Kong. 2022. Radio frequency assisted thermal processing for pasteurization of packaged whole milk powder surrounded by oil. *Food Control*. 135. 108762.

61. Danyluk, M. D., L. J. Harris, and D. W. Schaffner. 2006. Monte Carlo simulations assessing the risk of salmonellosis from consumption of almonds. *J. Food Prot*. 69:1594-1599.
62. Daryaei, H., W. Penaloza, I. Hildebrandt, K. Krishnamurthy, P. Thiruvengadam, and J. Wan. 2018. Heat inactivation of Shiga toxin-producing *Escherichia coli* in a selection of low moisture foods. *Food Cont*. 85:48-56.

63. Daryaei, H., Q. Sui, H. Liu, A. Rehkopf, W. Penaloza, A. Rytz, Y. Q. Luo, and J. Wan. 2020. Heat resistance of Shiga toxin-producing *Escherichia coli* and potential surrogates in wheat flour at two moisture levels. *Food Cont*. 108. 106788.

64. Dhowlaghar, N., J. M. Tang, and M. J. Zhu. 2021. Thermal inactivation of *Salmonella*, *Listeria monocytogenes* and *Enterococcus faecium* NRRL B-2354 in desiccated shredded coconut. *Lwt-Food Sci. Technol*. 149. 111851.

65. Dhowlaghar, N., and M. J. Zhu. 2022. Control of *Salmonella* in low-moisture foods: *Enterococcus faecium* NRRL B-2354 as a surrogate for thermal and non-thermal validation. *Crit. Rev. Food Sci. Nutr*. 62:5886-5902.

66. Dolan, K. D., and D. K. Mishra. 2013. Parameter estimation in food science. *Annu. Rev. Food Sci. Technol.* 4:401-422.

67. Dolan, K. D., V. P. Valdramidis, and D. K. Mishra. 2013. Parameter estimation for dynamic microbial inactivation: which model, which precision? *Food Cont*. 29:401-408.

68. Du, W. X., S. J. Abd, K. L. McCarthy, and L. J. Harris. 2010. Reduction of *Salmonella* on inoculated almonds exposed to hot oil. *J. Food Prot*. 73:1238-1246.

69. Eastwood, L. C., T. M. Taylor, J. W. Savell, K. B. Gehring, and A. N. Arnold. 2021. Efficacy of antimicrobial interventions in reducing *Salmonella enterica*, Shiga toxin-producing *Escherichia coli*, *Campylobacter*, and *Escherichia coli* biotype I surrogates on non-chilled and chilled, skin-on and skinless pork. *Meat Sci*. 172. 108309.

70. Enache, E., Y. H. Chen, G. Awuah, A. Economides, and V. N. Scott. 2006. Thermal resistance parameters for pathogens in white grape juice concentrate. *J. Food Prot.* 69:564-569.

71. Enache, E., A. Kataoka, D. G. Black, C. D. Napier, R. Podolak, and M. M. Hayman. 2015. Development of a dry inoculation method for thermal challenge studies in low-moisture foods by using talc as a carrier for *Salmonella* and a Surrogate (*Enterococcus faecium*). *J. Food Prot*. 78:1106-1112.

72. Farakos, S. M. S., J. F. Frank, and D. W. Schaffner. 2013. Modeling the influence of temperature, water activity and water mobility on the persistence of *Salmonella* in low-moisture foods. *Int. J. Food Microbiol.* 166:280-293.

73. Farakos, S. M. S., J. W. Hicks, and J. E. Frank. 2014. Temperature resistance of *Salmonella* in low-water activity whey protein powder as influenced by salt content. *J. Food Prot.* 77:631-634.

74. Farakos, S. M. S., R. Pouillot, G. R. Davidson, R. Johnson, I. Son, N. Anderson, and J. M. Van Doren. 2019. A quantitative risk assessment of human salmonellosis from consumption of walnuts in the United States. *J. Food Prot.* 82:45-57.

75. Farakos, S. M. S., R. Pouillot, G. R. Davidson, R. Johnson, J. Spungen, I. Son, N. Anderson, and J. M. Van Doren. 2018. A quantitative risk assessment of human salmonellosis from consumption of pistachios in the United States. *J. Food Prot.* 81:1001-1014.

76. Farakos, S. M. S., R. Pouillot, R. Johnson, J. Spungen, I. Son, N. Anderson, G. R. Davidson, and J. M. Van Doren. 2017. A quantitative assessment of the risk of human salmonellosis arising from the consumption of pecans in the United States. *J. Food Prot.* 80:1574-1591.

77. Farakos, S. M. S., R. Pouillot, R. Johnson, J. Spungen, I. Son, N. Anderson, and J. M. Van Doren. 2017. A quantitative assessment of the risk of human salmonellosis arising from the consumption of almonds in the United States: the impact of preventive treatment levels. *J. Food Prot.* 80:863-878.

78. Farakos, S. M. S., R. Pouillot, and S. E. Keller. 2017. *Salmonella* survival kinetics on pecans, hazelnuts, and pine nuts at various water activities and temperatures. *J. Food Prot.* 80:879-885.

79. Farakos, S. M. S., D. W. Schaffner, and J. F. Frank. 2014. Predicting survival of *Salmonella* in low-water activity foods: an analysis of literature data. *J. Food Prot.* 77:1448-1461.

80. Federal Register. 2007. Almonds grown in California; outgoing quality control requirements. Available at: http://www.gpo.gov/fdsys/granule/FR-2007-03-30/07-1557 Accessed March 10, 2024

81. Feinberg, M. 2007. Validation of analytical methods based on accuracy profiles. *J. Chromatogr. A.* 1158:174-183.

82. Feinberg, M., B. Boulanger, W. Dewé, and P. Hubert. 2004. New advances in method validation and measurement uncertainty aimed at improving the quality of chemical data. *Anal. Bioanal. Chem*. 380:502-514.

83. Finn, S., O. Condell, P. McClure, A. Amézquita, and S. Fanning. 2013. Mechanisms of survival, responses and sources of *Salmonella* in low-moisture environments. *Front. Microbiol.* 4. 331.

84. Fleischman, G. J. 2015. Reducing the experimental effort in measuring D and z values for microorganism inactivation kinetics. *J. Food Eng*. 155:1-9.

85. Forghani, F., M. den Bakker, J.-Y. Liao, A. S. Payton, A. N. Futral, and F. Diez-Gonzalez. 2019. *Salmonella* and enterohemorrhagic *Escherichia coli* serogroups O45, O121,O145 in wheat flour: effects of long-term storage and thermal treatments. *Front. Microbiol.* 10. 323.

86. Gabriel, A. A., and M. P. V. Azanza. 2010. D-72°C values of *Salmonella* Typhimurium in citrus juices: predictive efficacy of a model. *J. Food Proc. Eng*. 33:506-518.

87. Gabriel, A. A., E. B. Barrios, and M. P. V. Azanza. 2008. Modeling the thermal death of *Salmonella* Typhimurium in citrus systems. *J. Food Proc. Eng*. 31:640-657.

88. Gabriel, A. A., and E. E. C. Estilo. 2015. Influences of malic acid and nisin supplementations on the decimal reduction times of *Escherichia coli* O157:H7 in mildly-heated young coconut liquid endosperm. *Food Cont*. 50:645-651.

89. Gabriel, A. A., and H. Nakano. 2009. Inactivation of *Salmonella*, *E. coli* and *Listeria monocytogenes* in phosphate-buffered saline and apple juice by ultraviolet and heat treatments. *Food Cont*. 20:443-446.

90. Garces-Vega, F., and B. P. Marks. 2014. Use of simulation tools to illustrate the effect of data management practices for low and negative plate counts on the estimated parameters of

microbial reduction models. *J. Food Prot*. 77:1372-1379.

91. Garces-Vega, F. J., E. T. Ryser, and B. P. Marks. 2019. Relationships of water activity and moisture content to the thermal inactivation kinetics of *Salmonella* in low-moisture foods. *J. Food Prot.* 82:963-970.

92. Garre, A., A. Pielaat, M. H. Zwietering, H. Besten, and J. H. Smid. 2022. Critical comparison of statistical methods for quantifying variability and uncertainty of microbial responses from experimental data. *Int. J. Food Microbiol*. 383. 109935.

93. Georgalis, L., A. Psaroulaki, A. Aznar, P. S. Fernandez, and A. Garre. 2022. Different model hypotheses are needed to account for qualitative variability in the response of two strains of *Salmonella spp*. under dynamic conditions. *Food Res. Int*. 158. 11477.

94. Gil, M. M., F. A. Miller, T. R. S. Brandao, and C. L. M. Silva. 2017. Mathematical models for prediction of temperature effects on kinetic parameters of microorganisms' inactivation: tools for model comparison and adequacy in data fitting. *Food Bioprocess Technol*. 10:2208-2225.

95. Gil, M. M., F. A. Miller, C. L. M. Silva, and T. R. S. Brandao. 2014. Application of optimal experimental design concept to improve the estimation of model parameters in microbial thermal inactivation kinetics. *J. Food Eng*. 134:59-66.

96. Grasso-Kelley, E. M., X. Y. Liu, L. A. Halik, and B. Douglas. 2021. Evaluation of hot-air drying to inactivate *Salmonella* and *Enterococcus faecium* on apple pieces. *J. Food Prot*. 84:240-248.

97. Grijspeerdt, K., and L. Herman. 2003. Inactivation of *Salmonella* Enteritidis during boiling of eggs. *Int. J. Food Microbiol.* 82:13-24.

98. Grocery Manufacturers Association. 2008. Guidelines for validation of consumer cooking instructions for not-ready-to-eat (NRTE) products. Available at: https://www.gmaonline.org/forms/store/ProductFormPublic/guidelines-for-validation-of-consumer-cooking-instructions-for-not-ready-to-eat-nrte-products Accessed April 8, 2019.

99. Grocery Manufacturers Association. 2009. Control of *Salmonella* in low-moisture foods. Available at: https://www.gmaonline.org/forms/store/ProductFormPublic/SalmonellaControlGuidance Accessed April 8, 2019.

100. Gu, K. J., A. S. Sekhon, J. K. Richter, Y. Yang, E. Pietrysiak, M. Michael, and G. M. Ganjyal. 2022. Heat resistance comparison of *Salmonella* and *Enterococcus faecium* in cornmeal at different moisture levels. *Int. J. Food Microbiol*. 368. 109608.

101. Guan, J. W., A. Lacombe, J. M. Tang, D. F. Bridges, S. Sablani, B. Rane, and V. C. H. Wu. 2021. Use of mathematic models to describe the microbial inactivation on baby carrots by

gaseous chlorine dioxide. *Food Cont*. 123. 107832.

102. Harris, L. J., A. R. Uesugi, S. J. Abd, and K. L. McCarthy. 2012. Survival of *Salmonella* Enteritidis PT30 on inoculated almond kernels in hot water treatments. *Food Res. Int.* 45:1093-1098.

103. He, Y. S., D. J. Guo, J. Y. Yang, M. L. Tortorello, and W. Zhang. 2011. Survival and heat resistance of *Salmonella enterica* and *Escherichia coli* O157:H7 in peanut butter. *Appl. Environ. Microbiol.* 77:8434-8438.

104. He, Y. S., Y. Li, J. K. Salazar, J. Y. Yang, M. L. Tortorello, and W. Zhang. 2013. Increased water activity reduces the thermal resistance of *Salmonella enterica* in peanut butter. *Appl. Environ. Microbiol.* 79:4763-4767.

105. Henz, S., R. Nitzsche, M. Kiessling, K. Aganovic, V. Heinz, and C. Hertel. 2020. Surrogate for electron beam inactivation of *Salmonella* on pumpkin seeds and flax seeds. *J. Food Prot*. 83:1775-1781.

106. Hildebrandt, I. M., N. O. Hall, M. K. James, E. T. Ryser, and B. P. Marks. 2021. Process humidity affects *Salmonella* lethality at the surface and core of impingement-cooked meat and poultry products. *J. Food Prot*. 84:1512-1523.

107. Hildebrandt, I. M., B. P. Marks, V. K. Juneja, M. Osoria, N. O. Hall, and E. T. Ryser. 2016. Cross-laboratory comparative study of the impact of experimental and regression methodologies on *Salmonella* thermal inactivation parameters in ground beef. *J. Food Prot.* 79:1097-1106.

108. Hildebrandt, I. M., B. P. Marks, E. T. Ryser, R. Villa-Rojas, J. Tang, F. J. Garces-Vega, and S. E. Buchholz. 2016. Effects of inoculation procedures on variability and repeatability of *Salmonella* thermal resistance in wheat flour. *J. Food Prot.* 79:1833-1839.

109. Hildebrandt, I. M., C. Hu, E. M. Grasso-Kelley, P. Ye, N. M. Anderson, and S. E. Keller. 2017. Dry transfer inoculation of low-moisture spices containing antimicrobial compounds. *J. Food Prot*. 80:338-344.

110. Hu, M., and J. B. Gurtler. 2017. Selection of surrogate bacteria for use in food safety challenge studies: a review. *J. Food Prot.* 80:1506-1536.

111. Huang, L. 2016. Mathematical modeling and validation of growth of *Salmonella* Enteritidis and background microorganisms in potato salad - One-step kinetic analysis and model development. *Food Cont*. 68:69-76.

112. Huang, L. H. 2009. Thermal inactivation of *Listeria monocytogenes* in ground beef under isothermal and dynamic temperature conditions. *J. Food Eng.* 90:380-387.

113. Huang, L. H., and V. K. Juneja. 2001. A new kinetic model for thermal inactivation of microorganisms: Development and validation using *Escherichia coli* O157 : H7 as a test organism. *J. Food Prot*. 64:2078-2082.

114. Institute of Food Technologists Science Advisory Board. 2000. Overarching principles: kinetics and pathogens of concern for all technologies. *J. Food Sci*. 65:16-31.

115. International Commission on Microbiological Specifications for Food. 2011. Microorganisms in foods 8: Use of data for assessing process control and product acceptance. Springer US : Imprint: Springer

116. Izurieta, W. P., and E. Komitopoulou. 2012. Effect of moisture on *Salmonella* spp. heat resistance in cocoa and hazelnut shells. *Food Res. Int.* 45:1087-1092.

117. Jagannath, A., and T. Tsuchido. 2003. Predictive microbiology: a review. *Biocont. Sci*. 8:1-7.

118. Janssen, M., A. Verhulst, V. Valdramidis, F. Devlieghere, J. F. Van Impe, and A. H. Geeraerd. 2008. Inactivation model equations and their associated parameter values obtained under static acid stress conditions cannot be used directly for predicting inactivation under dynamic conditions. *Int. J. Food Microbiol.* 128:136-145.

119. Jeong, S., B. P. Marks, and M. K. James. 2017. Comparing thermal process validation methods for *Salmonella* inactivation on almond kernels. *J. Food Prot.* 80:169-176.

120. Jeong, S., B. P. Marks, and A. Orta-Ramirez. 2009. Thermal inactivation kinetics for *Salmonella* Enteritidis PT30 on almonds subjected to moist-air convection heating. *J. Food Prot.* 72:1602-1609.

121. Jeong, S., B. P. Marks, and E. T. Ryser. 2011. Quantifying the performance of *Pediococcus* sp. (NRRL B-2354: *Enterococcus faecium*) as a nonpathogenic surrogate for *Salmonella* Enteritidis PT30 during moist-air convection heating of almonds. *J. Food Prot.* 74:603-609.

122. Jeong, S., B. P. Marks, E. T. Ryser, and J. B. Harte. 2012. The effect of X-ray irradiation on *Salmonella* inactivation and sensory quality of almonds and walnuts as a function of water activity. *Int. J. Food Microbiol.* 153:365-371.

123. Jewell, K. 2012. Comparison of 1-step and 2-step methods of fitting microbiological models. *Int. J. Food Microbiol*. 160:145-161.

124. Jiang, W. T., C. Waldman, K. W. Li, J. Jaczynski, and C. L. Shen. 2021. Survival of *Salmonella* and the surrogate *Enterococcus faecium* in cooking of moisture enhanced reconstructed comminuted chicken patties by double pan-broiling. *Poult. Sci*. 100. 101171.

125. Jin, Y., T. Juming, and Z. Mei-Jun. 2020. Water activity influence on the thermal resistance of *Salmonella* in soy protein powder at elevated temperatures. *Food Cont*. 113:107160-

107160.

126. Jin, Y., S. R. Pickens, I. M. Hildebrandt, S. J. Burbick, E. M. Grasso-Kelley, S. E. Keller, and N. M. Anderson. 2018. Thermal inactivation of *Salmonella* Agona in low-water activity foods: predictive models for the combined effect of temperature, water activity, and food component. *J. Food Prot.* 81:1411-1417.

127. Jin, Y., and J. Tang. 2019. Improved design of aluminum test cell to study the thermal resistance of *Salmonella enterica* and *Enterococcus faecium* in low-water activity foods. *Food Cont*. 104:343-348.

128. Jongenburger, I., H. M. W. den Besten, and M. H. Zwietering. 2015. Statistical Aspects of Food Safety Sampling. *Annu. Rev. Food Sci. Technol.* 6:479-503.

129. Juneja, V. K., and B. S. Eblen. 2000. Heat inactivation of *Salmonella* Typhimurium DT104 in beef as affected by fat content. *Lett. Appl. Microbiol.* 30:461-467.

130. Juneja, V. K., B. S. Eblen, and H. M. Marks. 2001. Modeling non-linear survival curves to calculate thermal inactivation of *Salmonella* in poultry of different fat levels. *Int. J. Food Microbiol.* 70:37-51.

131. Juneja, V. K., B. S. Eblen, and G. M. Ransom. 2001. Thermal inactivation of *Salmonella* spp. in chicken broth, beef, pork, turkey, and chicken: determination of D- and Z-values. *J. Food Sci.* 66:146-152.

132. Juneja, V. K., H. M. Marks, and T. Mohr. 2003. Predictive thermal inactivation model for effects of temperature, sodium lactate, NaCl, and sodium pyrophosphate on *Salmonella* serotypes in ground beef. *Appl. Environ. Microbiol.* 69:5138-5156.

133. Jung, J., and D. W. Schaffner. 2022. Enterobacter aerogenes B199A may be an effective surrogate for quantifying transfer of *Salmonella* Newport 96E01152C-TX from cucumber peel to edible flesh and peeler during peeling. *J. Food Prot*. 85:1452-1457.

134. Jung, J., and D. W. Schaffner. 2022. Thermal inactivation of *Salmonella enterica* and nonpathogenic bacterial surrogates in wheat flour by baking in a household oven. *J. Food Prot*. 85:1431-1438.

135. Karolenko, C. E., J. Wilkinson, and P. M. Muriana. 2022. Evaluation of various lactic acid bacteria and generic *E. coli* as potential nonpathogenic surrogates for in-plant validation of biltong dried beef processing. *Microorganisms*. 10. 1648.

136. Keller, S. E., J. M. VanDoren, E. M. Grasso, and L. A. Halik. 2013. Growth and survival of *Salmonella* in ground black pepper (*Piper nigrum*). *Food Microbiol.* 34:182-188.

137. Kimber, M. A., H. Kaur, L. X. Wang, M. D. Danyluk, and L. J. Harris. 2012. Survival of *Salmonella*, *Escherichia coli* O157:H7, and *Listeria monocytogenes* on inoculated almonds

and pistachios stored at -19, 4, and 24°C. *J. Food Prot.* 75:1394-1403.

138. Kornacki, J. 2012. *Enterococcus faecium* NRRL B-2354: Tempest in a teapot or serious foodborne pathogen? *Food Safety Magazine.* Available at: https://www.food-safety.com/articles/3786-enterococcus-faecium-nrrl-b-2354-tempest-in-a-teapot-or-serious-foodborne-pathogen Accessed: March 10, 2024

139. Krapf, T., and C. Gantenbein-Demarchi. 2010. Thermal inactivation of *Salmonella* spp. during conching. *Lwt-Food Sci. Technol.* 43:720-723.

140. Lakshmikantha, H. C., M. Michael, J. C. Acuff, K. Lopez, R. K. Phebus, T. Harshavardhan, and G. Milliken. 2019. Validation of a nut muffin baking process and thermal resistance characterization of a 7-serovar *Salmonella* inoculum in batter when introduced via flour or walnuts. *Int. J. Food Microbiol.* 294:27-30.

141. Lakshmikantha, H. C., M. Michael, J. C. Acuff, R. K. Phebus, H. Thippareddi, M. Olewnik, and G. Milliken. 2017. Validation of the baking process as a kill-step for controlling *Salmonella* in muffins. *Int. J. Food Microbiol.* 250:1-6.

142. Lambertini, E., J. Barouei, D. W. Schaffner, M. D. Danyluk, and L. J. Harris. 2017. Modeling the risk of salmonellosis from consumption of pistachios produced and consumed in the United States. *Food Microbiol*. 67:85-96.

143. Lambertini, E., M. D. Danyluk, D. W. Schaffner, C. K. Winter, and L. J. Harris. 2012. Risk of salmonellosis from consumption of almonds in the North American market. *Food Res. Int.* 45:1166-1174.

144. Lau, S. K., R. Panth, B. D. Chaves, C. L. Weller, and J. Subbiah. 2021. Thermal inactivation kinetics of *Salmonella* and *Enterococcus faecium* NRRL-B2354 on whole chia seeds (*Salvia hispanica L.*). *J. Food Prot*. 84:1357-1365.

145. Li, C. C., L. H. Huang, and J. Q. Chen. 2014. Comparative study of thermal inactivation kinetics of *Salmonella* spp. in peanut butter and peanut butter spread. *Food Cont*. 45:143-149.

146. Li, H., G. H. Xie, and A. Edmondson. 2007. Evolution and limitations of primary mathematical models in predictive microbiology. *Brit. Food J.* 109:608-626.

147. Li, K. W., Y. C. Chiu, W. T. Jiang, L. Jones, X. Etienne, and C. L. Shen. 2020. Comparing the efficacy of two triple-wash procedures with sodium hypochlorite, a lactic-citric acid blend, and a mix of peroxyacetic acid and hydrogen peroxide to inactivate *Salmonella*, *Listeria monocytogenes*, and surrogate *Enterococcus faecium* on cucumbers and tomatoes. *Front. Sustain. Food Syst*. 4. 19.

148. Lianou, A., and K. P. Koutsoumanis. 2013. Evaluation of the strain variability of *Salmonella enterica* acid and heat resistance. *Food Microbiol.* 34:259-267.

149. Limcharoenchat, P., S. E. Buchholz, M. K. James, N. O. Hall, E. T. Ryser, and B. P. Marks. 2018. Inoculation protocols influence the thermal resistance of *Salmonella* Enteritidis PT 30 in fabricated almond, wheat, and date products. *J. Food Prot.* 81:606-613.

150. Limcharoenchat, P., M. K. James, and B. P. Marks. 2019. Survival and thermal resistance of *Salmonella* Enteritidis PT 30 on almonds after long-term storage. *J. Food Prot.* 82:194-199.

151. Liu, S., S. Ozturk, J. Xu, F. Kong, P. Gray, M.-J. Zhu, S. S. Sablani, and J. Tang. 2018. Microbial validation of radio frequency pasteurization of wheat flour by inoculated pack studies. *J. Food Eng.* 217:68-74.

152. Liu, S., R. V. Rojas, P. Gray, M.-J. Zhu, and J. Tang. 2018. *Enterococcus faecium* as a *Salmonella* surrogate in the thermal processing of wheat flour: Influence of water activity at high temperatures. *Food Microbiol.* 74:92-99.

153. Liu, S., J. Tang, R. K. Tadapaneni, R. Yang, and M.-J. Zhu. 2018. Exponentially increased thermal resistance of *Salmonella* spp. and *Enterococcus faecium* at reduced water activity. *Appl. Environ. Microbiol.* 84. e02742-17.

154. Liu, S., J. Xu, L. Xie, M.-J. Zhu, and J. Tang. 2019. Dry inoculation methods for nonfat milk powder. *J. Dairy Sci.* 102:77-86.

155. Liu, Z. S., L. A. Sheng, S. S. Canakapalli, and L. X. Wang. 2022. Evaluation of the microbial control efficacies of commonly used home-drying and storage practices of dried peaches. *Lwt-Food Sci. Technol*. 165. 113734.

156. Livingstone, D., A. M. Davis, and Royal Society of Chemistry (Great Britain). 2012. Drug design strategies quantitative approaches. p. xvii, 498 p. *In*, RSC drug discovery series no 13 RSC Pub, Cambridge.

157. Lucore, L. A., A. J. Gualtieri, and S. J. Abd. 2017. A thermal process lethality model for low water activity food. *Food Prot. Trends*. 37:43-55.

158. Ma, L., J. L. Kornacki, G. D. Zhang, C. M. Lin, and M. P. Doyle. 2007. Development of thermal surrogate microorganisms in ground beef for in-plant critical control point validation studies. *J. Food Prot.* 70:952-957.

159. Ma, L., G. D. Zhang, P. Gerner-Smidt, V. Mantripragada, I. Ezeoke, and M. P. Doyle. 2009. Thermal inactivation of *Salmonella* in peanut butter. *J. Food Prot.* 72:1596-1601.

160. Mafart, P., O. Couvert, S. Gaillard, and I. Leguerinel. 2002. On calculating sterility in thermal preservation methods: application of the Weibull frequency distribution model. *Int.*

J. Food Microbiol. 72:107-113.

161. Manas, P., R. Pagan, J. Raso, and S. Condon. 2003. Predicting thermal inactivation in media of different pH of *Salmonella* grown at different temperatures. *Int. J. Food Microbiol.* 87:45-53.

162. Martino, K. G., and B. P. Marks. 2007. Comparing uncertainty resulting from two-step and global regression procedures applied to microbial growth models. *J. Food Prot.* 70:2811-2818.

163. Mattick, K. L., F. Jorgensen, P. Wang, J. Pound, M. H. Vandeven, L. R. Ward, J. D. Legan, H. M. Lappin-Scott, and T. J. Humphrey. 2001. Effect of challenge temperature and solute type on heat tolerance of *Salmonella* serovars at low water activity. *Appl. Environ. Microbiol.* 67:4128-4136.

164. Mattick, K. L., J. D. Legan, T. J. Humphrey, and M. Peleg. 2001. Calculating *Salmonella* inactivation in nonisothermal heat treatments from isothermal nonlinear survival curves. *J. Food Prot.* 64:606-613.

165. McCann, M. S., D. A. McDowell, and J. J. Sheridan. 2009. Effects of reduction in beef surface water activity on the survival of *Salmonella* Typhimurium DT104 during heating. *J. Appl. Microbiol.* 106:1901-1907.

166. McClure, P., T. Roberts, J. Baranyi, and C. Adair. 2004. Thermal inactivation models - Validation and application. *Bulletin of the International Dairy Federation*. 111-115

167. McDonald, K., and D. W. Sun. 1999. Predictive food microbiology for the meat industry: a review. *Int. J. Food Microbiol*. 52:1-27.

168. McKellar, R. C., and X. Lu. 2004. Modeling microbial responses in food. CRC Press.

169. McMeekin, T. A., J. Olley, D. A. Ratkowsky, and T. Ross. 2002. Predictive microbiology: towards the interface and beyond. *Int. J. Food Microbiol*. 73:395-407.

170. Mee, R. W. 1984. Tolerance limits and bounds for proportions based on data subject to measurement error. *J. Qual. Technol*. 16:74-80.

171. Michael, M., J. Acuff, K. Lopez, D. Vega, R. Phebus, H. Thippareddi, and L. H. Channaiah. 2020. Comparison of survival and heat resistance of *Escherichia coli* O121 and *Salmonella* in muffins. *Int. J. Food Microbiol*. 317. 108422.

172. Miller, F. A., M. M. Gil, T. R. S. Brandao, P. Teixeira, and C. L. M. Silva. 2009. Sigmoidal thermal inactivation kinetics of *Listeria innocua* in broth: Influence of strain and growth phase. *Food Cont.* 20:1151-1157.

173. Miller, F. A., B. Ramos, M. M. Gil, T. R. S. Brandao, P. Teixeira, and C. L. M. Silva. 2009. Influence of pH, type of acid and recovery media on the thermal inactivation of *Listeria innocua. Int. J. Food Microbiol.* 133:121-128.

174. Miller, F. A., B. F. Ramos, M. M. Gil, T. R. S. Brandão, P. Teixeira, and C. L. M. Silva. 2011. Heat inactivation of *Listeria innocua* in broth and food products under non-isothermal conditions. *Food Cont.* 22:20–26.

175. Mogollon, M. A., B. P. Marks, A. M. Booren, A. Orta-Ramirez, and E. T. Ryser. 2009. Effect of beef product physical structure on *Salmonella* thermal inactivation. *J. Food Sci.* 74:M347-M351.

176. Motulsky, H., and A. Christopoulos. 2004. Fitting models to biological data using linear and nonlinear regression: a practical guide to curve fitting. Oxford University Press, USA.

177. Moussavi, M., J. C. Frelka, I. M. Hildebrandt, B. P. Marks, and L. J. Harris. 2020. Thermal resistance of foodborne pathogens and *Enterococcus faecium* NRRL B-2354 on inoculated pistachios. *J. Food Prot.* 83(7):1125-1136

178. Murdoch, M., A. Waser, G. Morantes, B. Dubovcova, G. Akepsimaidis, A. Currie, and S. D. Pillai. 2022. A new proposed validation method for low energy electron beam processing of dry spices. *Innov. Food Sci. Emerg. Technol*. 81. 103141.

179. Murphy, R. Y., L. K. Duncan, B. L. Beard, and K. H. Driscoll. 2003. D and z values of *Salmonella*, *Listeria innocua*, and *Listeria monocytogenes* in fully cooked poultry products. *J. Food Sci.* 68:1443-1447.

180. Murphy, R. Y., L. K. Duncan, E. R. Johnson, M. D. Davis, and J. A. Marcy. 2002. Thermal inactivation of *Salmonella* Senftenberg and *Listeria innocua* in beef/turkey blended patties cooked via fryer and/or air convection oven. *J. Food Sci*. 67:1879-1885.

181. Murphy, R. Y., L. K. Duncan, E. R. Johnson, M. D. Davis, and J. N. Smith. 2002. Thermal inactivation D- and z-values of *Salmonella* serotypes and *Listeria innocua* in chicken patties, chicken tenders, franks, beef patties, and blended beef and turkey patties. *J. Food Prot.* 65:53-60.

182. Murphy, R. Y., B. P. Marks, E. R. Johnson, and M. G. Johnson. 2000. Thermal inactivation kinetics of *Salmonella* and *Listeria* in ground chicken breast meat and liquid medium. *J. Food Sci.* 65:706-710.

183. Murphy, R. Y., E. M. Martin, L. K. Duncan, B. L. Beard, and J. A. Marcy. 2004. Thermal process validation for *Escherichia coli* O157:H79 *Salmonella*, and *Listeria monocytogenes* in ground turkey and beef products. *J. Food Prot.* 67:1394-1402.

184. Murphy, R. Y., T. Osaili, L. K. Duncan, and J. A. Marcy. 2004. Thermal inactivation of *Salmonella* and *Listeria monocytogenes* in ground chicken thigh/leg meat and skin. *Poult.*

Sci. 83:1218-1225.

185. National Advisory Committee on Microbiological Criteria for Foods. 2010. Parameters for determining inoculated pack/challenge study protocols. *J. Food Prot*. 73:140-202.

186. Natrella, M. 1963. Experimental Statistics, National Bureau of Standards Handbook 91. *In* U.S. Department of Commerce.

187. Newkirk, J. J., J. Wu, J. C. Acuff, C. B. Caver, K. Maffikaijunan, B. D. Wiersema, R. C. Williams, and M. A. Ponder. 2018. Inactivation of *Salmonella enterica* and surrogate *Enterococcus faecium* on whole black peppercorns and cumin seeds using vacuum steam pasteurization. *Front. Sustain. Food Syst*. 2. 48.

188. OECD. 2014. Guidance document on the validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] models. Available at: https://www.oecd-ilibrary.org/docserver/9789264085442-en.pdf?expires=1710121243&id=id&accname=guest&checksum=D7CE3E490EA74B3F576BEBBA9620FBE7 Accessed March 10, 2024

189. Orta-Ramirez, A., B. P. Marks, C. R. Warsow, A. M. Booren, and E. T. Ryser. 2005. Enhanced thermal resistance of *Salmonella* in whole muscle compared to ground beef. *J. Food Sci.* 70:M359-M362.

190. Oscar, T. P. 2009. General regression neural network and Monte Carlo simulation model for survival and growth of *Salmonella* on raw chicken skin as a function of serotype, temperature, and time for use in risk assessment. *J. Food Prot.* 72:2078-2087.

191. Oscar, T. P. 2013. Validation of a predictive model for survival and growth of *Salmonella* Typhimurium DT104 on chicken skin for extrapolation to a previous history of frozen storage. *J. Food Prot*. 76:1035-1040.

192. Oscar, T. R. 2005. Validation of lag, time and growth rate models for *Salmonella* Typhimurium: acceptable prediction zone method. *J. Food Sci.* 70:M129-M137.

193. Ozturk, S., F. B. Kong, and R. K. Singh. 2020. Evaluation of *Enterococcus faecium* NRRL B-2354 as a potential surrogate of *Salmonella* in packaged paprika, white pepper and cumin powder during radio frequency heating. *Food Cont*. 108. 106833.

194. Peleg, M., and M. B. Cole. 1998. Reinterpretation of microbial survival curves. *Crit. Rev. Food Sci. Nutr.* 38:353-380.

195. Peleg, M., and M. D. Normand. 2004. Calculating microbial survival parameters and predicting survival curves from non-isothermal inactivation data. *Crit. Rev. Food Sci. Nutr*. 44:409-418.

196. Perez-Reyes, M. E., X. Jie, M. J. Zhu, J. M. Tang, and G. V. Barbosa-Canovas. 2021. Influence of low water activity on the thermal resistance of *Salmonella* Enteritidis PT30 and *Enterococcus faecium* as its surrogate in egg powders. *Food Sci. Technol. International*. 27:184-193.

197. Pflug, I. J. 2005. Variability in the data generated by laboratories measuring D-values of bacterial spores. *PDA J. Pharm. Sci. Technol*. 59:3-9.

198. Picard, R. R., and R. D. Cook. 1984. Cross-validation of regression-models. *J. Am. Stat. Assoc.* 79:575-583.

199. Piyasena, P., R. C. McKellar, and F. M. Bartlett. 2003. Thermal inactivation of *Pediococcus* sp in simulated apple cider during high-temperature short-time pasteurization. *Int. J. Food Microbiol*. 82:25-31.

200. Portela, J. B., P. T. Coimbra, L. P. Cappato, V. O. Alvarenga, R. B. A. Oliveira, K. S. Pereira, D. R. P. Azeredo, A. S. Sant'Ana, J. S. Nascimento, and A. G. Cruz. 2019. Predictive model for inactivation of *Salmonella* in infant formula during microwave heating processing. *Food Cont.* 104:308-312.

201. Quinn, A. R., R. F. Liao, F. M. Steele, L. K. Jefferies, and B. J. Taylor. 2021. Isothermal inactivation of *Salmonella, Listeria monocytogenes*, and *Enterococcus faecium* NRRL B-2354 in peanut butter, powder infant formula, and wheat flour. *Food Contr*. 121. 107582.

202. Rachon, G., W. Penaloza, and P. A. Gibbs. 2016. Inactivation of *Salmonella*, *Listeria monocytogenes* and *Enterococcus faecium* NRRL B-2354 in a selection of low moisture foods. *Int. J. Food Microbiol.* 231:16-25.

203. Rane, B., A. Lacombe, S. Sablani, D. F. Bridges, J. M. Tang, J. W. Guan, and V. C. H. Wu. 2021. Effects of moisture content and mild heat on the ability of gaseous chlorine dioxide against *Salmonella* and *Enterococcus faecium* NRRL B-2354 on almonds. *Food Contr*. 123. 107732.

204. Redemann, M. A., J. Brar, S. E. Niebuhr, L. M. Lucia, G. R. Acuff, J. S. Dickson, and M. Singh. 2018. Evaluation of thermal process lethality for non-pathogenic *Escherichia coli* as a surrogate for *Salmonella* in ground beef. *Lwt-Food Sci. Technol.* 90:290-296.

205. Ross, T. 1996. Indices for performance evaluation of predictive models in food microbiology. *J. Appl. Bacteriol.* 81:501-508.

206. Ross, T., T. A. McMeekin, and J. Baranyi. 2014. Predictive Microbiology and Food Safety. *In* R.K. Robinson, C.A. Batt, and P.D. Patel (ed.), Encyclopedia of food microbiology, vol. 3. Academic Press.

207. Scharff, R. L. 2012. Economic burden from health losses due to foodborne illness in the United States. *J. Food Prot*. 75:123-131.

208. Schultze, D. M., R. Couto, F. Temelli, L. M. McMullen, and M. Ganzle. 2020. Lethality of high-pressure carbon dioxide on Shiga toxin-producing *Escherichia coli*, *Salmonella* and surrogate organisms on beef jerky. *Int. J. Food Microbiol*. 321. 108550.

209. Shah, U. R., Q. Y. Wang, S. Kathariou, and D. Salvi. 2023. Optimization of plasma-activated water and validation of a potential surrogate for *Salmonella* for future egg washing processes. *J. Food Prot*. 86. 100029.

210. Shen, X. Y., Y. Su, Z. Hua, J. Cong, N. Dhowlaghar, Q. Sun, S. N. Lin, T. N. Green, M. Perrault, M. Galeni, I. Hanrahan, T. V. Suslow, and M. J. Zhu. 2020. Verification of peroxyacetic acid treatment against *L. monocytogenes* on fresh apples using *E. faecium* NRRL B-2354 as a surrogate in commercial spray-bar operations. *Food Microbiol*. 92. 103590.

211. Sheng, L. N., X. Y. Shen, and M. J. Zhu. 2020. Screening of non-pathogenic surrogates of *Listeria monocytogenes* applicable for chemical antimicrobial interventions of fresh apples. *Food Cont*. 110. 106977.

212. Smith, D. F., I. M. Hildebrandt, K. E. Casulli, K. D. Dolan, and B. P. Marks. 2016. Modeling the effect of temperature and water activity on the thermal resistance of *Salmonella* Enteritidis PT 30 in wheat flour. *J. Food Prot.* 79:2058-2065.

213. Smith, D. F., and B. P. Marks. 2015. Effect of rapid product desiccation or hydration on thermal resistance of *Salmonella enterica* serovar Enteritidis PT 30 in wheat flour. *J. Food Prot*. 78:281-286.

214. Smith, S. E., J. L. Maurer, A. Orta-Ramirez, E. T. Ryser, and D. M. Smith. 2001. Thermal inactivation of *Salmonella* spp., *Salmonella* Typhimurium DT104, and *Escherichia coli* O157:H7 in ground beef. *J. Food Sci.* 66:1164-1168.

215. Stasiewicz, M. J., B. P. Marks, A. Orta-Ramirez, and D. M. Smith. 2008. Modeling the effect of prior sublethal thermal history on the thermal inactivation rate of *Salmonella* in ground turkey. *J. Food Prot.* 71:279-285.

216. Steghofer, S., R. Limburn, and E. Margas. 2021. Microbiological assessment of heat treatment of broiler mash at laboratory scale to evaluate *Salmonella* reduction during feed conditioning. *J. Appl. Poult. Res*. 30. 100122.

217. Steinbrunner, P., B. P. Marks, E. T. Ryser, Q. J. Suehr, and S. Jeong. 2021. Fate of *Salmonella* and *Enterococcus faecium* during pilot-scale spray drying of soy protein isolate. *J. Food Prot*. 84:674-679.

218. Steinbrunner, P. J., P. Limcharoenchat, Q. J. Suehr, E. T. Ryser, B. P. Marks, and S. Jeong. 2019. Effect of food structure, water activity, and long-term storage on X-ray irradiation for inactivating *Salmonella* Enteritidis PT30 in low-moisture foods. *J. Food Prot.* 82:1405-1411.

219. Steyerberg, E. W., F. E. Harrell, G. Borsboom, M. J. C. Eijkemans, Y. Vergouwe, and J. D. F. Habbema. 2001. Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. *J. Clin. Epidemiol*. 54:774-781.

220. Steyerberg, E. W., and F. E. Harrell, Jr. 2016. Prediction models need appropriate internal, internal-external, and external validation. *J. Clin. Epidemiol*. 69:245-247.

221. Sudarsan, A., and K. M. Keener. 2022. Inactivation of *Salmonella enterica* serovars and *Escherichia coli* O157:H7 surrogate from baby spinach leaves using high voltage atmospheric cold plasma (HVACP). *Lwt-Food Sci. Technol*. 155. 112903.

222. Suehr, Q. J., N. M. Anderson, and S. E. Keller. 2019. Desiccation and thermal resistance of *Escherichia coli* O121 in wheat flour. *J. Food Prot.* 82:1308-1313.

223. Suehr, Q. J., X. Liu, E. M. Grasso-Kelley, and N. M. Anderson. 2021. Predictive microbial modeling of *Enterococcus faecium* NRRL B-2354 inactivation during baking of a multicomponent low-moisture food. *J. Food Prot*. 84:1990-2001.

224. Syamaladevi, R. M., R. K. Tadapaneni, J. Xu, R. Villa-Rojas, J. M. Tang, B. Carter, S. Sablani, and B. Marks. 2016. Water activity change at elevated temperatures and thermal resistance of *Salmonella* in all purpose wheat flour and peanut butter. *Food Res. Int.* 81:163-170.

225. Tadapaneni, R. K., R. M. Syamaladevi, R. Villa-Rojas, and J. Tang. 2017. Design of a novel test cell to study the influence of water activity on the thermal resistance of *Salmonella* in low-moisture foods. *J. Food Eng.* 208:48-56.

226. Tadapaneni, R. K., J. Xu, R. Yang, and J. Tang. 2018. Improving design of thermal water activity cell to study thermal resistance of *Salmonella* in low-moisture foods. *Lwt-Food Sci. Technol.* 92:371-379.

227. Takhar, P. S., K. L. Head, K. M. Hendrix, and D. M. Smith. 2009. Predictive modeling of *Salmonella* species inactivation in ground pork and turkey during cooking. *Int. J. Food Eng.* 5:20.

228. Taylor, M. H., H. C. Tsai, B. Rasco, J. M. Tang, and M. J. Zhu. 2018. Stability of *Listeria monocytogenes* in wheat flour storage and isothermal treatment. *Food Cont*. 91:434-439.

229. Temple, J., R. Stearns, C. Coe, H. Chaney, J. Tou, A. Freshour, Y. C. Luo, and C. L. Shen. 2023. Evaluation of the efficacy of a mixer of hydrogen peroxide and peroxyacetic acid to mitigate microbial cross-contamination of *Salmonella* Typhimurium and the surrogate

Enterococcus faecium during triple-washing of butternut squash. *Lwt-Food Sci. Technol*. 173. 114313.

230. Tenorio-Bernal, M. I., B. P. Marks, E. T. Ryser, and A. M. Booren. 2013. Evaluating the predictive ability of a path-dependent thermal inactivation model for *Salmonella* subjected to prior sublethal heating in ground turkey, beef, and pork. *J. Food Prot.* 76:220-226.

231. Tsai, H. C., K. F. Ballom, S. Xia, J. M. Tang, B. P. Marks, and M. J. Zhu. 2019. Evaluation of *Enterococcus faecium* NRRL B-2354 as a surrogate for *Salmonella* during cocoa powder thermal processing. *Food Microbiol.* 82:135-141.

232. U.S. Centers for Disease Control and Prevention. 2024. Outbreaks Involving Salmonella | CDC. Available at https://www.cdc.gov/salmonella/outbreaks.html Accessed: March 10, 2024

233. U.S. Department of Agriculture Agricultural Marketing Service. 2020. Commercial item description crackers. Available at: https://www.ams.usda.gov/sites/default/files/media/CID%20Crackers.pdf Accessed July 1, 2020.

234. U.S. Food and Drug Administration. 2018. FDA Food Safety Modernization Act (FSMA), Available at: https://www.fda.gov/food/guidance-regulation-food-and-dietary-supplements/food-safety-modernization-act-fsma Accessed March 10, 2024

235. U.S. Food and Drug Administration. 2023. Hazard analysis and risk-based preventive controls for human food: guidance for industry draft guidance. Available at: https://www.fda.gov/media/172318/download?attachment Accessed: March 10, 2024

236. Uesugi, A. R., M. D. Danyluk, and L. J. Harris. 2006. Survival of *Salmonella* Enteritidis phage type 30 on inoculated almonds stored at -20, 4, 23, and 35°C. *J. Food Prot*. 69:1851-1857.

237. Unger, P., L. H. Channaiah, A. Singh, A. S. Sekhon, M. Babb, Y. Yang, and M. Michael. 2021. Validation of brownie baking step for controlling *Salmonella* and *Listeria monocytogenes. Food Sci. Nutr*. 9:1574-1583.

238. Valdramidis, V. P., A. H. Geeraerd, J. E. Gaze, A. Kondjoyan, A. R. Boyd, H. L. Shaw, and J. F. Van Impe. 2006. Quantitative description of *Listeria monocytogenes* inactivation kinetics with temperature and water activity as the influencing factors; model prediction and methodological validation on dynamic data. *J. Food Eng.* 76:79-88.

239. van Asselt, E. D., and M. H. Zwietering. 2006. A systematic approach to determine global thermal inactivation parameters for various food pathogens. *Int. J. Food Microbiol.* 107:73-82.

240. van Boekel, M. 2002. On the use of the Weibull model to describe thermal inactivation of microbial vegetative cells. *Int. J. Food Microbiol.* 74:139-159.

241. van Boekel, M. 2008. Kinetic modeling of food quality: A critical review. *Compr. Rev. Food Sci. Food Saf.* 7:144-158.

242. van Boekel, M. 2021. To pool or not to pool: That is the question in microbial kinetics*. *Int. J. Food Microbiol*. 354. 109283.

243. van Boekel, M. 1996. Statistical aspects of kinetic modeling for food science problems. *J. Food Sci.* 61:477-+.

244. Velasquez, A., T. J. Breslin, B. P. Marks, A. Orta-Ramirez, N. O. Hall, A. M. Booren, and E. T. Ryser. 2010. Enhanced thermal resistance of *Salmonella* in marinated whole muscle compared with ground pork. *J. Food Prot.* 73:372-375.

245. Verma, T., B. D. Chaves, S. Irmak, and J. Subbiah. 2021. Pasteurization of dried basil leaves using radio frequency heating: A microbial challenge study and quality analysis. *Food Cont*. 124. 107932.

246. Verma, T., and J. Subbiah. 2019. Conical twin-screw extrusion is an effective inactivation process for *Salmonella* in low-moisture foods at temperatures above 65 °C. *Lwt-Food Sci. Technol.* 114. 108369.

247. Verma, T., X. Wei, S. K. Lau, A. Bianchini, K. M. Eskridge, and J. Subbiah. 2018. Evaluation of *Enterococcus faecium* NRRL B-2354 as a surrogate for *Salmonella* during extrusion of low-moisture food. *J. Food Sci.* 83:1063-1072.

248. Verma, T., X. Y. Wei, B. D. Chaves, T. Howell, and J. Subbiah. 2022. Antimicrobial efficacy of gaseous chlorine dioxide for inactivation of *Salmonella* and *Enterococcus faecium* NRRL B-2354 on dried basil leaves. *Lwt-Food Sci. Technol*. 153. 112488.

249. Verma, T., X. Y. Wei, S. K. Lau, A. Bianchin, K. M. Eskrisge, J. Stratton, N. M. Anderson, H. Thippareddi, and J. Subbiah. 2018. Response surface methodology for *Salmonella* inactivation during extrusion processing of oat flour. *J. Food Prot.* 81:815-826.

250. Verma, T., X. Y. Wei, S. K. Lau, A. Bianchini, K. M. Eskridge, and J. Subbiah. 2018. Evaluation of *Enterococcus faecium* NRRL B-2354 as a surrogate for *Salmonella* during extrusion of low-moisture food. *J. Food Sci*. 83:1063-1072.

251. Villa-Rojas, R., J. M. Tang, S. J. Wang, M. X. Gao, D. H. Kang, J. H. Mah, P. Gray, M. E. Sosa-Morales, and A. Lopez-Malo. 2013. Thermal inactivation of *Salmonella* Enteritidis PT30 in almond kernels as influenced by water activity. *J. Food Prot.* 76:26-32.

252. Villa-Rojas, R., M.-J. Zhu, B. P. Marks, and J. Tang. 2017. Radiofrequency inactivation of *Salmonella* Enteritidis PT 30 and *Enterococcus faecium* in wheat flour at different water

activities. *Biosyst. Eng.* 156:7-16.

253. Wang, H. Y., Z. Chen, M. Dharmasena, A. K. Greene, B. M. Gardener, B. Holden, and X. P. Jiang. 2021. Plant-scale validation of physical heat treatment of poultry litter composts using surrogate and indicator microorganisms for *Salmonella*. *Appl. Environ. Microbiol*. 87. e02234-20

254. Wang, Y., Z. Zhang, A. Xu, and H. Zhang. 2017. A study on methods for establishing surrogate safety goals for nuclear power. *Nucl. Technol*. 198:327-341.

255. Wason, S., and J. Subbiah. 2023. Gaseous chlorine dioxide for inactivating *Salmonella enterica* and *Enterococcus faecium* NRRL B-2354 on chia seeds. *Food Cont*. 150. 109736.

256. Wason, S., T. Verma, S. Irmak, and J. Subbiah. 2022. In-package pasteurization of dried basil leaves using radiofrequency heating. *Food Cont*. 141. 109179.

257. Wason, S., T. Verma, X. Y. Wei, A. Mauromoustakos, and J. Subbiah. 2022. Thermal inactivation kinetics of *Salmonella enterica* and *Enterococcus faecium* NRRL B-2354 as a function of temperature and water activity in fine ground black pepper. *Food Res. Int*. 157. 111393.

258. Wei, X., S. K. Lau, B. D. Chaves, M.-G. C. Danao, S. Agarwal, and J. Subbiah. 2020. Effect of water activity on the thermal inactivation kinetics of *Salmonella* in milk powders. *J. Dairy Sci*. 103:6904-6917.

259. Wei, X. Y., S. Agarwal, and J. Subbiah. 2021. Evaluation of *Enterococcus faecium* NRRL B-2354 as a surrogate for *Salmonella enterica* in milk powders at different storage times and temperatures. *J. Dairy Sci*. 104:198-210.

260. Wei, X. Y., L. Chen, B. D. Chaves, M. A. Ponder, and J. Subbiah. 2021. Modeling the effect of temperature and relative humidity on the ethylene oxide fumigation of *Salmonella* and *Enterococcus faecium* in whole black peppercorn. *Lwt-Food Sci. Technol*. 140. 110742.

261. Wei, X. Y., S. K. Lau, B. S. Reddy, and J. Subbiah. 2020. A microbial challenge study for validating continuous radio-frequency assisted thermal processing pasteurization of egg white powder. *Food Microbiol*. 85. 103306.

262. Wei, X. Y., S. Vasquez, H. Thippareddi, and J. Subbiah. 2021. Evaluation of *Enterococcus faecium* NRRL B-2354 as a surrogate for *Salmonella* in ground black pepper at different water activities. *Int. J. Food Microbiol*. 344. 109114.

263. Wei, X. Y., T. Verma, M. G. C. Danao, M. A. Ponder, and J. Subbiah. 2021. Gaseous chlorine dioxide technology for improving microbial safety of spices. *Innov. Food Sci. Emerg. Technol*. 73. 102783.

264. Wei, X. Y., T. Verma, S. Irmak, and J. Subbiah. 2023. Effect of storage on microbial reductions after gaseous chlorine dioxide treatment of black peppercorns, cumin seeds, and dried basil leaves. *Food Cont*. 148. 109627.

265. Whiting, R. C. 1995. Microbial modeling in foods. *Crit. Rev. Food Sci. Nutr.* 35:467-494.

266. Whiting, R. C., and R. L. Buchanan. 1993. A classification of models in predictive microbiology - reply. *Food Microbiol*. 10:175-177.

267. Wiser, B., S. E. Niebuhr, and J. S. Dickson. 2022. Impact of interventions on the survival of *Salmonella enterica* I 4, 5 ,1 2:i:- in pork. *J. Food Prot*. 85:27-30.

268. Woerner, D. R., I. Geornaras, J. N. Martin, K. E. Belk, G. R. Acuff, and J. S. Dickson. 2018. Use of nonpathogenic *Escherichia coli* surrogates as predictors of the survival of nontyphoidal *Salmonella*, non-O157 shiga toxin-producing *Escherichia coli*, and *Escherichia coli* O157 populations after high hydrostatic pressure processing. *J. Food Prot.* 81:1068-1072.

269. Xu, J., S. Liu, J. Song, J. Tang, M.-J. Zhu, P. Gray, and R. Villa-Rojas. 2018. Dry-inoculation method for thermal inactivation studies in wheat flour using freeze-dried *Enterococcus faecium* NRRL B-2354. *Lwt-Food Sci. Technol.* 89:10-17.

270. Xu, J., S. X. Liu, J. M. Tang, S. Ozturk, F. B. Kong, and D. H. Shah. 2018. Application of freeze-dried *Enterococcus faeciuin* NRRL B-2354 in radio-frequency pasteurization of wheat flour. *Lwt-Food Sci. Technol.* 90:124-131.

271. Xu, J., J. M. Tang, Y. Q. Jin, J. X. Song, R. Yang, S. S. Sablani, and M. J. Zhu. 2019. High temperature water activity as a key factor influencing survival of *Salmonella* Enteritidis PT30 in thermal processing. *Food Cont*. 98:520-528.

272. Yang, W. T., Y. J. Yi, and B. Xia. 2023. Unveiling the duality of *Pantoea dispersa*: A mini review. *Sci. Total Environ*. 873. 162320.

273. Zhou, Z. J., S. Zuber, M. Campagnoli, M. Moser, F. Devlieghere, and M. Uyttendaele. 2020. Decontamination effect of hot-air drying against bacterial pathogen and surrogate strains on basil leaves, from laboratory to pilot scale settings. *Lwt-Food Sci. Technol*. 122. 109035.

274. Zhu, M., X. Song, X. Shen, and J. Tang. 2020. *Listeria monocytogenes* in almond meal: desiccation stability and isothermal inactivation. *Front. Microbiol*. 11. 1689.

275. Zhu, M. J., X. Song, H. C. Tsai, X. Y. Shen, M. Taylor, and J. M. Tang. 2021. Desiccation and thermal resistance of *Salmonella* and *Enterococcus faecium* NRRL B-2354 in almond meal as impacted by water activity and storage temperature. *Food Cont*. 126. 108037.

APPENDIX A. EXPERIMENTAL CONTEXT AND SUMMARY OF MODEL COMPLETENESS AND PREDICTIVITY FOR STUDIES REVIEWED IN CHAPTER 3.

The material included in this appendix summarizes the general experimental context for the 100 studies reviewed as part of the analysis included in Chapter 3. The reference numbers (Ref. #) included in the table correspond to the reference table. In this table, "Complete" refers to whether the studied adhered to the Good Standards of Practice for Predictive Modeling principles 1-3 (a defined endpoint, unambiguous methodologies and model, reported appropriate measures of model estimates, uncertainty, and variability) when reporting the predictive inactivation model(s). Models were also defined as predictive (implicitly and explicitly described the domain of applicability), unqualified predictive (implicitly defines an applicability domain but overall applicability was not explicitly qualified by the authors), and non-predictive (inactivation model utilized for non-predictive purposes).

*Cultivation of cultures was followed by a drying step for a dry inoculation of the product

** Experiment type listed refers to portion of the study used for model development or to which the model was applied

Table A.1 Experimental context and summary of model completeness and predictivity

Ref. #	Year	Product	Last Cultivation Medium	Microorganism(s) Studied	Experiment Type**	Complete	Applicability Domain
(1)	2012	Almond	Agar	*Salmonella enterica*	Thermal Treatment		Non-predictive
(6)	2019	Almond, Talc	Agar	*Enterococcus faecium*	Isothermal	X	Non-predictive
(7)	2007	Broth	Broth	*Salmonella enterica*	Isothermal		Non-predictive
(17)	2017	Oat Flour	Agar	*Salmonella enterica*	Thermal Treatment		Predictive
(22)	2015	Broth	Broth	*Listeria monocytogenes*	Isothermal		Non-predictive
(26)	2020	Non-fat Dry Milk Powder	Agar	*Listeria monocytogenes* Cocktail	Isothermal		Predictive
(29)	2008	Raspberry, Strawberry	Broth	*Salmonella enterica* Cocktail, *Escherichia coli* Cocktail	Chemical, Radiation		Unqualified Predictive
(30)	2012	Pet Food	Broth	*Enterococcus faecium*	Thermal Treatment		Non-predictive
(34)	2012	Walnut	Agar	*Escherichia coli* Cocktail, *Salmonella enterica* Cocktail, *Listeria monocytogenes* Cocktail	Survival		Non-predictive
(40)	2005	Ground Turkey	Broth	*Salmonella enterica* Cocktail	Isothermal		Non-predictive

Table A.1 (cont'd)

Ref. #	Year	Product	Last Cultivation Medium	Microorganism(s) Studied	Experiment Type**	Complete	Applicability Domain
(44)	2015	Pet Food	Agar	*Salmonella enterica* Cocktail, *Enterococcus faecium*, *Pediococcus acidilactici*	Isothermal		Non-predictive
(45)	2016	Hamburger Bun Dough	Agar	*Salmonella enterica* Cocktail, *Enterococcus faecium*, *Saccharomyces cerevisiae*	Isothermal		Predictive
(46)	2018	Donut Batter	Agar	*Salmonella enterica* Cocktail	Isothermal		Predictive
(53)	2007	Broth, Rawhide	Broth	*Salmonella enterica*	Isothermal		Non-predictive
(55)	2007	Potato	Broth	*Escherichia coli*	Isothermal		Non-predictive
(62)	2018	Dry Cat Mix, Chicken Meat Powder, Seasoning, Confectionery	Agar	*Escherichia coli* Cocktail	Thermal Treatment	X	Predictive
(68)	2010	Almond	Agar	*Salmonella enterica* Cocktail	Thermal Treatment		Predictive
(70)	2006	Grape Juice Concentrate	Broth	*Salmonella enterica* Cocktail, *Escherichia coli* Cocktail, *Listeria monocytogenes* Cocktail	Isothermal		Predictive
(72)	2013	Whey Protein, Peanut Meal, Cocoa Powder, Wheat Flour, Non-Fat Dry Milk	Broth*	*Salmonella enterica* Cocktail	Isothermal, survival		Predictive
(73)	2014	Whey Protein	Broth*	*Salmonella enterica* Cocktail	Isothermal	X	Non-predictive

Table A.1 (cont'd)

Ref. #	Year	Product	Last Cultivation Medium	Microorganism(s) Studied	Experiment Type**	Complete	Applicability Domain
(85)	2019	Wheat Flour	Broth	*Escherichia coli* Cocktail, *Salmonella enterica* Cocktail	Isothermal, Survival		Non-predictive
(87)	2008	Citrus	Broth	*Salmonella enterica*	Isothermal	X	Predictive
(88)	2015	Coconut	Broth	*Escherichia coli* Cocktail	Isothermal		Predictive
(89)	2009	Broth, Apple Juice	Broth	*Salmonella enterica*, *Escherichia coli*, *Listeria monocytogenes*	Isothermal, Radiation		Non-predictive
(91)	2019	Almond	Agar	*Salmonella enterica*	Isothermal		Non-predictive
(97)	2003	Egg	Broth	*Salmonella enterica*	Thermal Treatment		Predictive
(102)	2012	Almond	Agar	*Salmonella enterica*, *Enterococcus faecalis*	Thermal Treatment		Non-predictive
(103)	2011	Peanut Butter	Broth*	*Salmonella enterica* Cocktail, *Escherichia coli* Cocktail	Isothermal		Non-predictive
(104)	2013	Peanut Butter	Broth	*Salmonella enterica*, *Salmonella enterica* Cocktail	Isothermal		Non-predictive
(108)	2016	Ground Beef	Broth	*Salmonella enterica* Cocktail	Isothermal	X	Predictive

Table A.1 (cont'd)

Ref. #	Year	Product	Last Cultivation Medium	Microorganism(s) Studied	Experiment Type**	Complete	Applicability Domain
(109)	2016	Wheat flour	Broth, Agar	*Salmonella enterica*	Isothermal	X	Non-predictive
(112)	2009	Ground Beef	Broth	*Listeria monocytogenes* Cocktail	Isothermal, Thermal Treatment		Unqualified Predictive
(113)	2001	Ground Beef	Broth	*Escherichia coli* Cocktail	Isothermal		Unqualified Predictive
(116)	2012	Cocoa Shells, Hazelnut Shells	Agar	*Salmonella enterica*	Isothermal		Non-predictive
(118)	2008	Broth	Broth	*Listeria monocytogenes*	Chemical		Non-predictive
(120)	2009	Almond	Agar	*Salmonella enterica*	Thermal Treatment		Predictive
(122)	2012	Almond, Walnut	Agar	*Salmonella enterica*	Radiation		Non-predictive
(126)	2018	High Protein Dough, High Fat Dough	Agar	*Salmonella enterica*	Isothermal	X	Predictive
(127)	2019	Brown Rice Flour	Agar	*Salmonella enterica* Cocktail, *Enterococcus faecium*	Isothermal		Non-predictive (Surrogate)
(129)	2000	Ground Beef, Chicken Broth	Broth	*Salmonella enterica* Cocktail	Isothermal		Predictive

Table A.1 (cont'd)

Ref. #	Year	Product	Last Cultivation Medium	Microorganism(s) Studied	Experiment Type**	Complete	Applicability Domain
(130)	2001	Ground Chicken, Ground Turkey	Broth	*Salmonella enterica* Cocktail	Isothermal		Predictive
(131)	2001	Chicken Broth, Beef, Pork, Turkey, Chicken	Broth	*Salmonella enterica*, *Salmonella enterica* Cocktail	Isothermal		Predictive
(132)	2003	Ground Beef	Broth	*Salmonella enterica* Cocktail	Isothermal	X	Unqualified Predictive
(137)	2012	Almond, Pistachio	Agar	*Salmonella enterica* Cocktail, *Escherichia coli* Cocktail, *Listeria monocytogenes* Cocktail	Survival		Non-predictive
(139)	2010	Chocolate, Cocoa Butter	Agar	*Salmonella enterica*	Isothermal		Non-predictive
(140)	2019	Nut Muffin Batter, Walnut	Agar	*Salmonella enterica* Cocktail	Isothermal		Predictive
(141)	2017	Muffin Batter	Agar	*Salmonella enterica* Cocktail	Isothermal		Predictive
(145)	2014	Peanut Butter, Peanut Butter Spread	Broth	*Salmonella enterica* Cocktail	Isothermal		Predictive
(158)	2007	Ground Beef	Broth	*Listeria monocytogenes*, *Escherichia faecium*, *Escherichia coli*	Isothermal		Non-predictive (Surrogate)
(148)	2013	Broth	Broth	*Salmonella enterica*	Isothermal, chemical	X	Non-predictive

Table A.1 (cont'd)

Ref. #	Year	Product	Last Cultivation Medium	Microorganism(s) Studied	Experiment Type**	Complete	Applicability Domain
(149)	2018	Almond, Wheat, Date	Agar	*Salmonella enterica*	Isothermal	X	Non-predictive
(150)	2019	Almond	Agar	*Salmonella enterica*	Isothermal	X	Non-predictive
(151)	2018	Wheat Flour	Agar	*Salmonella enterica*, *Enterococcus faecium*	Isothermal, Thermal Treatment		Predictive
(152)	2018	Wheat Flour	Agar	*Salmonella enterica*, *Enterococcus faecium*	Isothermal		Non-predictive (Surrogate)
(153)	2018	Silicon Dioxide	Agar	*Salmonella enterica*, *Enterococcus faecium*	Isothermal		Non-predictive (Surrogate)
(154)	2019	Non-fat Dry Milk Powder	Agar	*Salmonella enterica* Cocktail, *Enterococcus faecium*	Isothermal		Non-predictive (Surrogate)
(159)	2009	Peanut Butter	Broth	*Salmonella enterica* Cocktail	Isothermal		Unqualified Predictive
(161)	2003	Broth	Broth	*Salmonella enterica*	Isothermal		Non-predictive
(163)	2001	High Sugar Broth	Broth	*Salmonella enterica*	Isothermal		Non-predictive
(164)	2001	Broth	Broth	*Salmonella enterica*	Isothermal, Thermal Treatment		Non-predictive

Table A.1 (cont'd)

Ref. #	Year	Product	Last Cultivation Medium	Microorganism(s) Studied	Experiment Type**	Complete	Applicability Domain
(165)	2009	Beef	Broth	*Salmonella enterica*	Isothermal		Non-predictive
(171)	2020	Muffin Batter	Agar	*Escherichia coli* Cocktail, *Salmonella enterica* Cocktail	Isothermal		Predictive
(172)	2009	Broth	Broth	*Listeria innocua*	Isothermal		Non-predictive
(173)	2009	Broth	Broth	*Listeria innocua*	Isothermal		Non-predictive
(174)	2011	Meat Pocket, Broth, Parsley	Broth	*Listeria innocua*	Thermal Treatment		Non-predictive
(175)	2009	Beef, Ground Beef	Broth	*Salmonella enterica* Cocktail	Isothermal		Non-predictive
(179)	2003	Chicken, Turkey, Duck	Broth	*Salmonella enterica* Cocktail, *Listeria monocytogenes* Cocktail, *Listeria innocua*	Isothermal		Predictive
(180)	2002	Blended Ground Beef/Turkey	Broth	*Salmonella enterica*, *Listeria innocua*	Thermal Treatment		Predictive
(181)	2002	Chicken, Ground Chicken, Franks, Blended Ground Beef/Turkey	Broth	*Salmonella enterica* Cocktail, *Listeria innocua*	Isothermal		Unqualified Predictive
(182)	2000	Ground Chicken, Broth	Broth	*Salmonella enterica* Cocktail, *Listeria monocytogenes*	Isothermal		Non-predictive

Table A.1 (cont'd)

Ref. #	Year	Product	Last Cultivation Medium	Microorganism(s) Studied	Experiment Type**	Complete	Applicability Domain
(183)	2004	Ground Beef, Ground Turkey	Broth	*Salmonella enterica* Cocktail, *Listeria monocytogenes* Cocktail, *Escherichia coli* Cocktail	Isothermal, Thermal Treatment		Predictive
(184)	2004	Ground Chicken	Broth	*Salmonella enterica* Cocktail, *Listeria monocytogenes* Cocktail	Isothermal		Predictive
(189)	2005	Beef, Ground Beef	Broth	*Salmonella enterica* Cocktail	Isothermal		Non-predictive
(200)	2019	Infant Formula	Broth	*Salmonella enterica*	Thermal Treatment		Predictive
(202)	2016	Confectionary, Seasoning, Chicken Meat Powder, Pet Food	Agar	*Salmonella enterica* Cocktail, *Listeria monocytogenes* Cocktail, *Enterococcus faecium*	Isothermal		Unqualified Predictive
(204)	2018	Ground Beef	Broth	*Escherichia coli* Cocktail, *Salmonella enterica* Cocktail	Isothermal		Non-predictive
(212)	2016	Wheat flour	Agar	*Salmonella enterica*	Isothermal		Predictive
(214)	2001	Ground Beef	Broth	*Salmonella enterica*, *Salmonella enterica* Cocktail, *Escherichia coli*	Isothermal		Non-predictive
(215)	2008	Ground Turkey	Broth	*Salmonella enterica* Cocktail	Isothermal, Thermal Treatment		Non-predictive
(218)	2019	Almond, Wheat, Date	Agar	*Salmonella enterica*	Radiation		Non-predictive

Table A.1 (cont'd)

Ref. #	Year	Product	Last Cultivation Medium	Microorganism(s) Studied	Experiment Type**	Complete	Applicability Domain
(222)	2019	Wheat Flour	Agar	*Escherichia coli* Cocktail	Isothermal		Non-predictive
(224)	2016	Wheat Flour, Peanut Butter	Agar	*Salmonella enterica*	Isothermal		Non-predictive
(225)	2017	Wheat Flour	Agar	*Salmonella enterica*	Isothermal		Non-predictive
(226)	2018	Wheat Flour	Agar	*Salmonella enterica*	Isothermal		Non-predictive
(227)	2009	Turkey, Pork	Broth	*Salmonella enterica* Cocktail	Isothermal, Thermal Treatment		Unqualified Predictive
(228)	2018	Wheat Flour	Agar	*Listeria monocytogenes* Cocktail	Isothermal		Unqualified Predictive
(230)	2013	Ground Beef, Ground Turkey, Ground Pork	Broth	*Salmonella enterica* Cocktail	Thermal Treatment		Predictive
(231)	2019	Cocoa	Agar	*Salmonella enterica* Cocktail, *Enterococcus faecium*	Isothermal		Non-predictive (Surrogate)
(238)	2006	Potato	Broth	*Listeria monocytogenes*	Isothermal, Thermal Treatment		Non-predictive
(244)	2010	Pork, Ground Pork	Broth	*Salmonella enterica* Cocktail	Isothermal		Non-predictive

Table A.1 (cont'd)

Ref. #	Year	Product	Last Cultivation Medium	Microorganism(s) Studied	Experiment Type**	Complete	Applicability Domain
(246)	2019	Oat flour	Agar	*Salmonella enterica* Cocktail	Thermal Treatment		Predictive
(247)	2018	Oat Flour	Agar	*Salmonella enterica* Cocktail, *Enterococcus faecium*	Thermal Treatment		Non-predictive (Surrogate)
(249)	2018	Oat Flour	Agar	*Salmonella enterica* Cocktail	Thermal Treatment	X	Predictive
(251)	2013	Almond Flour	Broth	*Salmonella enterica*	Isothermal		Unqualified Predictive
(258)	2020	Milk Powders	Agar	*Salmonella enterica* Cocktail	Isothermal		Predictive
(269)	2018	Wheat Flour	Broth*	*Enterococcus faecium*	Isothermal		Non-predictive
(270)	2018	Wheat Flour	Broth*	*Enterococcus faecium*	Isothermal, Thermal Treatment		Non-predictive (Surrogate)
(271)	2019	Wheat Flour, Almond Flour, Whey Protein	Agar	*Salmonella enterica*	Isothermal		Non-predictive
(125)	2020	Soy Protein	Agar	*Salmonella enterica*	Isothermal		Non-predictive
(274)	2020	Almond Meal	Agar	*Listeria monocytogenes* Cocktail	Isothermal		Non-predictive

APPENDIX B. SUMMARY OF INACTIVATION MODEL REGRESSION AND REPORTING PRACTICES FOR THE STUDIES EXAMINED IN CHAPTER 3

The material included in this appendix summarizes the general experimental context for the 100 studies reviewed as part of the analysis included in Chapter 3. The reference numbers (Ref. #) included in the table correspond to the reference table. In this table, "Complete" refers to whether the studied adhered to the Good Standards of Practice for Predictive Modeling principles 1-3 (a defined endpoint, unambiguous methodologies and model, reported appropriate measures of model estimates, uncertainty, and variability) when reporting the predictive inactivation model(s). Models were also defined as predictive (implicitly and explicitly described the domain of applicability), unqualified predictive (implicitly defines an applicability domain but overall applicability was not explicitly qualified by the authors), and non-predictive (inactivation model utilized for non-predictive purposes).

Key for the columns Error Estimates, Goodness-of-fit Metrics, and Model Selection

AIC - Akaike information criterion

CI - Confidence interval of parameter estimates

MSE - Mean square error of the model

R2 - Coefficient of determination (standard or adjusted)

RE - Relative error of the parameter estimate

RMSE - Root mean square error of the model

RSE - Relative standard error of the parameter estimates

SD - Standard deviation of parameter estimates

SE - Standard error of the parameter estimates

SEM - Standard error of the mean of parameter estimates

SHW - Standard half-width of the confidence interval

Table B.1 Summary of inactivation model regression and reporting practices

Ref. #	Primary Models	Secondary Models	Regression Features**	Error Estimates*	Goodness-of-fit Metrics	Model Selection	Complete	Applicability Domain
(1)	Weibull (Eqn. 2)		N/A	CI	R2			Non-predictive
(6)	Log-linear		N/A	RMSE, CI	RMSE, CI		X	Non-predictive
(7)	Log-linear	Bigelow-type	Incomplete, Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	SDp				Non-predictive
(17)	Response Surface Model		N/A	SE				Predictive
(22)	Weibull (Eqn. 3)	Bigelow-type	Two-step Estimation, Missing Secondary Parameter(s)	RMSE, CIs, SDp				Non-predictive
(26)	Log-linear	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	SEM	R2			Predictive
(29)	Linear, Weibull (Eqn. 3)		N/A	RMSE	R2	R2, RMSE, Validation		Unqualified Predictive
(30)	Response Surface Model		N/A	RMSE	R2			Non-predictive
(34)	Linear, Baranyi, Gompertz		N/A		R2	R2		Non-predictive
(40)	Linear	Arrhenius	Two-step Estimation, Missing Secondary Parameter(s)	RMSE, SDp	R2	RMSE		Non-predictive

Table B.1 (cont'd)

Ref. #	Primary Models	Secondary Models	Regression Features**	Error Estimates*	Goodness-of-fit Metrics	Model Selection	Complete	Applicability Domain
(44)	Log-linear	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	CIp	R2			Non-predictive
(45)	Log-linear	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	SE	R2			Predictive
(46)	Log-linear	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	SE	R2			Predictive
(53)	Weibull (Eqn. 2)		Incomplete	SDp	R2			Non-predictive
(55)	Log-linear	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	SD	R2			Non-predictive
(62)	Log-linear, Weibull (Eqn. 3)	Bigelow-type	One-step Estimation	RMSE, SE	R2, RMSE	Convergence Failure	X	Predictive
(68)	Weibull (Eqn. 2)		N/A	CI	R2			Predictive
(70)	Log-linear	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	SD				Predictive
(72)	Log-linear, Weibull (Eqn. 3), Biphasic, Geeraerd, Baranyi	Response Surface Model	Two-step Estimation	RMSEp, SE	R2	R2, RMSE, F-test		Predictive
(73)	Weibull (Eqn. 3)		N/A	RMSE, SE	R2		X	Non-predictive

Table B.1 (cont'd)

Ref. #	Primary Models	Secondary Models	Regression Features**	Error Estimates*	Goodness-of-fit Metrics	Model Selection	Complete	Applicability Domain
(85)	Weibull (Eqn. 3)		N/A	SE	R2			Non-predictive
(87)	Log-linear	Response Surface Model	Linearized Approximation, Two-step Estimation	MSE, SE, SD	MSE	Significance Testing	X	Predictive
(88)	Log-linear		Linearized Approximation	SD				Predictive
(89)	Log-linear		Linearized Approximation	SD				Non-predictive
(91)	Log-linear, Weibull (Eqn. 3)		N/A	RMSE, RSE, CI	RMSE, RE, CI, Analysis of Residuals	High RE		Non-predictive
(97)	Geeraerd	Bigelow-type	One-step Estimation	CI	F-score	F-test		Predictive
(102)	Log-linear	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)		R2			Non-predictive
(103)	Log-linear, Weibull (Eqn. 3)		Incomplete, Missing Primary Parameter(s)	SD	R2			Non-predictive
(104)	Log-linear, Weibull (Eqn. 3)		Incomplete, Missing Primary Parameter(s)	SD	R2			Non-predictive
(108)	Log-linear, Weibull (Eqn. 3)	Bigelow-type	One-step Estimation, Two-step Estimation	RMSE, SE	RMSE	AIC	X	Predictive

Table B.1 (cont'd)

Ref. #	Primary Models	Secondary Models	Regression Features**	Error Estimates*	Goodness-of-fit Metrics	Model Selection	Complete	Applicability Domain
(109)	Log-linear		N/A	RMSE, SE	RMSE		X	Non-predictive
(112)	Log-linear, Weibull (Eqn. 2), Gompertz	Bigelow-type, Log-linear, Linear	Linearized Approximation, Two-step Estimation	RMSE	R2, RMSE	RMSE		Unqualified Predictive
(113)	Log-linear, Modified Sigmoidal	Bigelow-type, Polynomial	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	SE	R2, Visual Inspection	R2, Visual Inspection		Unqualified Predictive
(116)	Log-linear	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	SD				Non-predictive
(118)	Weibull (Eqn. 3)	pH model	One-step Estimation	RMSE, CI	R2			Non-predictive
(120)	Log-linear	Bigelow-type	One-step Estimation	RMSE	R2, RMSE	Validation		Predictive
(122)	Log-linear		N/A	CI				Non-predictive
(126)	Log-linear, Weibull (Eqn. 3)	Bigelow-type	One-step Estimation	RMSE, SE	RMSE, Residual Analysis	Significance Testing	X	Predictive
(127)	Log-linear, Weibull (Eqn. 3)	Bigelow-type, Log-logistic	Two-step Estimation, Missing Secondary Parameter(s)	RMSE	RMSE			Non-predictive (Surrogate)
(129)	Log-linear	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	RMSE, SDp	R2			Predictive

Table B.1 (cont'd)

Ref. #	Primary Models	Secondary Models	Regression Features**	Error Estimates*	Goodness-of-fit Metrics	Model Selection	Complete	Applicability Domain
(130)	Log-linear	Bigelow-type, Polynomial	Two-step Estimation, Missing Secondary Parameter(s)	SE, CI	R2, RMSE			Predictive
(131)	Log-linear	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	RMSE, SD	R2			Predictive
(132)	Modified Sigmoidal	Polynomial	Two-step Estimation	RMSE, SE	R2, Regression Error	Likelihood Elimination	X	Unqualified Predictive
(137)	Linear, Baranyi, Gompertz		N/A		R2	R2		Non-predictive
(139)	Log-linear	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)					Non-predictive
(140)	Log-linear	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	SE	R2			Predictive
(141)	Log-linear	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	SE	R2			Predictive
(145)	Weibull (Eqn. 2)	Linear	Incomplete, Two-step Estimation	SD	R2			Predictive
(158)	Log-linear	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	SDp	R2			Non-predictive (Surrogate)
(148)	Log-linear, Geeraerd		N/A	RMSE, SD	R2		X	Non-predictive

Table B.1 (cont'd)

Ref. #	Primary Models	Secondary Models	Regression Features**	Error Estimates*	Goodness-of-fit Metrics	Model Selection	Complete	Applicability Domain
(149)	Log-linear, Weibull (Eqn. 3)		N/A	SE, RMSE, Replication Error	RMSE	AIC	X	Non-predictive
(150)	Log-linear, Weibull (Eqn. 3)		N/A	RMSE, SE	RMSE	AIC	X	Non-predictive
(151)	Log-linear, Weibull (Eqn. 3)	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	RMSE, SE	R2, RMSE			Predictive
(152)	Log-linear, Weibull (Eqn. 3)	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	RMSE, SDp	RMSE, Accuracy/ Bias Factors	RMSE, Accuracy/ Bias Factors		Non-predictive (Surrogate)
(153)	Log-linear	Bigelow-type	Incomplete, Two-step Estimation, Missing Secondary Parameter(s)	RMSE, SDp	R2			Non-predictive (Surrogate)
(154)	Log-linear, Weibull (Eqn. 3)	Bigelow-type	Two-step Estimation, Missing Secondary Parameter(s)	SEMp	R2			Non-predictive (Surrogate)
(159)	Log-linear, Weibull (Eqn. 3)	Bigelow-type, Log-linear	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	SDp	R2			Unqualified Predictive
(161)	Log-linear	Bigelow-type, Response Surface Model	Linearized Approximation, Two-step Estimation	CI	R2			Non-predictive
(163)	Weibull (Eqn. 2)	Response Surface Model	Two-step Estimation		R2	F-test		Non-predictive
(164)	Weibull (Eqn. 2)	Exponential	Two-step Estimation					Non-predictive

Table B.1 (cont'd)

Ref. #	Primary Models	Secondary Models	Regression Features**	Error Estimates*	Goodness-of-fit Metrics	Model Selection	Complete	Applicability Domain
(165)	Negative Exponential		N/A	SD	% Variance			Non-predictive
(171)	Log-linear	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	SE	R2			Predictive
(172)	Gompertz	Bigelow-type	Two-step Estimation, Missing Secondary Parameter(s)	SHW, CI	R2, Analysis of Residuals			Non-predictive
(173)	Gompertz	Bigelow-type	Two-step Estimation	SHW	R2			Non-predictive
(174)	Gompertz	Arrhenius	One-step Estimation	SHW	R2, Analysis of Residuals			Non-predictive
(175)	Log-linear		Linearized Approximation	SD	R2			Non-predictive
(179)	Log-linear	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	SDp	R2			Predictive
(180)	Linear		N/A	SE, CI	R2	Significance Testing		Predictive
(181)	Log-linear	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)		R2			Unqualified Predictive
(182)	Log-linear, Arrhenius	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)		R2			Non-predictive

Table B.1 (cont'd)

Ref. #	Primary Models	Secondary Models	Regression Features**	Error Estimates*	Goodness-of-fit Metrics	Model Selection	Complete	Applicability Domain
(183)	Log-linear	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	RE	R2			Predictive
(184)	Log-linear	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	SDp	R2			Predictive
(189)	Log-linear	Arrhenius	Two-step Estimation	SDp	R2			Non-predictive
(200)	Log-linear, Geeraerd	Root	Two-step Estimation	RMSE, SDp	R2, RMSE	R2, RMSE		Predictive
(202)	Log-linear, Weibull (Eqn. 3)		Incomplete	SD	R2			Unqualified Predictive
(204)	Log-linear	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	SDp				Non-predictive
(212)	Log-linear, Weibull (Eqn. 3)	Bigelow-type, Response Surface Model, Other	One-step Estimation	RMSE, CI, SE	RMSE	AIC, RMSE		Predictive
(214)	Log-linear	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	SE	R2			Non-predictive
(215)	Weibull (Eqn. 3)	Arrhenius, Tau Thermal Injury	One-step Estimation	RMSE		Validation		Non-predictive
(218)	Log-linear		N/A	CI	RMSE			Non-predictive

Table B.1 (cont'd)

Ref. #	Primary Models	Secondary Models	Regression Features**	Error Estimates*	Goodness-of-fit Metrics	Model Selection	Complete	Applicability Domain
(222)	Log-linear, Weibull (Eqn. 3)	Bigelow-type	One-step Estimation	RMSE, SE, RE	R2, RMSE, RE, AIC	R2, AIC		Non-predictive
(224)	Log-linear		Linearized Approximation	SD	R2			Non-predictive
(225)	Log-linear		N/A	SE	% Variance			Non-predictive
(226)	Log-linear		Linearized Approximation	SD				Non-predictive
(227)	Linear, Weibull (Eqn. 2)	Interpolation	Incomplete	RMSE	R2, RMSE	R2, RMSE		Unqualified Predictive
(228)	Log-linear	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	RMSE, SEM, CI	R2			Unqualified Predictive
(230)	Weibull (Eqn. 2)	Arrhenius, Tau Thermal Injury	One-step Estimation	RMSE, Residual Bias	RMSE	AIC, F-test		Predictive
(231)	Log-linear, Weibull (Eqn. 3)	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	RMSE, SE, CI	RMSE, Accuracy/ Bias Factors	RMSE, Accuracy/ Bias Factors		Non-predictive (Surrogate)
(238)	Modified Sigmoidal	Bigelow-type, Cerf, Arrhenius	Incomplete		R2	R2		Non-predictive
(244)	Log-linear		Linearized Approximation	SE				Non-predictive

Table B.1 (cont'd)

Ref. #	Primary Models	Secondary Models	Regression Features**	Error Estimates*	Goodness-of-fit Metrics	Model Selection	Complete	Applicability Domain
(246)	Response Surface Model		N/A	SE	R2, Significance Testing	Significance Testing		Predictive
(247)	Response Surface Model		N/A	SE	R2	Significance Testing		Non-predictive (Surrogate)
(249)	Response Surface Model		N/A	RMSE, SE	R2, RMSE	Significance Testing	X	Predictive
(251)	Log-linear, Weibull (Eqn. 3)	Bigelow-type, Response Surface Model, Log-logistic	Two-step Estimation		R2	R2, Significance Testing		Unqualified Predictive
(258)	Log-linear	Bigelow-type, Response Surface Model	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	RMSE, SD	R2, RMSE	AIC, RMSE		Predictive
(269)	Log-linear		Linearized Approximation	SD	R2			Non-predictive
(270)	Log-linear	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	SDp	R2			Non-predictive (Surrogate)
(271)	Log-linear	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	SDp	R2			Non-predictive
(125)	Log-linear	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	RMSE, CI	R2			Non-predictive
(274)	Log-linear	Bigelow-type	Linearized Approximation, Two-step Estimation, Missing Secondary Parameter(s)	SD	R2			Non-predictive

APPENDIX C. RAW DATA FROM MULTI-LABORATORY COMPARISON STUDY DESCRIBED IN CHAPTER 4

Table C.1 *Salmonella* isothermal inactivation results in oat flour collected by the control group

Batch Code	A	Batch Code	B	Batch Code	D	F	G
Batch a_w	0.450	Batch a_w	0.468	Batch a_w	0.464	0.453	0.462
Time (min)	Log CFU/g	Time (min)	Log CFU/g	Time (min)	Log CFU/g	Log CFU/g	Log CFU/g
0	7.19	0	7.96	0	8.04	7.34	8.37
0	7.41	0	7.76	0	8.06	7.42	8.35
0	6.99	0	7.83	0	8.14	8.37	7.80
10	6.16	7	7.25	7	7.02	6.92	7.66
10	6.35	7	7.65	7	7.10	7.29	7.16
10	5.94	7	7.28	7	6.91	7.50	7.15
20	6.06	17	7.06	22	5.97	6.65	6.09
20	6.25	17	6.97	22	5.72		6.10
20	6.65	17	7.12	22	6.40		6.68
30	5.67	27	6.41	37	4.56	6.76	7.22
30	5.67	27	6.47	37	4.78	6.84	6.17
30	5.98	27	6.34	37	5.47	6.49	6.78
40	4.80	37	6.86	52	3.43	5.73	7.02
40	5.39	37	6.53	52	4.32	5.85	6.11
40	5.72	37	6.08	52	4.07	5.19	6.66
50	5.54	47	6.45	67	2.73	4.94	5.94
50	5.07	47	6.29	67	4.26	5.08	5.69
50	5.91	47	6.51	67	3.37		5.98
		57	6.66	82	2.36	4.75	6.12
		57	6.11	82	2.53	5.44	4.80
		57	6.15	82	3.08	4.88	4.77
				97	3.12	5.16	5.80
				97	2.46	4.54	
				97	2.52	5.47	

Table C.2 *Salmonella* isothermal inactivation results in oat flour collected by IIT and FDA groups

Batch Code	IIT.A	Batch Code	IIT.B	Batch Code	IIT.F	Batch Code	FDA.A	Batch Code	FDA.B	Batch Code	FDA.H
Batch a_w	0.437	Batch a_w	0.437	Batch a_w	0.452	Batch a_w	0.454	Batch a_w	0.448	Batch a_w	0.474
Time (min)	Log CFU/g	Time (min)	Log CFU/g	Time (min)	Log CFU/g	Time (min)	Log CFU/g	Time (min)	Log CFU/g	Time (min)	Log CFU/g
0	7.70	0	8.05	0	7.54	0	7.93	0	7.90	0	8.33
0	8.00	0	7.99	0	7.99	0	7.59	0	8.27	0	8.41
0	7.58	0	8.05	0	8.14	0	7.42	0	7.73	0	8.15
7	7.26	7	7.67	7	7.30	7	6.87	7	7.61	7	7.42
7	7.32	7	7.39	7	7.30	7	7.21	7	7.10	7	7.31
7	7.13	7	7.13	7	7.51	7	6.67	7	7.09	7	7.15
17	6.62	17	6.99	22	6.96	22	6.29	17	6.63	22	6.60
17	6.68	17	6.60	22	6.92	22	6.59	17	6.74	22	6.60
17	7.05	17	6.71	22	6.61	37	5.50	17	7.02	22	6.62
27	5.75	37	6.64	37	6.08	37	6.22	27	7.07	37	5.81
27	6.87	37	6.36	37	6.72	37	6.12	27	6.78	37	5.82
27	6.61	37	6.23	37	5.48	52	5.95	37	6.22	37	5.86
37	5.99	57	6.06	52	6.05	52	6.40	37	6.47	52	6.21
37	6.09	57	5.80	52	6.12	52	6.18	37	7.28	52	6.39
37	7.06	57	5.88	52	5.68	67	5.27	47	6.17	52	6.11
47	5.63	77	5.31	67	5.68	67	5.18	47	6.45	67	5.16
47	5.72	77	5.65	67	6.28	67	5.89	47	6.15	67	5.93
47	5.82	77	5.99	82	5.11	82	5.44	57	5.31	82	5.50
57	6.03	97	5.90	82	4.67	82	5.45	57	6.33	82	5.15
57	5.79	97	4.98	82	5.91	82	4.91	57	5.87	82	4.83
57	5.11	97	5.46	97	4.99	97	5.23	67	6.07	97	4.55
				97	5.11	97	5.17	67	5.68	97	5.76
				97	4.65	97	5.27			97	4.65

Table C.3 *Salmonella* isothermal inactivation results in oat flour collected by two groups at MSU with different operators

Batch Code	MSU1.A	Batch Code	MSU1.B	MSU1.D	MSU1.F	Batch Code	MSU2.B	MSU2.C	MSU2.E	MSU2.H
Batch a_w	0.429	Batch a_w	0.466	0.465	0.455	Batch a_w	0.473	0.453	0.475	0.472
Time (min)	Log CFU/g	Time (min)	Log CFU/g	Log CFU/g	Log CFU/g	Time (min)	Log CFU/g	Log CFU/g	Log CFU/g	Log CFU/g
0	7.78	0	7.97	7.82	7.51	0	6.70	7.01	7.61	8.15
0	7.55	0	7.42	8.19	8.43	0	8.19	8.30	7.99	7.31
0	7.60	0	7.63	7.78	8.35	0	8.28	7.78	7.55	7.40
10	6.53	15	6.75	5.97	7.53	12	7.34	6.82	6.74	6.18
10	7.51	15	6.67	4.48	7.46	12	7.43	7.31	8.08	6.77
10	8.11	15	6.90		6.02	12	7.24		6.00	6.70
20	6.17	30	5.80	3.71	6.24	24	6.40	6.57	6.59	6.24
20	8.32	30	6.17	5.08	6.56	24	6.57	8.00	5.40	5.91
20	6.92	30	6.19	4.54	6.31	24	7.54	6.91	5.65	5.93
30	6.13	45	5.99	3.55	6.30	36	7.68	5.98	5.74	5.51
30	8.03	45	6.24	4.06	5.29	36	7.42	5.65	5.06	
30	6.13	45	5.82	4.79	6.06	36	5.85	6.91		
40	6.14	60	5.35	3.78	6.65	48	5.97	5.99	4.31	6.36
40	5.25	60	5.94	3.81	6.24	48	6.27	6.09		5.69
40	6.30	60	5.98	3.75	5.44	48	6.38	5.95		6.47
50	5.15	75	5.56	3.61	5.44	60	5.46	6.36	3.30	5.89
50	5.63	75	5.04	3.58		60	5.51	5.23	2.71	5.95
50	5.68	75	5.07			60		6.26		5.97
60	6.06	90	4.53	2.30	4.86	72	5.93	6.23	3.30	4.16
60	6.61	90	5.69	2.30	5.09	72	5.39	5.74	5.27	4.31
60	5.65	90			5.05	72			3.30	5.71
70	5.31	105	4.42	2.30	4.61	84	5.35	5.76	2.70	4.62
70	6.42	105	4.71	1.70	5.54	84	5.44	5.94	3.31	5.25
70	5.24	105	4.89		5.37	84		6.31	3.60	4.81

Table C.4 *Salmonella* isothermal inactivation results in oat flour collected by WSU and UNL

Batch Code	WSU.A	WSU.B	WSU.G	WSU.H	Batch Code	UNL.F	UNL.G	UNL.H
Batch a_w	0.422	0.455	0.452	0.450	Batch a_w	0.448	0.470	0.461
Time (min)	Log CFU/g	Log CFU/g	Log CFU/g	Log CFU/g	Time (min)	Log CFU/g	Log CFU/g	Log CFU/g
0	7.89	7.63	7.91	7.90	0	7.75	7.78	7.60
0	8.20	8.08	7.88	8.07	0	7.60	8.24	7.54
0	7.84	7.51	8.43	8.39	0	7.58	7.57	7.56
9	7.71	6.58	7.75	7.28	30	6.05	5.75	5.81
9	6.93	7.11	7.58	7.37	30	5.85	6.10	4.72
9	7.34	6.51	7.58	7.64	30	5.86	5.48	5.44
18	6.55	6.86	7.40	6.68	60	4.84	5.02	4.64
18	6.56	6.60	7.59	7.12	60	5.15	5.02	4.41
18	6.82	7.18	6.87	6.85	60	4.65	5.16	4.34
27	6.45	5.90	6.81	6.71	90	4.30	4.93	3.69
27	6.73	6.65	6.66	6.68	90	4.04	4.45	3.54
27	6.57	6.99	6.98	6.33	90	4.82	4.22	3.75
36	5.83	6.08	5.70	6.25	120	3.02	3.65	4.13
36	6.35	6.51	6.54	5.85	120	3.04	4.00	2.86
36	6.37	6.68	6.04	6.12	120	3.79	3.57	3.87
45	5.91	5.99	5.45	5.89	150	2.87	2.59	2.94
45	5.71	6.07	5.80	5.84	150	2.81	3.44	2.75
45	5.79	6.24	5.05	6.05	150	2.24	2.68	2.59
					180	2.43	2.75	2.58
					180	1.90	2.92	2.76
					180	2.71	2.69	2.59

APPENDIX D. REVIEW OF RECENT SURROGATE-PATHOGEN COMPARISON STUDIES WITH A FOCUS ON HOW THE SURROGATE-PATHOGEN RELATIONSHIP WAS EVALUATED (CHAPTER 6)

This material is supplemental to the *Current Opinion in Food Science Review* article entitled "Improving the Utility of Surrogates Intended for Foodborne Pathogen Preventive Control Validations" by Ian Hildebrandt and Bradley Marks at Michigan State University.

Studies were located using Web of Science (Clarivate, PA, USA) database by searching combinations of "pathogen", "surrogate", and the names of common pathogens/surrogates (e.g., *Salmonella* surrogate, *Listeria* surrogate, *E. coli* surrogate, *Listeria innocua*, *E. faecium*, *E. faecium* pathogen) in the topic field. Studies were selected that evaluated surrogate and pathogen reduction resistance for the purpose of qualifying a surrogate as appropriate. Studies that did not include a direct pathogen/surrogate comparison were excluded. Pathogens and surrogates are listed in the table as they were listed in the source materials. Reduction ratios listed were sourced from the source materials wherever possible; otherwise, the authors provided estimates of reduction ratios.

Table D.1 Summary of surrogate-pathogen study context, surrogate evaluation methodologies, and relative results

Ref.	Pathogen(s)	Surrogate(s)*	Subject(s)	Inter-vention(s)	Stated Desirable Red. Char.**	Surrogate-Pathogen Comparison Methods	Red. Ratio ***
(3)	Cocktail of *E. coli* O121:H19 FNW19M81, O157:H7 F4546; cocktail of *L. monocytogenes* 1/2a FSL R2-499, 1/2b FSL R2-502, 4b ScottA; cocktail of *S. enterica* Montevideo 1449, Newport, Tennessee K4643	*Pediococcus acidilactici* ATCC 8042	Raisins, apricots, and macadamia nuts	Thermal	>R, ≈R, ≯>R	Visual comparison of model parameter(s), Statistical comparison of population reductions after treatment(s), Statistical evaluation of reduction ratio(s)	1.5 - 3
(4)	Cocktail of *S. enterica* Agona 447967, Montevideo 488275, Mbandaka 698538, Tennessee K4643, Reading Moff 180418	*E. faecium* NRRL B-2354	Nonfat milk powder, peanut butter, wheat flour, ground black pepper, almond meal, date paste	Thermal	>R, ≯>R	Statistical comparison of model parameter(s)	0.9 - 4
(5)	Cocktail of *S. enterica* Agona 447967, Montevideo 488275, Mbandaka 698538, Tennessee K4643, Reading Moff 180418	*E. faecium* NRRL B-2354	Dairy powders	Thermal	>R, =R	Statistical comparison of model parameter(s)	1-2
(15)	Cocktail of *S. enterica* Enteritidis PT 30, Tennessee K4643, Agona 447967	*E. faecium* NRRL B-2354	Honey and high fructose corn syrup	Persistence	≈R	Visual comparison of population reductions after treatment(s)	1 -1.5

Table D.1 (cont'd)

Ref.	Pathogen(s)	Surrogate(s)*	Subject(s)	Inter-vention(s)	Stated Desirable Red. Char.**	Surrogate-Pathogen Comparison Methods	Red. Ratio ***
(21)	Cocktail of *S. enterica* Choleraesuis ATCC 13312, Enteritidis ATCC 4931, Newport ATCC 6962, Typhimurium ATCC 700720, Rissen SAL 4599, Montevideo SAL 1449, cocktail of *S. enterica* Rissen SAL 4599, Montevideo SAL 1449, cocktail of *S. enterica* Choleraesuis ATCC 13312, Enteritidis ATCC 4931, Newport ATCC 6962, Typhimurium ATCC 700720	*E. faecium* NRRL B-2354	Onion powder, dried oregano, cumin seeds, peppercorns, and talc	Gamma radiation	>R, ≈R	Statistical comparison of model parameter(s), fit model to reduction ratio(s)	1-1.3
(24)	Hepatitis A virus HM-175 ATCC VR-1402, cocktail of *E. coli* O157:H7 BRMSID 188, O157:H7 LFMFP 846, O157:H7 ATCC 700728; cocktail of *S. enterica* Thompson RM1987, Typhimurium SL1344, Typhimurium LFMFP 883; cocktail of L. monocytogenes LMG 23192, LMG 23194, LMG 26484	Murine norovirus S99; (Bacteriophage MS2); *E. faecium* ATCC 8459; *E.coli* P1 ATCC BAA-1427; *Listeria innocua* ATCC 33090	Blueberries	Osmotic dehydration	≈R	Statistical comparison of population reductions after treatment(s)	0.9 - 1.5
(25)	Cocktail of *S. enterica* Agona 447967, Tennessee K4643, Enteritidis PT 30; cocktail of *L. monocytogenes* NRRL B-57618, NRRL B-33053, NRRL B-33466	*E. faecium* NRRL B-2354, cocktail of *L. innocua* NRRL B-33197, *L. innocua* TVS470, *L. innocua* TVS 471	Cocoa powder	Thermal	>R, ≈R	Statistical comparison of model parameter(s), Statistical comparison of population reductions after treatment(s)	1

Table D.1 (cont'd)

Ref.	Pathogen(s)	Surrogate(s)*	Subject(s)	Inter-vention(s)	Stated Desirable Red. Char.**	Surrogate-Pathogen Comparison Methods	Red. Ratio ***
(36)	Individual and cocktail formulations of *S. enterica* Seftenberg 775W ATCC 43845, Enteritidis PT-30 ATCC BAA-1045, Tennessee K4643	*E. faecium* ATTC 8459; (*E. faecalis* ATCC 29212)	Peanuts, pecans	Thermal	>R	Statistical comparison of population reductions after treatment(s)	0.9 - 2
(50)	Cocktail of *S. enterica* Agona 447967, Montevideo 488275, Mbandaka 698538, Tennessee K4643, Reading Moff 180418	*E. faecium* NRRL B-2354	Inshell hazelnuts	Thermal	>R, ≈R	Statistical comparison of model parameter(s)	5-10
(51)	Cocktail of *S. enterica* Agona 447967, Montevideo 488275, Mbandaka 698538, Tennessee K4643, Reading Moff 180418	*E. faecium* NRRL B-2354	Cumin seeds	Gaseous ethylene oxide	>R, ≈R	Visual comparison of model parameter(s)	1 - 1.5
(52)	*S. enterica* Typhimurium IB43, cocktail of *S. enterica* Enteritidis ATCC 13076, Newport ATCC 6962, Typhimurium ATCC 14028	*E. faecium* NRRL B-2354	Spinach	Persistence	>R, =R	Statistical comparison of population reductions after treatment(s)	1
(56)	*S. enterica* Typhimurium ATCC 14028	*E. faecium* ATCC 8459	Mash broiler feed	Thermal	>R, ≈R	Statistical comparison of model parameter(s)	2
(60)	Cocktail of *S. enterica* Typhimurium, Agona, Enteritidis, Montevideo, Tennessee	*E. faecium* NRRL B-2354	Milk powders	Thermal	>R, ≈R, ⊁>>R	Visual comparison of model parameter(s)	1.5 - 2.5

Table D.1 (cont'd)

Ref.	Pathogen(s)	Surrogate(s)*	Subject(s)	Inter-vention(s)	Stated Desirable Red. Char.**	Surrogate-Pathogen Comparison Methods	Red. Ratio ***
(63)	*E. coli* O157:H7 M-11-0450i-1, O145:H- 97–3192, O157:H7 DMB EC1738, O26:H11 97–3250, O157:H7 ATCC 35150, O104:H4 ATCC BAA-2326, O104:H4 WS04003A, O111:H-95JB1, O157:H7 ATCC 43894, O103:H2 P030, O45:H- P034, O157:H7 Sakai, O157:H7 P092, O104:H4 P108, O157:H7 P091, O157:H7 ENT C9490, O157:H7 SEA13B88, O157:H7 ATCC 43895, O157:H7 09 PF 736, O104:H4 2011C-3493, O111:H- 95NR1, O111:H8 B99BE001203	(*E. coli* K12 LMM 1010), *E. coli* P1 BAA-1427; (*Listeria innocua* ATCC 33090); (*Lactobacillus plantarum* ATCC 8014); (*Bifidobacterium lactis* Nestec PP011); *E. faecium* NRRL B-2354	Wheat flour	Persistence, thermal	>R	Visual comparison of population reductions after treatment(s)	1-2
(64)	Cocktail of *S. enterica* Enteritidis PT 30, Tennessee K4643, Agona 447967; cocktail of *L. monocytogenes* 1/2a NRRL B-57618, 4b NRRL B-33053, 1/2b NRRL B-33466	*E. faecium* NRRL B-2354	Shredded coconut	Thermal	>R, ≈R	Statistical comparison of model parameter(s)	1.1 - 2.0
(69)	*S. enterica* Typhimurium LT2; cocktail of shiga-toxin forming *E. coli* consisting of O26, O111, O91, O145, O157 serogroups; cocktail of *Campylobacter jejuni* ATCC 29428, ATCC 33291, ATCC BAA-374	Cocktail of *E. coli* ATCC BAA 1427, ATCC BAA 1428, ATCC 1430	Pork	Antimicro-bial solution spray	≈R	Visual comparison of population reductions after treatment(s)	N/A

Table D.1 (cont'd)

Ref.	Pathogen(s)	Surrogate(s)*	Subject(s)	Inter-vention(s)	Stated Desirable Red. Char.**	Surrogate-Pathogen Comparison Methods	Red. Ratio ***
(96)	Cocktail of *S. enterica* Agona 447967, Montevideo 488275, Mbandaka 698538, Tennessee K4643, Reading Moff 180418	*E. faecium* NRRL B-2354	Apple	Thermal	>R, ≈R	Statistical comparison of population reductions after treatment(s)	1-2
(100)	Cocktail of *S. enterica* Typhimurium ATCC 14028, Seftenberg ATCC 43845, Enteritidis ATCC BAA-708, Newport ATCC 6962, Montevideo ATCC BAA-710	*E. faecium* ATTC 8459	Cornmeal	Thermal	>R, ≈R	Statistical comparison of model parameter(s)	1 - 4
(101)	Cocktail of *E. coli* O157:H7 ATCC 700599, O157:H7 ATCC 35150, O26:H11 SJ2, O45:H2 05-6545, O103:H11 SJ12, cocktail of *S. enterica* Typhimurium ATCC 14028, Heidelberg ATCC 45955, Enteritidis PT 30, Montevideo 51, Newport H1073	(Cocktail of E. coli O157 ATCC 700728, O157 ATCC 43888, O45:H10 ATCC BAA-2649, OR:H48:K ATCC 700926, O6 ATCC 25922); cocktail of attenuated *S. enterica* Typhimurium ATCC 53647, Typhimurium ATCC 54648, Typhimurium χ4096, Typhimurium χ3985, Typhimurium χ8098	Carrots	Gaseous chlorine dioxide	>R, ≈R	Visual comparison of model parameter(s)	0.9 - 1.3
(105)	Cocktail of *S. enterica* Enteritidis 15-SA02843, Gaminara 05-1527, Oranienburg 17-SA01525, Rubislaw 07-01143, Typhimurium 10-01906; cocktail of *E. coli* DSM 19206, DSM 5923	(*Geobacillus stearothermophilus* ATCC 7953); (*Deinococcus radiodurans* DSM 20539); (*E. coli* DSM 19206), (*E. coli* DSM 5923), *E. coli* DSM 18039; (*E. faecium* NRRL B-2354)	Pumpkin and flax seeds	Electron beam	>R, ≈R, ≯>R	Statistical comparison of model parameter(s), Statistical comparison of population reductions after treatment(s)	1

Table D.1 (cont'd)

Ref.	Pathogen(s)	Surrogate(s)*	Subject(s)	Inter-vention(s)	Stated Desirable Red. Char.**	Surrogate-Pathogen Comparison Methods	Red. Ratio ***
(124)	Cocktail of *S. enterica* Typhimurium ATCC 14028, Tennessee ATC 10722	*E. faecium* ATCC 8459	Chicken patties	Thermal	>R, ≈R	Statistical comparison of model parameter(s)	1-1.25
(133)	*S. enterica* Enteritidis PT 30 ATCC BAA-1045, Typhimurium PT 42	*(Enterobacter aerogenes* B199A), *Pantoea dispersa* JFS	Whole wheat flour	Thermal	>R	Visual comparison of model parameter(s)	1
(134)	Data from previous studies using cocktails of *S. enterica*, *E. coli* O157:H7, *L. monocytogenes*, and *Staphylococcus aureus*	*(E. faecium* 201224-016); (*Pediococcus acidilactici* Saga200); (cocktail of *Pediococcus acidilactici* ATCC 8042, PO2K5, ATCC 43200, FBB61-2); (cocktail of *E. coli* ATCC BAA-1427, ATCC BAA-1428, ATCC BAA-1429, ATCC BAA-1430); (cocktail of *Latilactobacillus sakei* GO-R2C, GO-R2D); cocktail of *Carnobacterium divergens* GO-R2E-B, GO-R1B; *Carnobacterium gallinarum* NB-R2A, NB-R2B	Beef	Curing	≈R	Visual comparison of population reductions after treatment(s)	1
(135)	*S. enterica* Newport 96E01152C-TX	*Enterobacter aerogenes* B199A	Cucumbers	Transfer	>R, ≈R	Statistical comparison of population reductions after treatment(s)	1

Table D.1 (cont'd)

Ref.	Pathogen(s)	Surrogate(s)*	Subject(s)	Inter-vention(s)	Stated Desirable Red. Char.**	Surrogate-Pathogen Comparison Methods	Red. Ratio ***
(144)	Cocktail of *S. enterica* Agona 447967, Montevideo 488275, Mbandaka 698538, Tennessee K4643, Reading Moff 180418	*E. faecium* NRRL B-2354	Chia seeds	Persistence, thermal	>R	Statistical comparison of model parameter(s)	1.6
(147)	Cocktail of *S. enterica* Typhimurium ATCC 14028, Tennessee ATCC 10722; *L. monocytogenes* L2624, L2625	*E. faecium* ATCC 8459	Cucumbers and tomatoes	Washing	>R, =R	Statistical comparison of population reductions after treatment(s)	0.6 - 2
(155)	Cocktail of *S. enterica* Enteritidis PT 30, Tennessee LJH 1244, Montevideo LJH 1245, Saintpaul LJH 1375, Gaminara LJH 1220	*E. faecium* ATCC 8459	Peaches	Various drying pre-treatments, heat-assisted dehydration, sun dehydration	≈R, ⊁>R	Statistical comparison of population reductions after treatment(s)	0.9 - 4
(177)	Individual and/or cocktail formulations of *S. enterica* Enteritidis PT 30, Enteritidis PT 9c, Tennessee, Montevideo, Saintpaul, Senftenberg; *E. coli* O157:H7 Odwalla strain 223, O157:H7 CDC 658, O157:H7 EC4042, O157:H7 EC1738, O157:H7 NML 11-1865; *L. monocytogenes* 4b LCDC81-861, 4b Micro FD 4-4-0-4-4, LIS0234, LIS0133, PTVS 308	*E. faecium* NRRL B-2354	Pistachios	Thermal	>R, =R	Statistical comparison of model parameter(s), Statistical comparison of population reductions after treatment(s), statistical evaluation of reduction ratio(s)	0.9 - 2

Table D.1 (cont'd)

Ref.	Pathogen(s)	Surrogate(s)*	Subject(s)	Inter-vention(s)	Stated Desirable Red. Char.**	Surrogate-Pathogen Comparison Methods	Red. Ratio ***
(178)	*S. enterica* Rissen	*E. faecium* NRRL B-2354	Black peppercorns	Electron beam	None	Statistical evaluation of reduction ratio(s)	5
(187)	Cocktail of *S. enterica* Montevideo, Tennessee K4643, Ball ARL-SE-085, Johannesburg ARL-SE-013	*E. faecium* ATCC 8459	Black peppercorns and cumin seeds	Thermal	>R, ≈R	Statistical comparison of population reductions after treatment(s), visual evaluation of reduction ratio(s)	1-1.1
(193)	Cocktail of *S. enterica* Typhimurium, Agona, Montevideo, Tennessee	*E. faecium* NRRL B-2354	Paprika, white pepper, cumin powder	Thermal	>R	Statistical comparison of model parameter(s), visual comparison of population reductions after treatment(s)	1.5 - 2.5
(196)	*S. enterica* Enteritidis PT 30	*E. faecium* NRRL B-2354	Egg powders	Thermal	>R	Statistical comparison of model parameter(s)	1.5 - 2.0

Table D.1 (cont'd)

Ref.	Pathogen(s)	Surrogate(s)*	Subject(s)	Inter-vention(s)	Stated Desirable Red. Char.**	Surrogate-Pathogen Comparison Methods	Red. Ratio ***
(201)	Cocktail of *L. monocytogenes* 1/2b FSL J1-177, 1/2a FSL C1-056, 4b FSL N3-013, 1/2a FSL R2-499, 4b FLS N1-227, 4b/Scott A ATCC 49594; cocktail of *S. enterica* Montevideo, Agona, Tennessee, Weltevreden ATCC BAA-2568, Senftenberg ATCC H385, Typhimurium PT 42	*E. faecium* NRRL B-2354	Peanut butter, powder infant formula, and wheat flour	Thermal	>R	Statistical comparison of model parameter(s)	1.25 - 2.5
(203)	Cocktail of *S. enterica* Typhimurium ATCC 14028, Heidelberg ATCC 45955, Enteritidis PT 30, Montevideo 51, Newport H1073	*E. faecium* NRRL B-2354	Almonds	Gaseous chlorine dioxide	=R	Statistical comparison of population reductions after treatment(s)	1
(208)	*S. enterica* Typhimurium ATCC 13311, Typhimurium ATCC 13311(pLHR), Typhimurium ATCC 13311(pRK767), Senftenberg ATCC 43845, FUA1934, FUA1946, FUA1955; cocktail of *E. coli* K-12 MG1655 ATCC 700926, 03-2832 O121:H19, 05-6544 O26:H11, C0283 O157:H7, PARC 449 O145:NM, 03-6430 O145:NM	*E. faecium* NRRL B-2354; *Pediococcus acidilactici* FUA3072; *Staphylococcus carnosus* FUA2133; *E. coli* AW1.7, *E. coli* AW1.7 ΔpHR1 (suitability not reported)	Beef jerky	High pressure carbon dioxide	>R, =R	Statistical comparison of population reductions after treatment(s)	1-5
(209)	Cocktail of *S. enterica* Agona 447867, Mbandaka 698538, Typhimurium DO 5, Tennessee K4643, Enteritidis PT 30, Montevideo 488275	Avirulent *S. enterica* Typhimurium MHM112	Egg	Washing with plasma activated water	≈R	Statistical comparison of population reductions after treatment(s)	1

Table D.1 (cont'd)

Ref.	Pathogen(s)	Surrogate(s)*	Subject(s)	Inter-vention(s)	Stated Desirable Red. Char.**	Surrogate-Pathogen Comparison Methods	Red. Ratio ***
(211)	Cocktail of *L. monocytogenes* NRRL B-57618, NRRL B-33053, NRRL 33466	*(Lactobacillus plantarum* NRRL B-531); (*Pediococcus acidilactici* NRRL B-1116); *E. faecium* NRRL B-2354; cocktail of *Listeria innocua* NRRL 33197, NRRL B-33314, NRRL B-33554	Apples	Chemical sanitizers	>R, ≈R	Statistical comparison of population reductions after treatment(s)	1-2
(216)	Cocktails consisting of three variants for each of *S. enterica* Senftenberg, Mbandaka, Montevideo, Agona, Tennessee. Surrogate comparison used *S. enterica* Agona RA 1052	*E. faecium* ATCC 8459	Mash broiler feed	Thermal	>R, =R	Statistical comparison of model parameter(s)	1 - 1.4
(217)	*S. enterica* Enteritidis PT 30	*E. faecium* NRRL B-2354	Soy protein	Thermal	>R	Statistical comparison of population reductions after treatment(s)	1 - 1.5
(221)	*E. coli* ATCC 25922; cocktail of *S. enterica* Typhimurium ATCC 700728, Enteritidis, Newport, Montevideo	*E. faecium* ATCC 8459 (suitability not reported)	Spinach	High voltage atmospheric cold plasma	≈R	Visual comparison of model parameter(s), Statistical comparison of population reductions after treatment(s)	2
(229)	*S. enterica* Typhimurium ATCC 14028	*E. faecium* ATCC 8459	Butternut squash	Antimicrobial wash	>R, =R	Statistical comparison of population reductions after treatment(s)	1

Table D.1 (cont'd)

Ref.	Pathogen(s)	Surrogate(s)*	Subject(s)	Inter-vention(s)	Stated Desirable Red. Char.**	Surrogate-Pathogen Comparison Methods	Red. Ratio ***
(231)	Cocktail of *S. enterica* Agona 447967, Tennessee K4643, Enteritidis PT 30	*E. faecium* NRRL B-2354	Cocoa powder	Persistence, thermal	>R	Visual comparison of model parameter(s)	0.7 - 1.4
(245)	Cocktail of *S. enterica* Agona 447967, Montevideo 488275, Mbandaka 698538, Tennessee K4643, Reading Moff 180418	*E. faecium* NRRL B-2354	Dried basil leaves	Thermal	>R, ≈R	Statistical comparison of population reductions after treatment(s)	1.8
(247)	Cocktail of *S. enterica* Agona 447967, Montevideo 488275, Mbandaka 698538, Tennessee K4643, Enteritidis PT 30	*E. faecium* NRRL B-2354	Oat flour	Thermal	>R, =R	Statistical comparison of model parameter(s), statistical evaluation of reduction ratio(s)	2.5
(248)	Cocktail of *S. enterica* Agona 447967, Montevideo 488275, Mbandaka 698538, Tennessee K4643, Reading Moff 180418	*E. faecium* NRRL B-2354	Dried basil leaves	Gaseous chlorine dioxide	>R, =R	Statistical comparison of model parameter(s)	1.2 - 1.9
(253)	*S. enterica* Seftenberg 775W ATCC 43845	*E. faecium* NRRL B-2354	Poultry litter	Thermal	>R	Statistical comparison of population reductions after treatment(s)	>2
(255)	Cocktail of *S. enterica* Agona 447967, Montevideo 488275, Mbandaka 698538, Tennessee K4643, Reading Moff 180418	*E. faecium* NRRL B-2354	Chia seeds	Gaseous chlorine dioxide	>R	Visual comparison of model parameter(s)	0.9 - 1.5

Table D.1 (cont'd)

Ref.	Pathogen(s)	Surrogate(s)*	Subject(s)	Inter-vention(s)	Stated Desirable Red. Char.**	Surrogate-Pathogen Comparison Methods	Red. Ratio ***
(256)	Cocktail of *S. enterica* Agona 447967, Montevideo 488275, Mbandaka 698538, Tennessee K4643, Reading Moff 180418	*E. faecium* NRRL B-2354	Dried basil leaves	Thermal	>R, ≈R	Statistical comparison of population reductions after treatment(s)	2
(257)	Cocktail of *S. enterica* Agona 447967, Montevideo 488275, Mbandaka 698538, Tennessee K4643, Reading Moff 180418	*E. faecium* NRRL B-2354	Ground black pepper	Thermal	>R, ≈R	Statistical comparison of model parameter(s)	1.5 - 10
(259)	Cocktail of *S. enterica* Agona 447967, Montevideo 488275, Mbandaka 698538, Tennessee K4643, Reading Moff 180418	*E. faecium* NRRL B-2354	Milk powders	Persistence, thermal	>R	Statistical comparison of model parameter(s)	2
(260)	Cocktail of *S. enterica* Agona 447967, Montevideo 488275, Mbandaka 698538, Tennessee K4643, Reading Moff 180418	*E. faecium* NRRL B-2354	Black peppercorn	Gaseous ethylene oxide	>R, ≈R	Visual comparison of model parameter(s), visual evaluation of reduction ratio(s)	1 - 1.25
(261)	Cocktail of *S. enterica* Agona 447967, Montevideo 488275, Mbandaka 698538, Tennessee K4643, Reading Moff 180418	*E. faecium* NRRL B-2354	Egg white powder	Thermal	>R, =R, ≈R	Statistical comparison of population reductions after treatment(s)	1-1.8
(262)	Cocktail of *S. enterica* Agona 447967, Montevideo 488275, Mbandaka 698538, Tennessee K4643, Reading Moff 180418	*E. faecium* NRRL B-2354	Ground black pepper	Thermal	>R, ≈R	Statistical comparison of model parameter(s)	1.5 - 2

Table D.1 (cont'd)

Ref.	Pathogen(s)	Surrogate(s)*	Subject(s)	Inter-vention(s)	Stated Desirable Red. Char.**	Surrogate-Pathogen Comparison Methods	Red. Ratio ***
(263)	Cocktail of *S. enterica* Agona 447967, Montevideo 488275, Mbandaka 698538, Tennessee K4643, Reading Moff 180418	*E. faecium* NRRL B-2354	Black peppercorns and cumin seeds	Gaseous chlorine dioxide	>R	Visual evaluation of reduction ratio(s)	1-1.5
(264)	Cocktail of *S. enterica* Agona 447967, Montevideo 488275, Mbandaka 698538, Tennessee K4643, Reading Moff 180418	*E. faecium* NRRL B-2354	Black peppercorns, cumin seeds, dried basil leaves	Gaseous chlorine dioxide, persistence	>R, =R	Statistical comparison of population reductions after treatment(s)	1 - 1.2
(267)	Cocktail of *S. enterica* Choleraesuis ATCC 13312, Enteritidis ATCC 4931, Newport ATCC 6962, Typhimurium ATCC 700720, Heidelberg ATCC 8326, cocktail of *S. enterica* I 4,[5],12:i:ISU-SAL0233-19, I 4,[5],12:i:ISU-SAL0234-19, I 4,[5],12:i:ISU-SAL0235-19, I 4,[5],12:i:ISU-SAL0236-19, I 4,[5],12:i:ISU-SAL0237-19	Cocktail of *E. coli* BAA-1427, BAA-1428, BAA-1429, BAA-1430, BAA-1431	Pork	Washing	None	Statistical comparison of population reductions after treatment(s)	1

Table D.1 (cont'd)

Ref.	Pathogen(s)	Surrogate(s)*	Subject(s)	Inter-vention(s)	Stated Desirable Red. Char.**	Surrogate-Pathogen Comparison Methods	Red. Ratio ***
(268)	Cocktail of *S. enterica* Choleraesuis ATCC 13312, Enteritidis ATCC 4931, Newport ATCC 6962, Typhimurium ATCC 700720, Typhimurium DT104; cocktail of *E. coli* O157 FSRL-Y24, O157 FSRL-Y25, O157 ATCC 43895, O157 ATCC 43894, O157 ATCC 35150, cocktail of *E. coli* O26:H11 ATCC BAA-2196, O45:H2 ATCC BAA-2193, O103:H11 ATCC BAA-2215, O111 ATCC BAA-2440, O121:H19 ATCC BAA-2219, O145 ATCC BAA-2192	Cocktail of *E. coli* ATCC BAA 1427, ATCC 1428, ATCC 1429, ATCC 1430, ATCC 1431	Beef	High hydrostatic pressure	>R	Statistical comparison of population reductions after treatment(s), fit model to reduction ratio(s)	1-1.2
(273)	*S. enterica* Senftenberg 775W ATCC 43845, Enteritidis PT 30 ATCC BAA-1045, Thompson RM1987; *L. monocytogenes* 4b; *E. coli* O157:H7 ATCC 700728	*E. coli* P1 ATCC BAA-1427; *Listeria innocua* ATCC 33090; (*E. faecium ATCC 8459)*	Basil leaves	Thermal	>R, ≈R, ⊁>R	Statistical comparison of population reductions after treatment(s)	1
(275)	Cocktail of *S. enterica* Enteritidis PT 30, Tennessee K4643, Agona 447967	*E. faecium* NRRL B-2354	Almond meal	Thermal	>R	Statistical comparison of model parameter(s)	1.3 - 1.8

* Surrogates in parentheses were reported as not appropriate by the original authors against the compared pathogen

** Stated desirable reduction characteristics key:

"=R" appropriate surrogates may have equivalent resistance to the mode of reduction as the pathogen

"≈R" appropriate surrogates may have similar resistance to the mode of reduction as the pathogen

">R" appropriate surrogates may have greater resistance to the mode of reduction than the pathogen

"⊁>R" appropriate surrogates may not have substantially greater resistance to the mode of reduction than the pathogen

*** Reduction ratios were estimated between surrogates and their pathogen counterparts with a focus at ranges of pathogen reduction > 4 log