CPS CENTER for PRODUCE SAFETY

CPS 2018 RFP
FINAL PROJECT REPORT

Project Title
Exploring the relationship between product testing and risk

Project Period
January 1, 2019 – December 31, 2019

Principal Investigator
Emma Hartnett
Risk Sciences International
55 Metcalfe St., Suite 700
Ottawa, ON K1P 6L5
Canada
T: 613-260-1424 x152
E: ehartnett@risksciences.com

Co-Principal Investigator
Donald W. Schaffner
Freehold, NJ 07728
E: don.schaffner@rutgers.edu

Objectives

1. *Develop a sampling-risk model that quantifies the relationship between product testing, lot rejection rates, and risk (which is directly related to prevalence and concentration in products post testing).*
2. *Provide detailed, fully documented, analyses of the relationship between product sampling variables driving the risk.*
3. *Support risk reduction initiatives through analyses that explicitly enable the exploration of risk management options, facilitating selection of actionable sampling strategies that have the biggest impact on risk reduction.*

Funding for this project provided by the Center for Produce Safety through:
CPS Campaign for Research

FINAL REPORT

Abstract

Microbiological sampling is one tool available to help verify product safety, but where and how to sample, in terms of the most effective risk reduction and efficient use of resources (money, labor, time), is not always clear. The focus of risk management strategies is to prevent contamination and subsequent growth if contamination has occurred. There are also many pressures that have driven industry toward product testing through sampling. To effectively employ sampling as a risk-management strategy, it is essential to have a clear understanding of the (sometimes complex) relationship between sampling activities in the supply chain and the residual consumer risk. Such an understanding can be supported by the quantitative definition of the relationship (i.e. a model) that is employed to explore how the components of the relationship interact to influence the risk. To meet this need, we have developed a sampling-risk model that quantifies the relationship between product testing, lot rejection rates, and risk, and performed detailed analyses of the relationship between product sampling variables driving the risk and location options for sampling.

Background

The produce industry currently faces a challenge in that there are many potential routes of contamination of produce with microbial pathogens. There is increasing pressure from regulatory bodies and product purchasers to adopt risk-based management practices. Microbiological sampling is one tool available to help ensure product safety, but where to sample, in terms of the most effective risk reduction and efficient use of resources (money, labor, time), is not always clear. End product testing is one option, but is not a preventative measure, so sampling at other points in the field-to-fork continuum may be more desirable in terms of supporting actionable risk reduction strategies.

Product sampling can be used either as a tool for verifying that other control measures have been applied effectively or as a means of directly identifying unacceptable lots, i.e. not complying with an established limit, and thus preventing their release into domestic or international trade (CAC, 2013; JEMRA, 2006, 2016). The development of sampling plans is a critical part of the decision-making process of ensuring safety while paying appropriate attention to practicality and cost. Choosing a sampling plan requires careful attention to the underlying statistical properties of microbial contamination, and the statistical properties of the sampling and analytical processes themselves, for example the presence-absence or number of micro-organisms per unit as indicated by a sampling plan. Foods deemed unacceptable on this basis may be unmarketable, and so producers need to understand the relationship between the level of microbial contamination in the product, the stringency of the plan applied, and the likelihood of lot rejection that results from any given sampling plan that may be employed. The traditional application of microbiological criteria is to individual product lots, with the goal of eliminating those lots that have an unacceptably high level of contamination.

Tools have been developed that link sampling plan attributes to concentration post-sampling (for example ICMSF, 2016; JEMRA 2011), but they either do not directly link to residual risk (ICMSF tool) or do not allow analysis of the relationship of all of the variables on the risk (JEMRA tool). Our analysis will help fill the gap between defining sampling activities and understanding the impact of sampling in the produce supply chain on consumer risk.

Research Methods and Results

There are 3 components to this report:

1) Describing the relationships between sampling plan components and sampling plan performance in terms of the **probability of rejection**
2) Describing the relationships between sampling plan components and sampling plan performance in terms of **consumer risk and risk reduction**
3) Exploring the impact of the choice of **location** of sampling on **consumer risk reduction**

Describing the Relationships Between Sampling Plan Components

Sampling plans for pathogens on produce are generally adopting a presence-absence (also referred to as two-class) based sampling approach. In this approach a number of samples (n) of a specified mass (m) are tested for the presence of the organism. If the organism is found the action depends upon the number of acceptable positives (c) that will be tolerated. If this number is exceeded the lot will be rejected and removed from the intended product flow (for example by being destroyed or repurposed in some way). Each of the components n, m, and c combine to determine the performance of a given sampling plan in terms of its ability to detect contamination of a lot at specific levels, described in terms of the concentration of organism in the lot (often described using the units Log[1] CFU/g). Conversion of Log CFU/g to CFU/g is provided in Table 1.

For our analyses we have developed a sampling model that calculates the probability of rejection[2] based upon a positive test result (and the complement to this: the probability of acceptance), for a tested lot. We adjust the value of the components and present the resulting probability of rejection of a lot for a range of concentration values.

Number of Samples. The number of samples taken and tested positively influences the probability of rejection. The relationship between probability of rejection and the number of samples is shown in Figure 1. In this figure 25g samples are taken, and no positives are accepted (c=0). Each line shows the probability of rejection for a specific number of samples used (see the color code in the graph legend on Figure 1). As the mean concentration increases, the probability of rejection increases for any given number of samples (as seen by following a single line on the graph). As the number of samples increases for a given concentration, the probability of rejection increases; this is shown by the difference between the lines in Figure 1.

Sample Mass. A similar relationship is seen with increasing sample size/mass and is shown in Figure 2. Sample size and sample mass combine to give the total mass examined which increases the likelihood that at least one organism will be present in the sample taken from the lot, thus increasing the probability of rejection.

[1] Base 10 is utilized throughout these analyses, that is Log_{10} CFU/g.

[2] Analyses assume a lognormal distribution of concentration with a standard deviation between lots of 0.8, and a within lot standard deviation of 0.5. Probability of rejection follows standard calculations as described by the Joint FAO/WHO Expert Meetings on Microbiological Risk Assessment (JEMRA). For example JEMRA (2006): FAO/WHO Expert meeting on *Enterobacter sakazakii* and *Salmonella* in PIF, Chapter 4.3 Microbiological Criteria; JEMRA (2011) Microbiological Sampling Plan Analysis Tool. http://www.fstools.org/sampling/

Acceptable Positives. For any given sampling plan, increasing the number of acceptable positives decreases the probability of rejection. Often when sampling for pathogens the plan implemented uses zero aceptable positives (c=0), this is the most stringent option. To show the influence of the number of acceptable positives, values in the range from 0 to 5 were analysed and the results are can be seen in Figure 3. Zero acceptable positives has the highest probabilty of rejection for all mean concentrations shown, for example, at a mean concentration of -2 log CFU/g the probabilty of rejection for the difference optipons for acceptable positives in show in is Table 2.

Table 1: Log CFU/g to CFU/g conversions.

Log CFU/g	CFU/g
2	100
1	10
0	1
-1	0.1 (1 in 10g)
-2	0.01 (1 in 100g)
-3	0.001 (1 in 1000g)

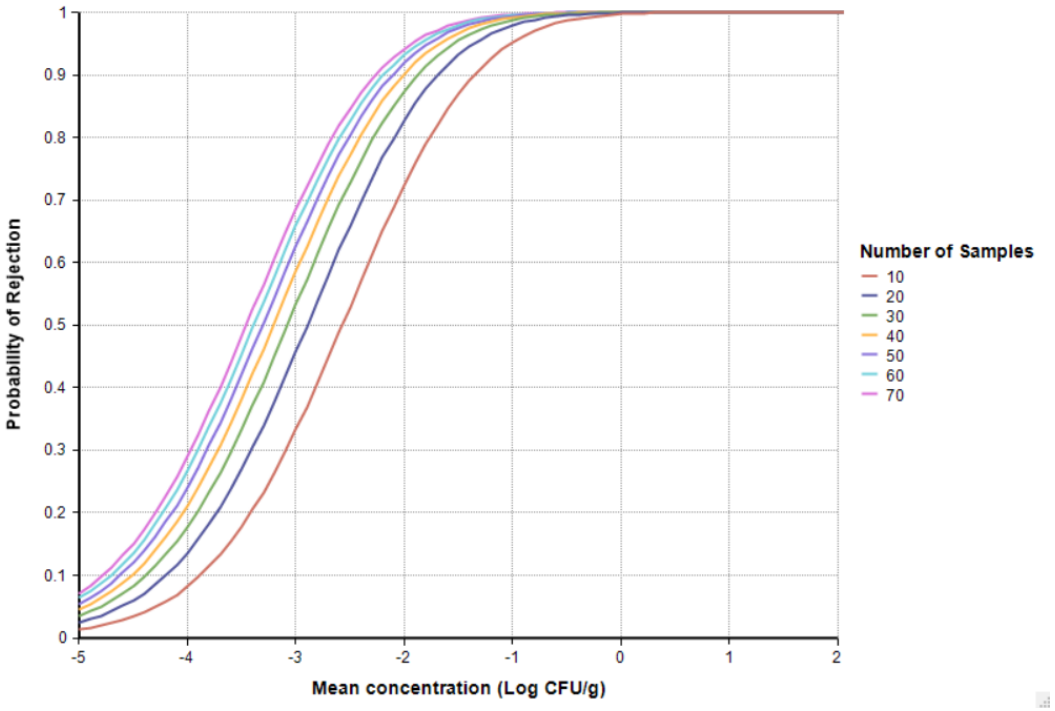


Figure 1: Operating characteristic (OC) curve showing the probability of rejecting a lot with different mean concentrations for sampling plans with 25g samples and varying options for the number of samples taken. Acceptable positives is set to c=0.

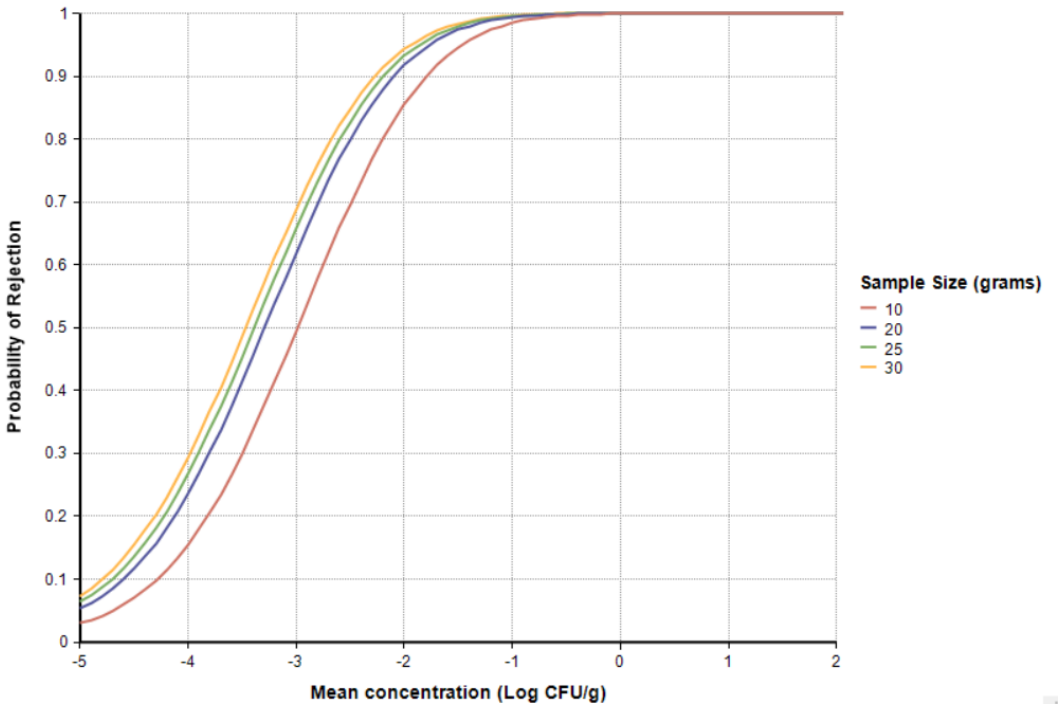


Figure 2: Operating characteristic (OC) curve showing the probability of rejecting a lot with different mean concentrations for sampling plans with 60 samples and varying options for the sample size/mass taken. Acceptable positives is set to c=0.

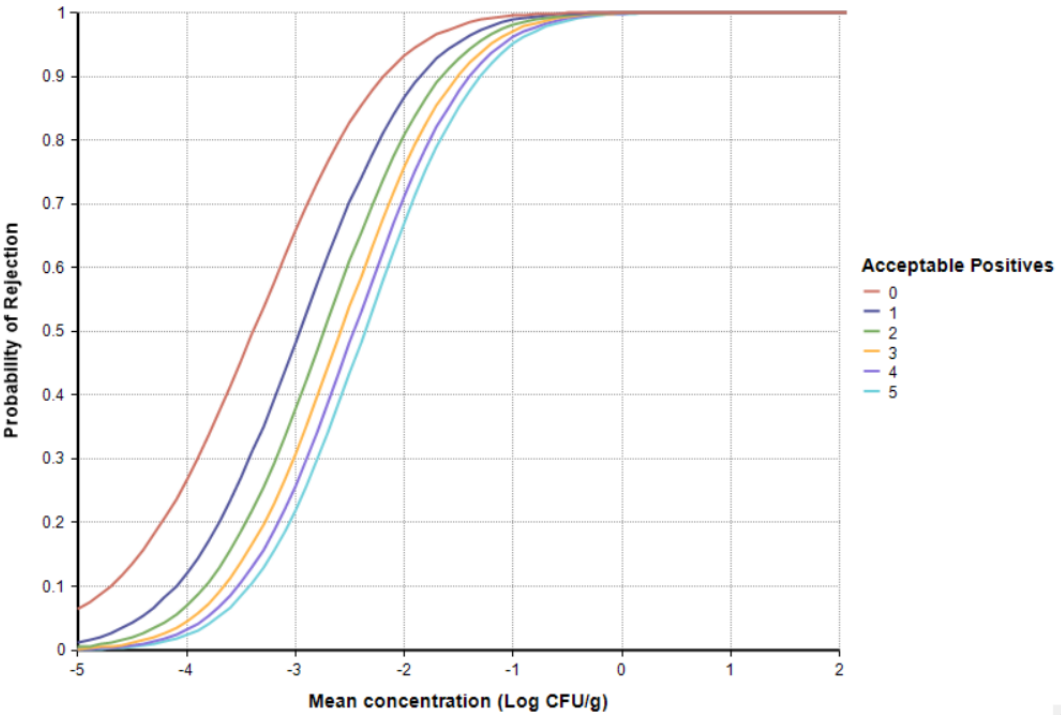


Figure 3: Operating characteristic (OC) curve showing the probability of rejecting a lot with different mean concentrations for sampling plans with 60 samples of 25g with varying options for number of acceptable positives.

Table 2: The probability of rejection (based upon detection of at least one positive sample) calculated for a sampling plan applied to a concentration of -1 Log CFU/g (i.e. 0.1 CFU/g or 1 in 10 g), for 60 samples taken of varying sample size.

Acceptable Positives	Probability of rejection			
	10g samples	20g samples	25g samples	30g samples
0	0.85	0.92	0.93	0.94
1	0.74	0.84	0.87	0.89
2	0.66	0.78	0.81	0.83
3	0.59	0.72	0.76	0.79
4	0.53	0.67	0.71	0.74
5	0.48	0.63	0.67	0.7

The Sampling-Risk Model

To examine the relationship between sampling plan options and consumer risk we developed a sampling-risk model using quantitative microbial risk assessment methodology to link the sampling calculations to estimates of consumer risk following consumption of product. The model describes the changes in pathogen prevalence and concentration along a defined production pathway, and translates prevalence and concentration of the organism into a dose ingested by the consumer. This dose is then used to calculate the risk of illness, which can then be used to determine the risk reduction that can be afforded from sampling strategies in the pathway. The model is built with the intention that it enables risk to be directly compared across sampling scenarios. The conceptual design of the sampling-risk model is presented in Figure 4. Using the sampling-risk model the impact of sampling on consumer risk can be estimated. For example, Figure 5 shows the impact that sampling can have on the number of cases of illness. Results are shown for 1 million servings of produce and a sampling plan of 10 samples of 10g, and no acceptable positives. (Note that this result is purely illustrative, it is not the intention of this project to estimate the number of cases in the US population from contamination of produce with *Salmonella*).

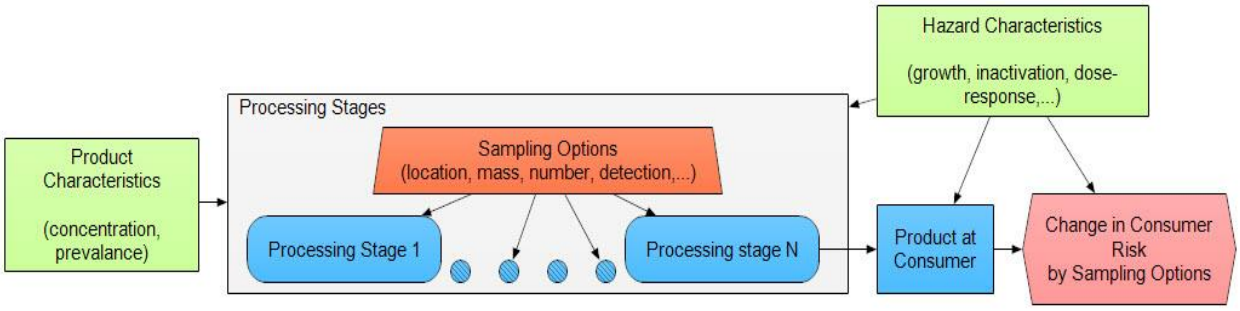


Figure 4: Conceptual design of the sampling-risk model.

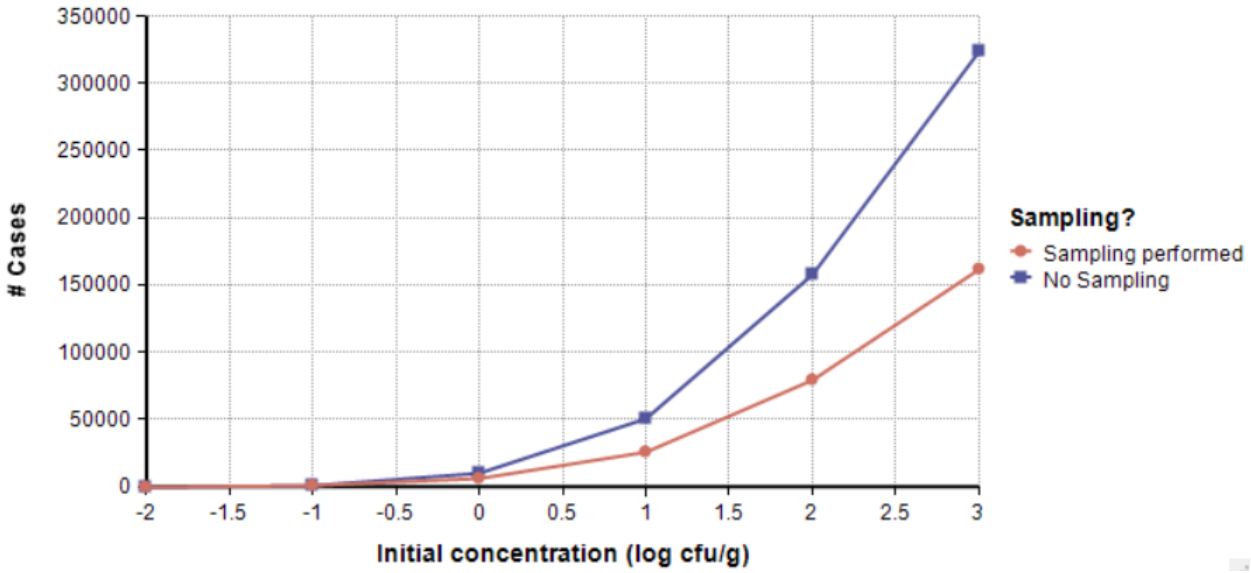


Figure 5: Illustration of the impact a sampling can have in terms of the number of cases of illness compared to when no sampling is performed. Results shown for a range of contamination concentrations when ten 10-g samples are tested and positive lots are destroyed.

Exploring Consumer Risk Reduction

The difference between the general behavior of a sampling plan in terms of lot rejection and exploring the impact in terms of consumer risk reduction is the consideration of the impact of the dose the consumer is exposed to through consumption of the product, taking account of any changes in contamination level that may occur during the process chain (e.g. as a result of inactivation, removal such as washing, or growth). This relationship is described by the dose-response curve. The dose-response curve shows the change in the probability of developing infection[3] or illness for a given number of organisms ingested (the dose). Each organism has a specific dose-response curve. To illustrate, the dose response curve for 5 pathogens is shown in Figure 6. As dose increases so does the probability of infection, reaching saturation at a probability of 1.

The impact of sampling in terms of risk reduction for organisms brings together the characteristics of the sampling plan that can be seen in the operating characteristic curve (OC curve) and the relationship in the dose-response curve. The presence-absence sampling plans employed in produce sampling have an OC curve that is largely to the left of the dose response curve for all 5 pathogens shown in Figure 6. To illustrate, the OC curve showing the probability of acceptance of a sampled lot (sampling plan is 10 samples of 100g each, c=0) and dose-response curve for *E. coli* O157 are compared in Figure 7 (the dose response curve assumes ingestion of 100g of product). The graph shows that as the concentration increases the probability of acceptance decreases to zero, reaching1 in 1 million at approximately -1.4 log CFU/g (0.04 CFU/g or equivalently 1 CFU per 25g of product) for this particular sampling plan.

[3] In dose-response models used in microbial risk assessment the term infection is described as the colonisation of the consumer with the organism, but the individual does not necessarily display any symptoms of illness.

Before reaching zero, there is overlap of the two curves where both curves are non-zero. A segment of this overlap is shown in Figure 8.

For the region where the OC curve has a probability of acceptance of 1 or close to 1 (occurring at the lowest concentrations in our examples) there is little to no impact on risk from sampling, as the plan is unable to distinguish between contaminated and uncontaminated lots. This is Region 1. In this region the risk is driven by the DR curve. Where there is overlap of the region where the OC curve is less than 1 and the DR curve, Region 2, the impact of the sampling on risk is driven by the attributes of the plan (sample size, sample mass, and number of acceptable positives). In areas where the OC curve has dropped to a probability of acceptance of zero (and all lots that are contaminated will be identified[4] and rejected) the impact of the sampling plan is driven by the proportion of lots tested (with consumer risk also influenced by the prevalence of contaminated lots). This is Region 3. The regions are illustrated in Figure 9.

To quantify the relationship between the sampling plan and risk, the assumption is made that samples are taken at harvest and there is no subsequent inactivation or growth of the organism before consumption. In other words, the number of organisms in the product at harvest is the same number of organisms at the point of consumption. The impact of inactivation or growth in the stages between harvest and consumption is explored further in section 3.

When concentrations are in Region 2, the relationships described in Section 1 of this report are observed this time with the impact on risk. For example, the number of samples taken has an impact on risk (see Figure 10). The more samples taken, the greater the risk reduction, where risk reduction is given by the ratio of risk with and without sampling.

The proportion of lots tested and prevalence of positive lots influence the risk to the consumer as they determine the number of contaminated lots that will reach the consumer. An increased prevalence of contamination results in an increased number of cases, and an increased absolute number of cases averted by sampling. However, it does not affect the magnitude of the risk reduction from sampling (in terms of the ratio of cases with and without sampling) with the relationship between the characteristics of the sampling plan and rejection rates not influenced by the prevalence of contamination across lots.

The influence of the proportion of lots tested is determined by the concentration; this is illustrated in Figure 11. As concentrations increase within Region 2 there is an increasing risk reduction for increasing proportions tested. As concentrations increase into Region 3 the gain is maximized, with the limit dictated by the proportion of lots tested (as shown by the saturation (flattening) of the curves).

[4] Additional factors that influence the identification of contaminated lots are related to the specifics of the testing methodology which may be specified in a microbiological criteria for example. These factors would include the sensitivity and specificity of the testing approach. These factors have not been incorporated here were the relationships assume a test that is able to detect contamination when it is present with 100% specificity and sensitivity.

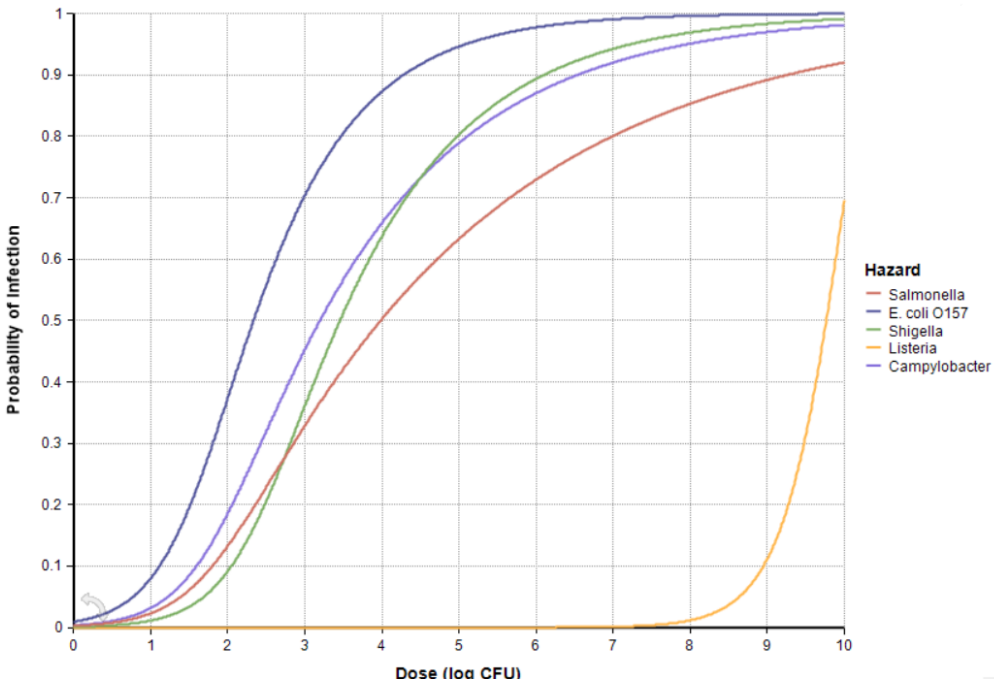


Figure 6: Illustration of the dose-response curves for 5 pathogens.

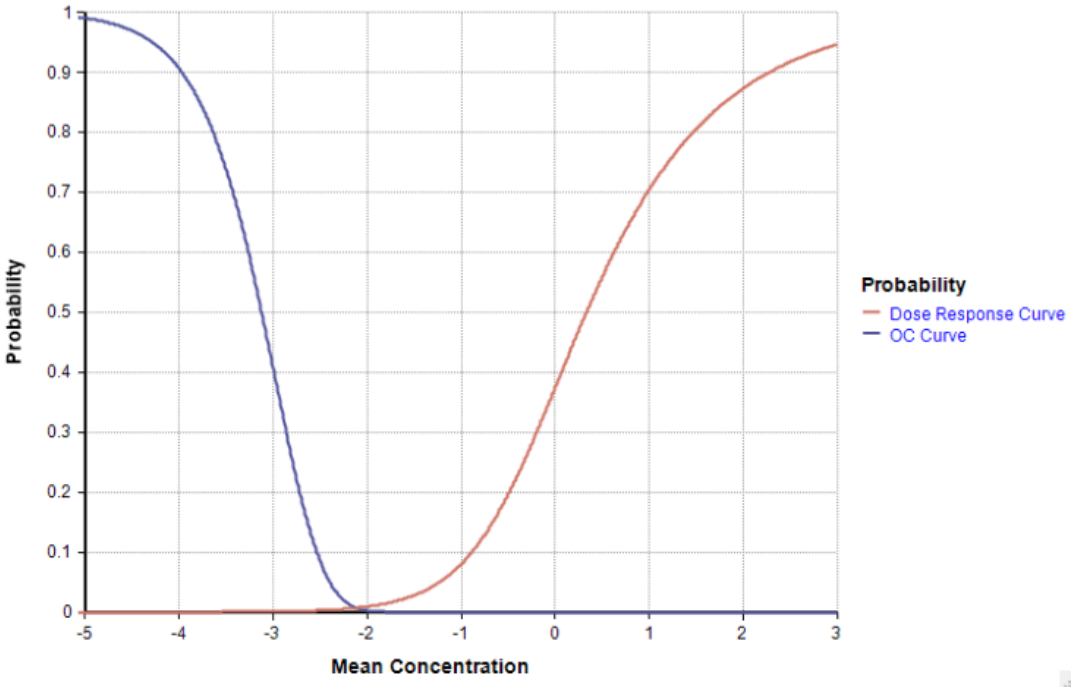


Figure 7: Comparing E. coli dose response curve to with the OC curve. (Note that the dose response calculation assumes ingestion of 100g of contaminated product to be consistent with estimates of risk in analyses detailed in the following sections.)

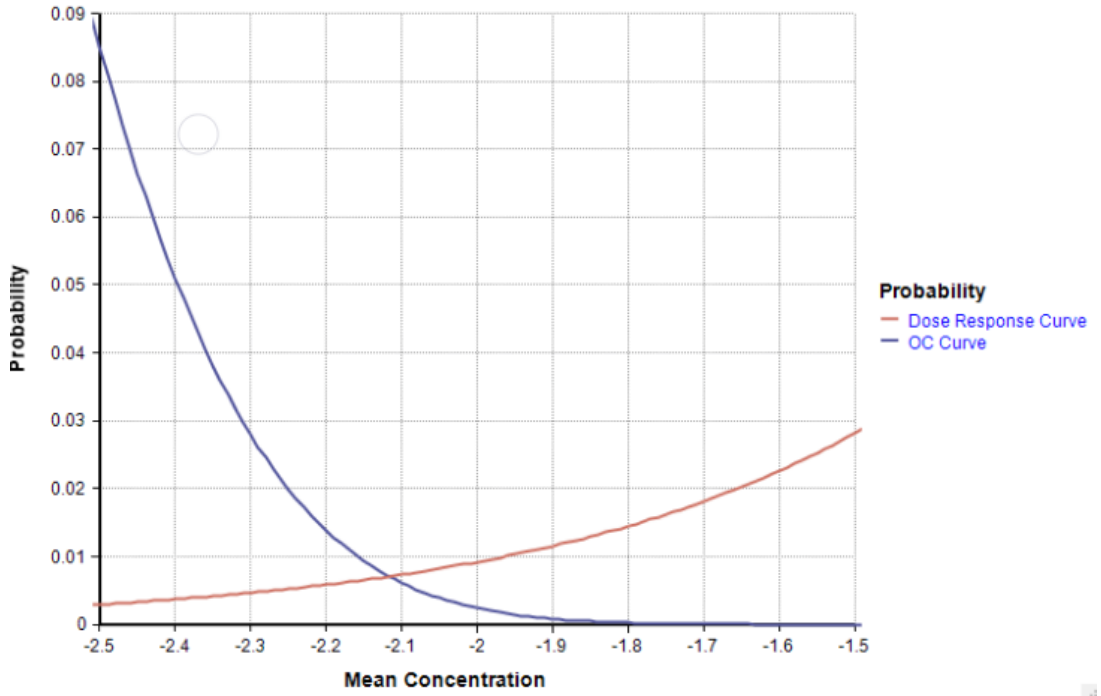


Figure 8: A closer look at the region where the curves intersect.

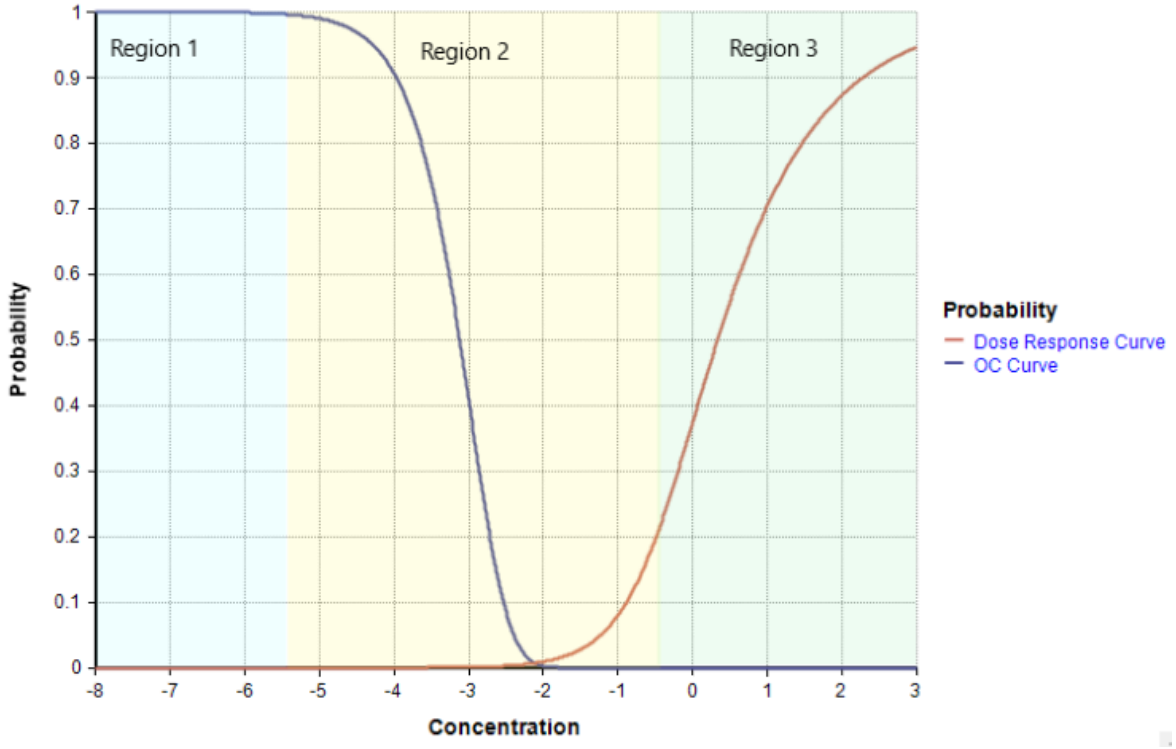


Figure 9: Illustration of the three regions of behavior when comparing a dose-response and OC curve.

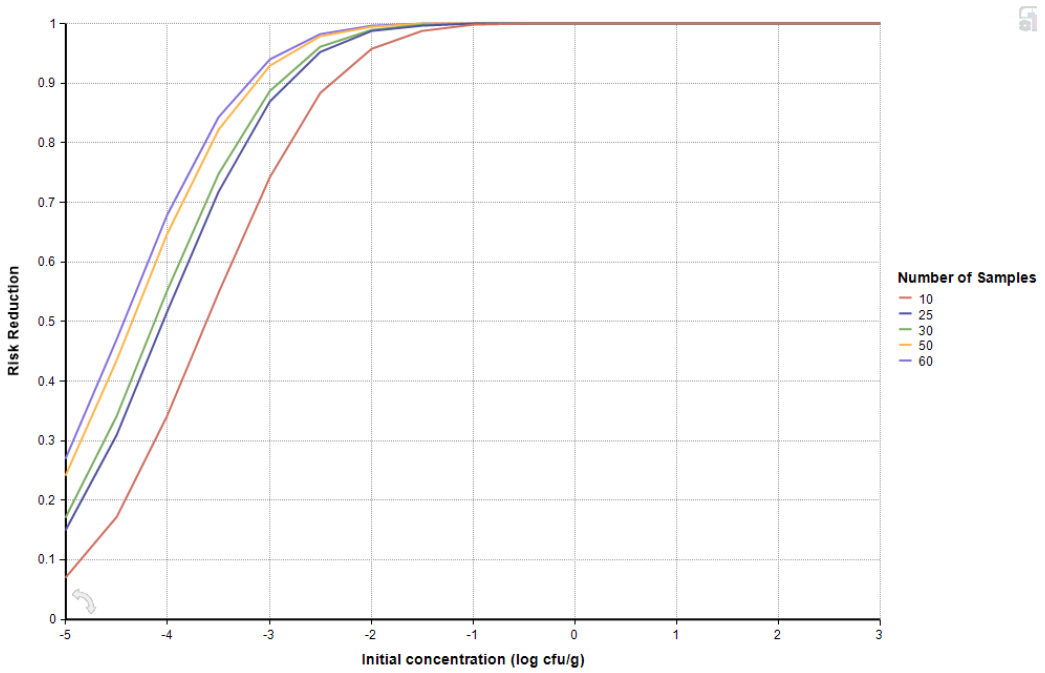


Figure 10: The risk reduction from sampling for a range of concentrations and number of samples. Result are based on Salmonella, 10g sample size, 0.1 (10%) prevalence (i.e. probability of being contaminated at any concentration), and c = 0. E. coli results are the same.

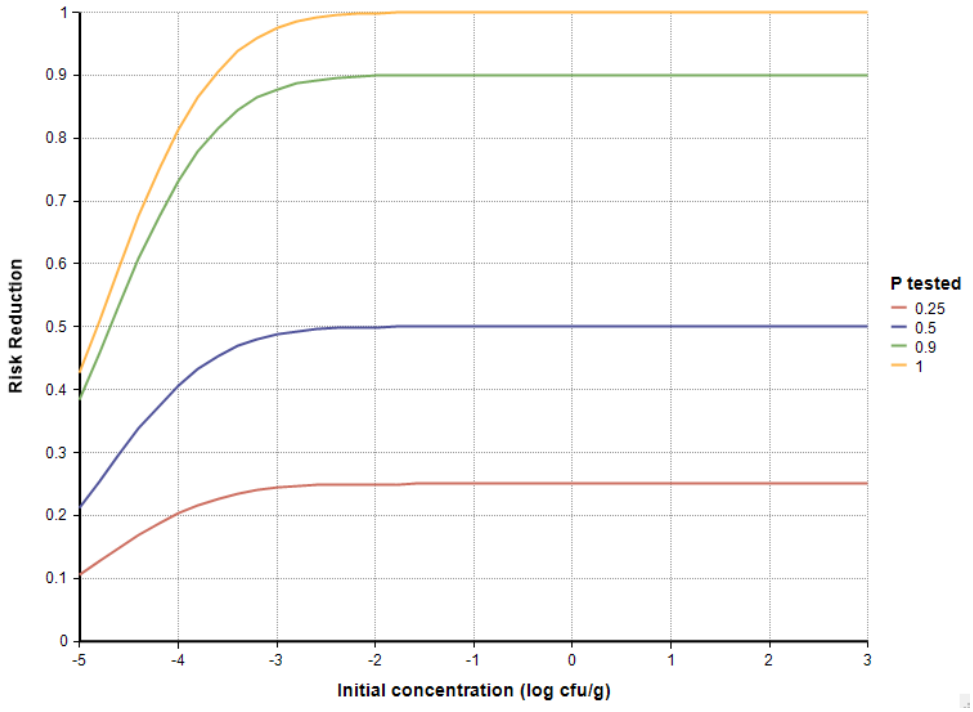


Figure 11: The risk reduction from sampling for a range of concentrations given different proportions of lots tested (P tested). Results are based on E. coli O157, 60 samples of 25g (c=0).

Impact of the Location of Sampling

Sampling is often conducted at the point of harvest; however there are other opportunities available to take samples. When choosing an appropriate location for sampling the aim is to achieve the best risk reduction coupled with the most practical strategy; for example if location does not affect the risk reduction obtained, then sampling earlier in the supply chain to lead to early identification of contamination may be the preferred option to avoid supply chain disruptions that could result from identification of contamination later in the supply chain (e.g. cascading recalls). This section explores the impact of the location in the production chain to begin the conversation around the best point in the chain for sampling to occur (in terms of consumer protection).

A change in the impact of sampling given a different location will only occur if there is a process occurring that changes either the prevalence or concentration of organisms present on the product. To explore the impact on risk reduction we consider two processes that can affect the contamination profile: a reduction step that can affect both prevalence and concentration (for example a washing step), and a thermal deviation that supports bacterial growth (examples include temperature abuse, or a break in the cold chain). Using the sampling-risk model, the model is adjusted to describe three different scenarios: 1) inclusion of a reduction step and no thermal deviation, 2) exclusion of a reduction step but including thermal deviation, 3) inclusion of both a reduction step and a thermal deviation. It is assumed that the reduction step occurs after harvest, and that the growth (if any) occurs after the reduction step. The sampling calculation is then evaluated at three locations: 1) “Post harvest” – this is at the first point in the process chain, 2) “Post reduction” – this is after step two of the process chain which may or may not include the reduction step, and 3) “Post growth” – this is after step three which may or may not include a thermal deviation. The samples are not taken sequentially. This process is illustrated in Figure 12. This is a simplistic view of a production chain; however if a particular product chain has similar characteristics then this information is provided to support development of more in-depth investigation into the best options for that product.

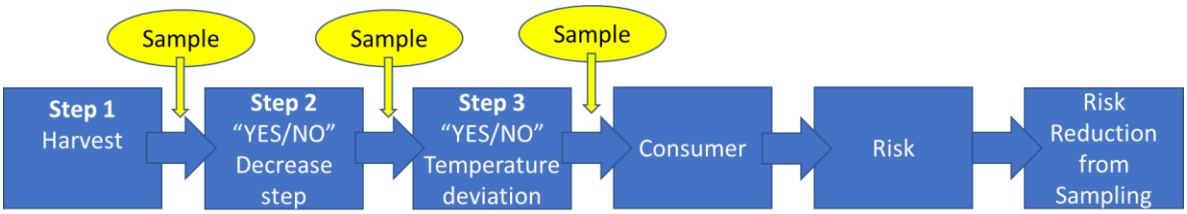


Figure 12: Summary of the steps in the sampling-risk model.

Understanding the Results

- Risk reduction is the risk compared to no sampling, and isolates the additional benefit of sampling in addition to any other impact. For example, if there is a reduction step, the reduction itself affords a risk reduction, however the risk reduction measure isolates the benefit from sampling in addition to this impact.
- Unless otherwise indicated the results are based upon sampling using 60 10g samples, with no acceptable positives.
- Unless otherwise indicated, results presented below are for scenarios with *Salmonella* growth and dose-response models, however equivalent relationships are observed for *E. coli* O157.

Scenario 1: inclusion of a contamination decrease step and no growth

Scenario 1 explores the impact of location when there is a reduction step following harvest (see Figure 13). Three different contamination decrease scenarios are compared: no decrease, a decrease process that results in a 2-log reduction, and a decrease process that results in a 4-log reduction. (The mechanics of how the decreases are achieved are not described. Any decrease process that results in a reduction of 2- or 4-logs would be represented). The results for the three contamination decrease scenarios are shown in Figure 14.

- When there is no decrease, sampling provides an additional risk reduction compared to no sampling, but the location has no impact on the magnitude of the benefit. This is because there is assumed to be no process that adjusts either concentration or prevalence in the production chain.
- When the decrease in contamination is a 2-log reduction, sampling brings an additional risk reduction through sampling at Step 1 (post harvest). Sampling at either Step 2 or Step 3 has gains over no sampling (a non-zero risk reduction), but there is no difference between Step 2 and 3 because both have the benefit of the contamination decrease step, but not other difference in contamination profile.
- When sampling before the contamination decrease (reduction) step occurs, the gain in risk reduction from including sampling is the greatest as the contamination decrease is performed on a concentration that has already been reduced by the sampling process.
- When sampling happens after the contamination decrease step the gain is dependent on the decrease (reduction) amount. For a 2-log process, sampling affords an additional benefit (in terms of risk reduction), but this benefit is not as great as when performed prior to the decrease step.
- For a decrease from a 4-log reduction process there is no appreciable additional benefit from sampling in terms of risk reduction (as shown by a risk reduction close to 0).
- The gain from sampling post harvest when there is a reduction step is maintained with increasing product concentration, see Figure 15.

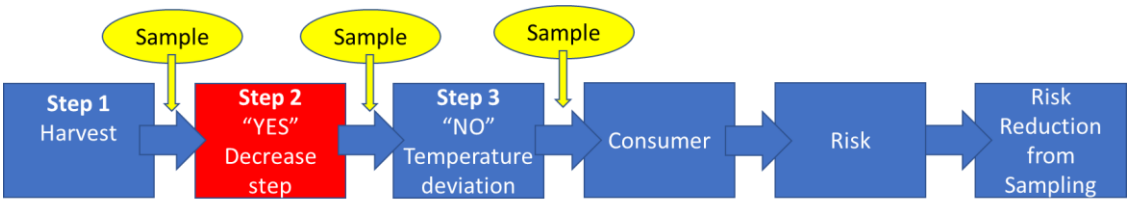


Figure 13: Illustration of the set-up for Scenario 1: inclusion of a contamination decrease step and no growth.

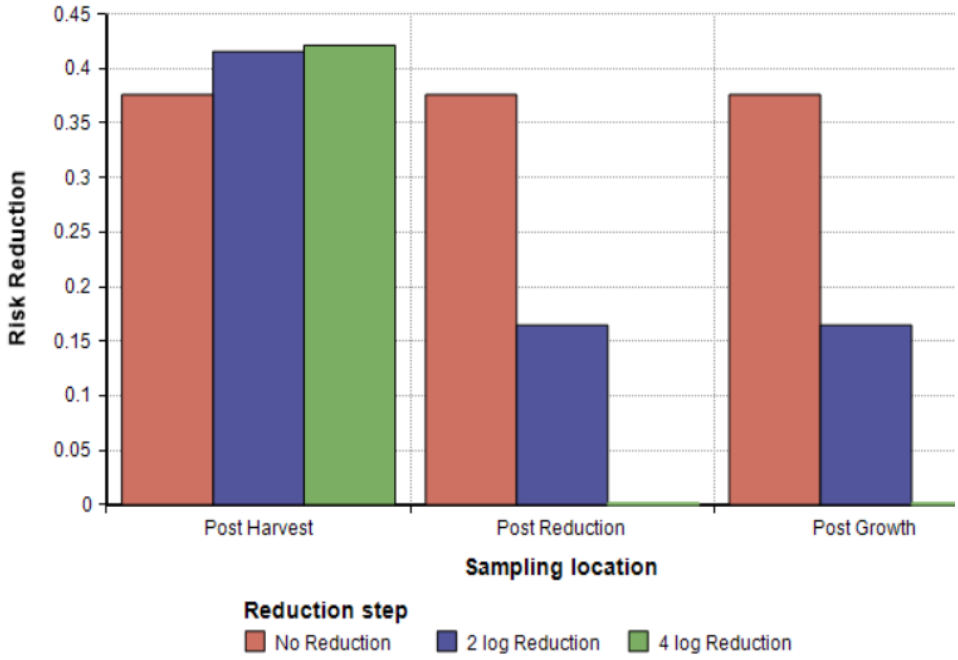


Figure 14: Comparison of the risk reduction for three contamination decrease scenarios (Reduction step) when the sampling location is changed. (Concentration is set to -2 log CFU/g (no growth). Risk reduction is compared to no sampling and isolates the additional benefit of sampling.)

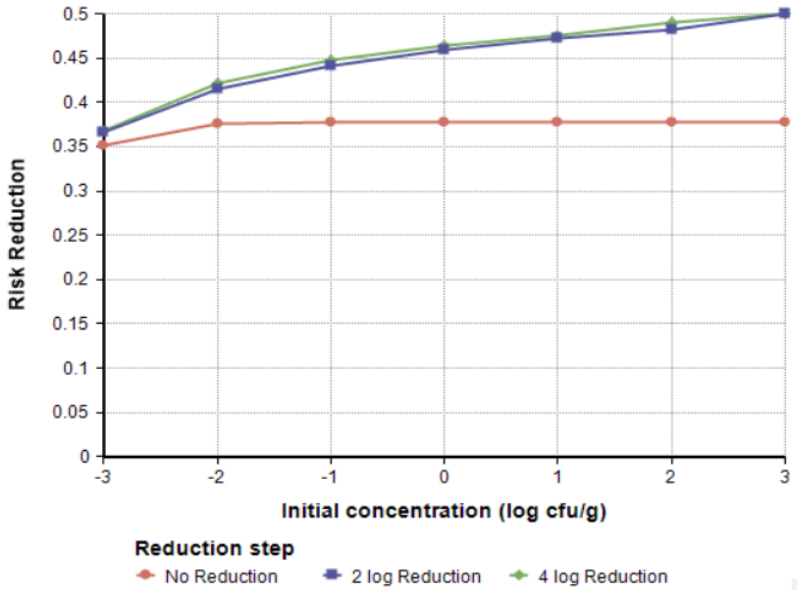


Figure 15: Risk reduction from sampling (compared to without sampling) for different contamination decrease (reduction step) scenarios, shown when sampling at Step 1, for a range of initial concentration assumptions (along the x-axis).

Scenario 2: Exclusion of a reduction step but including thermal deviation

A report from the Produce Marketing Association (PMA) stated that 70% of the food consumed in the U.S. is handled by the cold chains, and that 25% of all food products transported in the cold chain are wasted each year due to breaches that result in fluctuations in temperature. Scenario 2 looks at the gain from sampling by location in the presence of a temperature deviation, with no reduction step (see Figure 16). The temperature deviation is described by calculating the amount of growth that could occur for different deviation scenarios where the time window for growth is 2, 12, 24, or 72 hours. It is assumed the temperature for this time window is 20°C/ 68°F.

- Sampling has the biggest impact when the window for growth is 2 hours.
- As the time that growth occurs increases the benefit from sampling decreases when sampling occurs at either Step 1 or Step 2 (i.e. before the growth occurs). This is because the reduction in concentration from sampling is later reduced by the increase in concentration from growth.
- When sampling occurs after the temperature deviation the benefit of sampling is equivalent across the different deviation scenarios, as in each case the sampling is reducing the growth-increased concentrations.

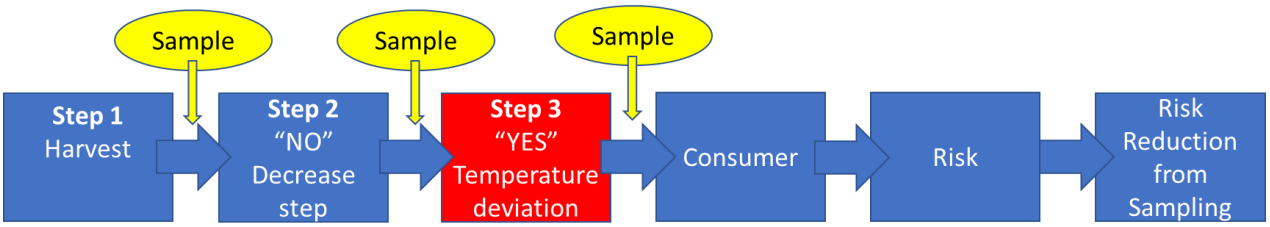


Figure 16: Illustration of the set-up for Scenario 2: Exclusion of a reduction step but including thermal deviation.

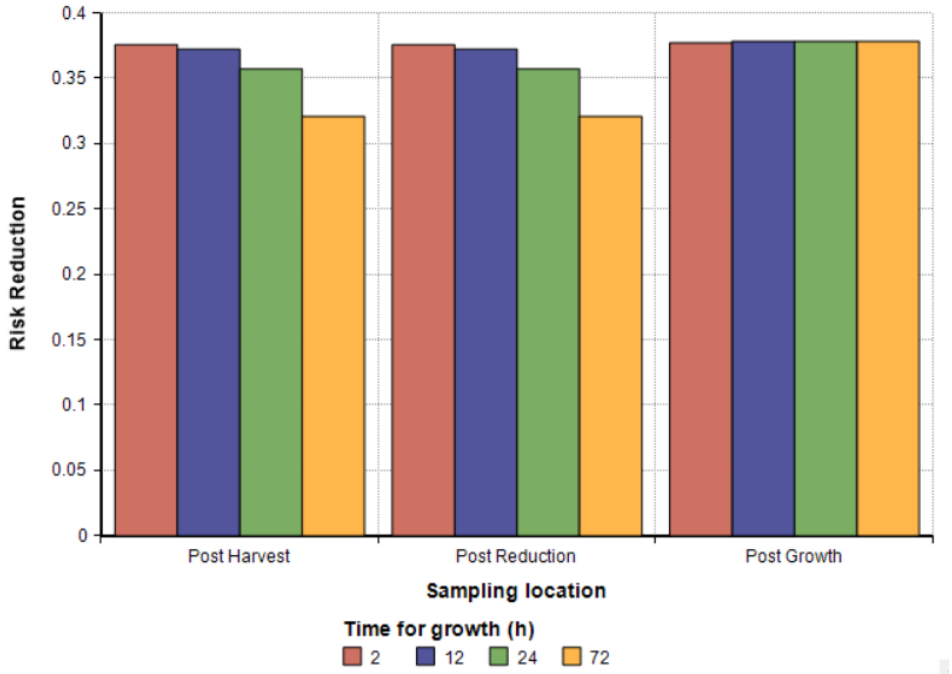


Figure 17: Comparison of the risk reduction for four growth scenarios by sampling location. Sampling plan is 60 samples of 10g, c=0.

Scenario 3: Inclusion of both a contamination decrease step and a thermal deviation

Scenario 3 includes both a contamination decrease step and a temperature deviation (i.e. growth) and explores the impact of location in terms of the benefit of sampling (compared to no sampling).

- As the time for growth increases the location with the best additional gain in risk reduction (compared to no sampling) switches from Step 1 (post harvest) to Step 3 (after the temperature deviation and therefore post growth).
- Sampling at Step 2 gives additional benefit compared to no sampling, but this gain is lower than sampling at Step 1 or Step 3 and decreases with increasing time-window for growth.

- With an extended time window (72 hours), sampling at Step 3 has the biggest benefit for all contamination decrease (reduction) scenarios. Sampling at Step 2 does not provide addition gain over sampling when there is no reduction, and only provides an addition gain over sampling when there is a contamination decrease of a 2-log reduction. This is because the reduction steps are lowering the concentration to a level that cannot be readily detected by the sampling plan (e.g. in the presence of a 4-log reduction step a -2 log concentration gets reduced to -6 Log/CFU which has a probability of acceptance close to 1 on the OC curve, limiting the utility of sampling).

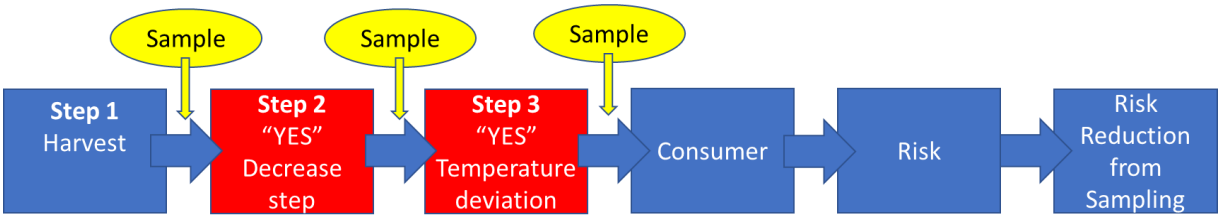


Figure 18: Illustration of the set-up for Scenario 3: Inclusion of both a reduction step and a thermal deviation.

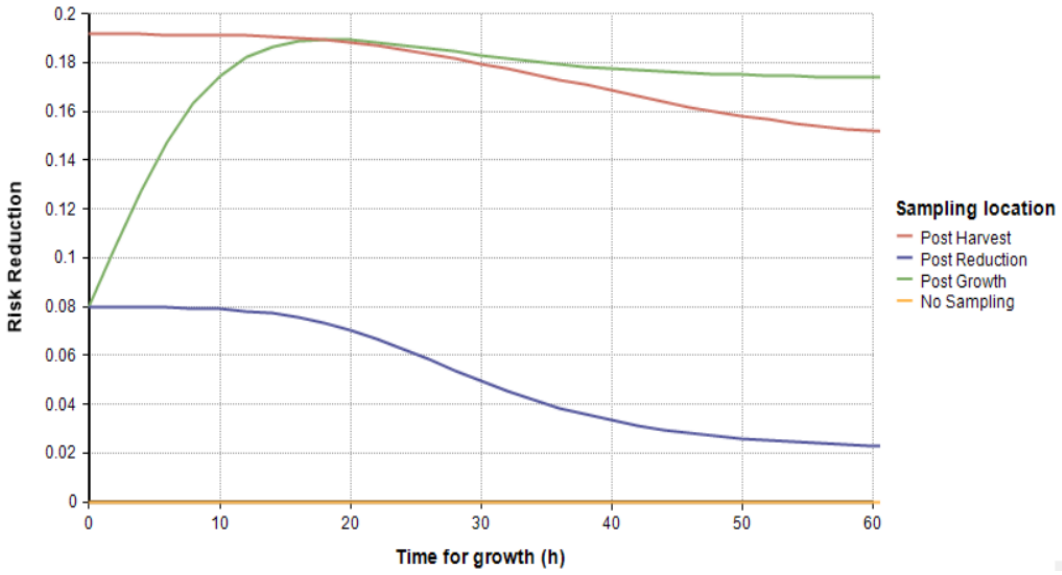


Figure 19: Impact of sampling location for different amounts of time for growth to occur. Risk reduction from sampling (compared to without sampling) is shown for a 2-log reduction scenario (reduction occurs prior to temperature abuse).

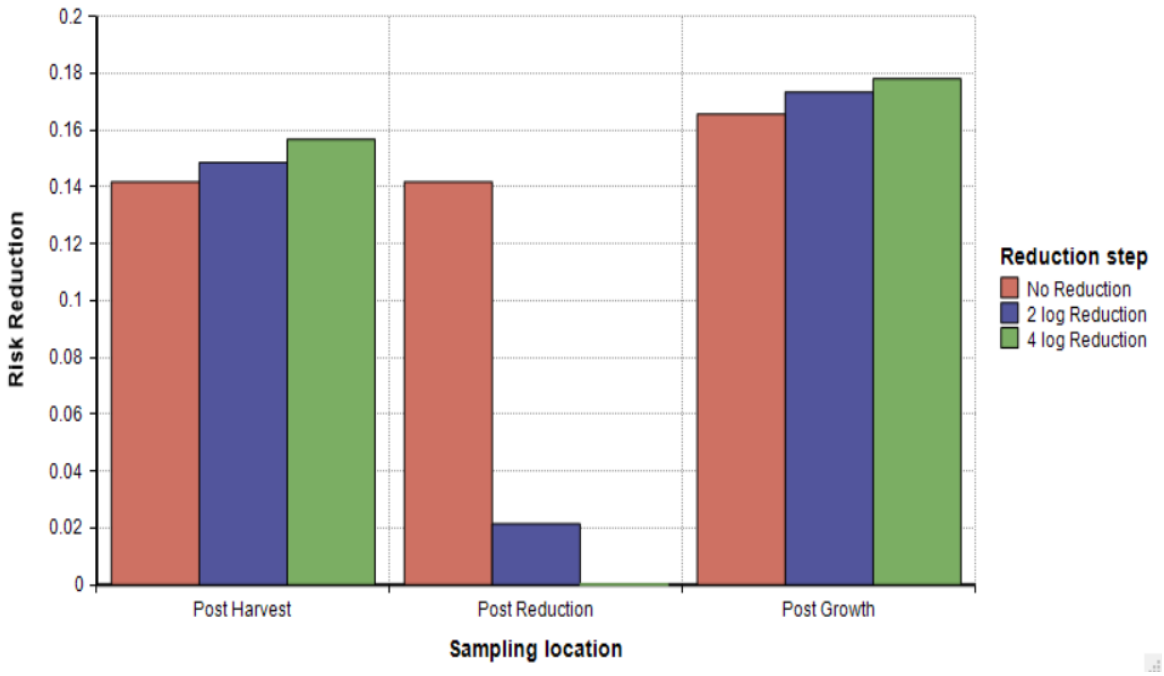


Figure 20: Risk reduction from sampling (compared to without sampling) for 72 hours of growth with a start -2 Log CFU/g mean conc. (Post Reduction, 4 log Reduction value for Risk Reduction is 0.003)

Outcomes and Accomplishments

A detailed analysis of the relationship is presented to describe the role of sampling plan components in determining the lot rejection rates.

Analyses are also presented that extend the role of sampling plans (and sampling plan components) in terms of consumer risk reduction. This analysis is extended to explore the impact of location of sampling in the process chain taking into account the presence of processes that may affect contamination of product (decrease processes, and/or growth).

Summary of Findings and Recommendations

- Exploring risk as a measure of sampling plan performance can yield insights beyond just considering the basic sampling performance metrics commonly described (e.g. lot rejection rates).
- Location of sampling can matter in terms of achieving the best risk reduction possible from a sampling strategy. This depends upon on the characteristics of the production chain and its influence on the microbial contaminated (e.g. reduction, growth). This highlights the importance of understanding the specifics of the production chain and how it influences pathogen behavior, when determining risk mitigation strategies.
- Cost and ease of sampling at different locations should be factored into risk management considerations, e.g. Cost-Benefit analysis. This is not the scope of this project and was not considered at this time.

APPENDICES

Publications and Presentations

Oral presentation: “Exploring the relationship between product testing of fresh produce and consumer risk,” Society for Risk Analysis 2019 Annual Meeting, Dec 8–12, Arlington, VA.

Budget Summary

The total funds awarded to this project were $65,595. Project funds were used to support labor in the model development, model analysis, presentation of results to a stakeholder group, subsequent refinement of analyses, and development of the final set of results presented in this report. Funds were sufficient to complete the work.

Suggestions to CPS

The above work could be extended to include considerations that were out of scope, including:

- Role of cost-benefit considerations (cost could include not only $ cost, but also practicality, time, etc.) when considering different strategies.
- Impact of test methodology employed. Our work assumed a “perfect test”; this is not the reality and could be influential in determining risk-benefit trade-offs for different sampling strategies.

References Cited

CAC (Codex Alimentarius Commission). (2013). Principles and Guidelines for the Establishment and Application of Microbiological Criteria Related to Foods. CAC/GL 21-1997 (Revision 1-2013).

ICMSF (International Commission on Microbiological Specifications for Foods). (2016). Microbiological sampling plans. http://www.icmsf.org/main/software_downloads.html

JEMRA (Joint FAO/WHO Expert Meetings on Microbiological Risk Assessment). (2006). *Enterobacter sakazakii* and *Salmonella* in powdered infant formula: Meeting report. FAO/WHO microbiological risk assessment series no. 10. Rome, Italy.

JEMRA (Joint FAO/WHO Expert Meetings on Microbiological Risk Assessment). (2011). Microbiological Sampling Plan Analysis Tool. http://www.fstools.org/sampling/.

JEMRA (Joint FAO/WHO Expert Meetings on Microbiological Risk Assessment). (2016). Statistical Aspects of Microbiological Criteria Related to Foods: A risk managers guide. FAO/WHO microbiological risk assessment series, no 24. Rome, Italy.

1 Appendix

During preparations for the 2020 CPS Research Symposium the question was posed by an industry member if there was a plan to make the Sampling-Risk model (described in ***CPS Final Report - Hartnett***) available to industry and others, for example in the form of an app-type product. The intention would be to enable industry to perform their own analyses independently for the specific parameters of interest to them. This was not part of the initial project plan and was not incorporated into the timeline or budget. However, the opportunity arose to explore this idea in a proof-of-concept manner. *The specific task is to develop a first "proof of concept" of an application or tool for the industry to use to do analyses themselves. This will be developed in Analytica (the software used to develop the tool) to show possible functionality, and a summary appendix for the project report will describe the "proof-of- concept" tool.* There are two goals:

- to prove that the Analytica model can be adjusted to allow for a non-expert user to interact with the model and perform calculations—this will be achieved by developing a user interface that avoids the need to understand the underlying complexities of the mathematical calculations and software implementation
- to prove that the types of user interactions developed in Analytica can be brought to an app-type environment

This appendix summarises the findings.

1.1 User interface development

To enable user interaction with the Sampling-Risk model, we developed a user interface that provides the ability to conduct the types of analyses described in the ***Final Report***. To provide a useful tool to be used by industry (and others) the capabilities of the Sampling-Risk model were expanded, and user controls were developed for the interface. Specific expansions developed to explore the feasibility of transforming the Sampling-Risk model into a tool to be used by others were:

- ability to select any combination of the process steps to explore the role of the location of sampling,
- ability to select any combination of the pre-defined log reduction options to compare the impact in risk reduction from sampling ('No Reduction' is included as an option),
- providing the option to include a user-defined custom log reduction for a decrease step, and the option to include/exclude this custom input,
- ability to include or exclude pathogen growth, and to choose the temperature.

These expansions provide flexibility beyond a fixed list of inputs as used to demonstrate the relationships between sampling and risk presented in the *Final Report*. These expansions demonstrate that the tool can be enhanced with additional flexibility that would allow industry to better reflect their particular circumstances in terms of how their processing situation may affect the pathogens present. The interface itself is shown in **Figure A1**. The interface was built to reflect an app-type process consisting of three sequential pages.

Page 1 is the front page with a basic introduction (that can be readily adjusted) and the initial inputs needed to define the specific sampling plan variables. Page 2 provides options to select the location(s) of sampling to be analysed and compared, the features of processing including the magnitude of a reduction step (if included), and the choice to include pathogen growth in the calculations with the temperature for growth specified. Page 3 shows the results provided as a graph in the same layout as

throughout the analyses in the ***Final Report***. Next and Back buttons were developed to provide navigation between the pages, and a Help button was developed to provide capacity for instructions, glossary, or other useful user support information.

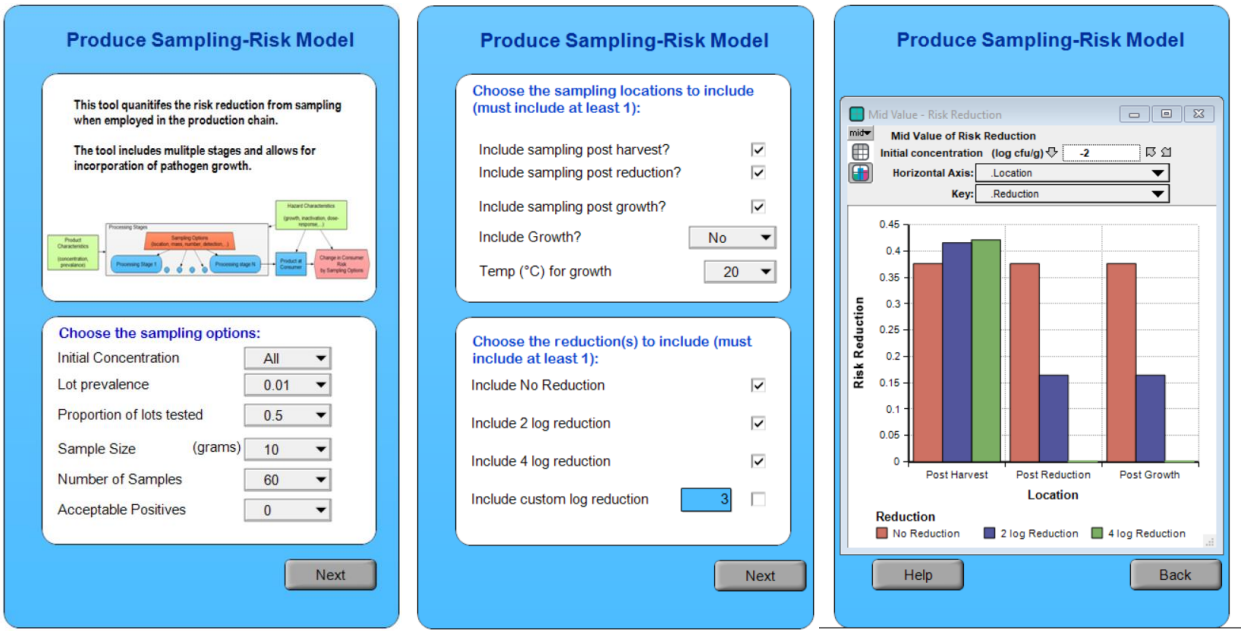
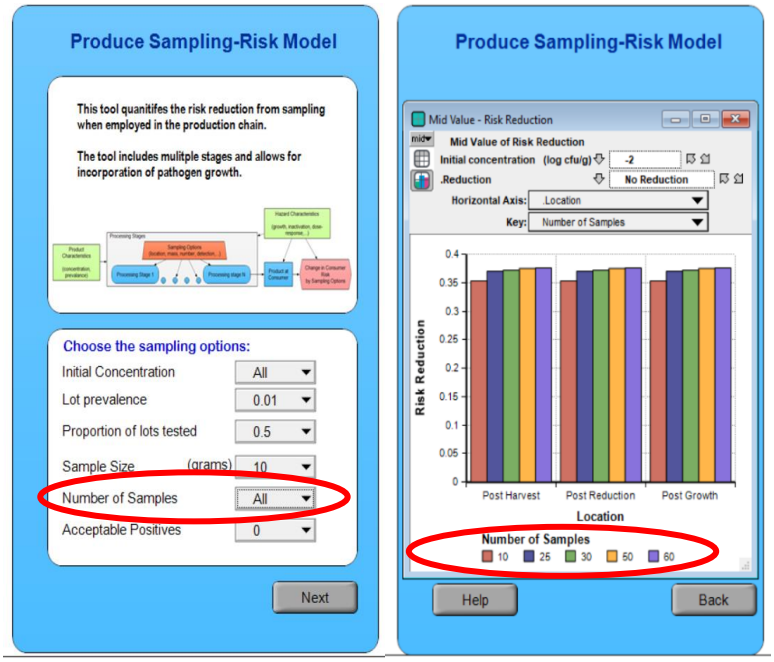


Figure A1: Screenshots of the three pages of the user interface for the Sampling-Risk tool within the Analytica software.



In the example in **Figure A2** (left), the user input for the number of samples has been adjusted to "All" which means that all of the options in the dropdown box are calculated in parallel. This change is then reflected in the results graph with different coloured bars for each number of samples run through the calculation.

Figure A2: Illustration of the capacity to make alternative selections on the user interface.

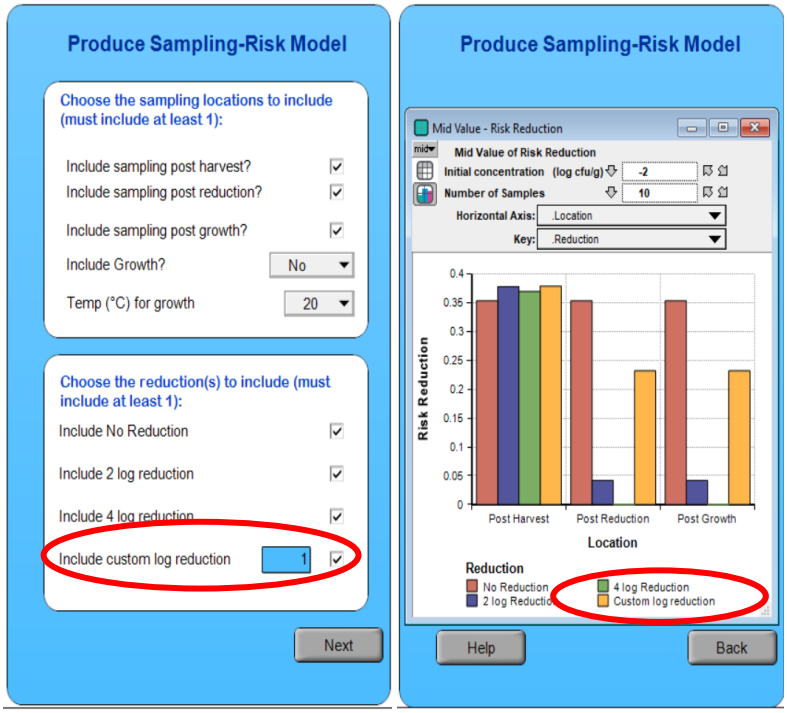


Figure A3: Illustration of the custom log reduction feature.

In the example in **Figure A3** (left), the custom log reduction feature is demonstrated. A log reduction of 1 is added to the analysis (the magnitude is entered by the user). This calculation is conducted in parallel with the other selected reductions, and the result is added to the graph (yellow bar).

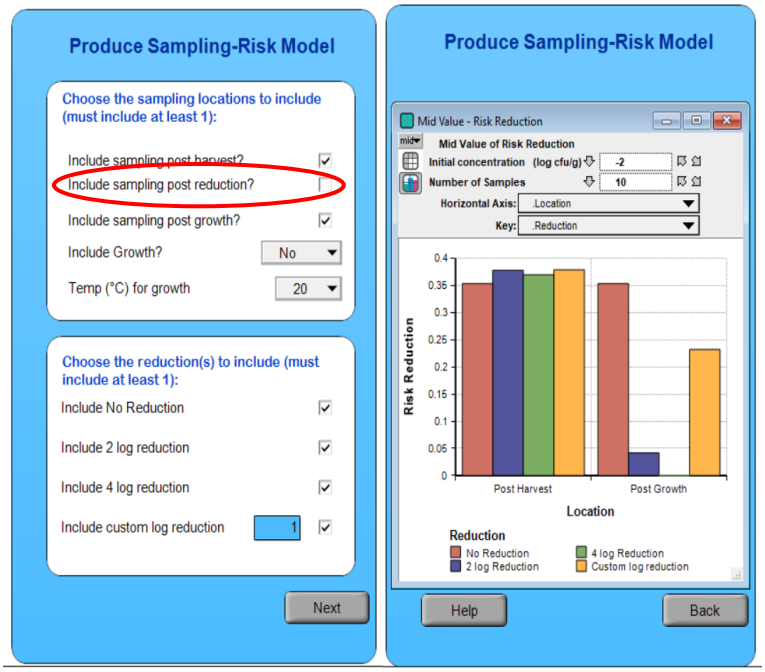


Figure A4: Illustration of the capacity to select which sampling locations are included in the analysis.

In the example in **Figure A4** (left), the ability to select which location(s) to assess (and compare) is shown. In the example the sampling post reduction step is removed from the analysis by unchecking the box. With this action, Post Reduction values are not shown in the results. This example also shows the inclusion of the custom log reduction, illustrating that multiple customisations can be selected at the same time.

1.2 Proving the tool with user interaction can be developed in a more app-focussed environment

The proof-of-concept user interface described in Section 1.1 was contained within the original modelling software Analytica and developed to demonstrate that a user could achieve meaningful interaction with the tool and be able to perform similar calculations as described in the ***Final Report*** without the need to make adjustments to the model itself.

To further prove out the idea of a user-focused app, we need to determine if the types of relationships we are able to determine within Analytica can also be described in a platform designed for app development. This was not immediately clear, as Analytica has a very powerful capacity for handling the multidimensional data that the Sampling-Risk model generates (as a result of considering multiple scenarios in parallel). This capacity allows users to explore many relationships at the same time (for example, Figure A4 examines both the relationship with sampling location and log reduction in the same single graphic). This power is the reasoning behind selection of Analytica as the modelling tool of choice for the Sampling-Risk model. However, it would not be the first choice for app development (nor is it intended to be by design). We therefore explored if the types of relationships being generated could be provided to the user in a similar manner using platforms intended for app-type products. To do this we utilised R, which has both modelling capacity and a specific developer environment Shiny that focusses on development of user interfaces for statistical tools to be used by others.

The available time and budget did not allow a re-implementation of the Sampling-Risk model within R, so to test the analysis capacity summary results data was extracted from the Analytica file and used as the basis for exploration. This is sufficient to meet the needs of the task.

A user interface was developed using R-Shiny that can be shown in a browser, as shown in **Figure A5**. The interface includes responsive design and therefore automatically re-sizes to the screen. The tool is responsive to changes in input selections by the user, enabling options to be explored rapidly. This is illustrated in **Figure A6**, which shows the ability to change the sampling locations that are included in the analysis.

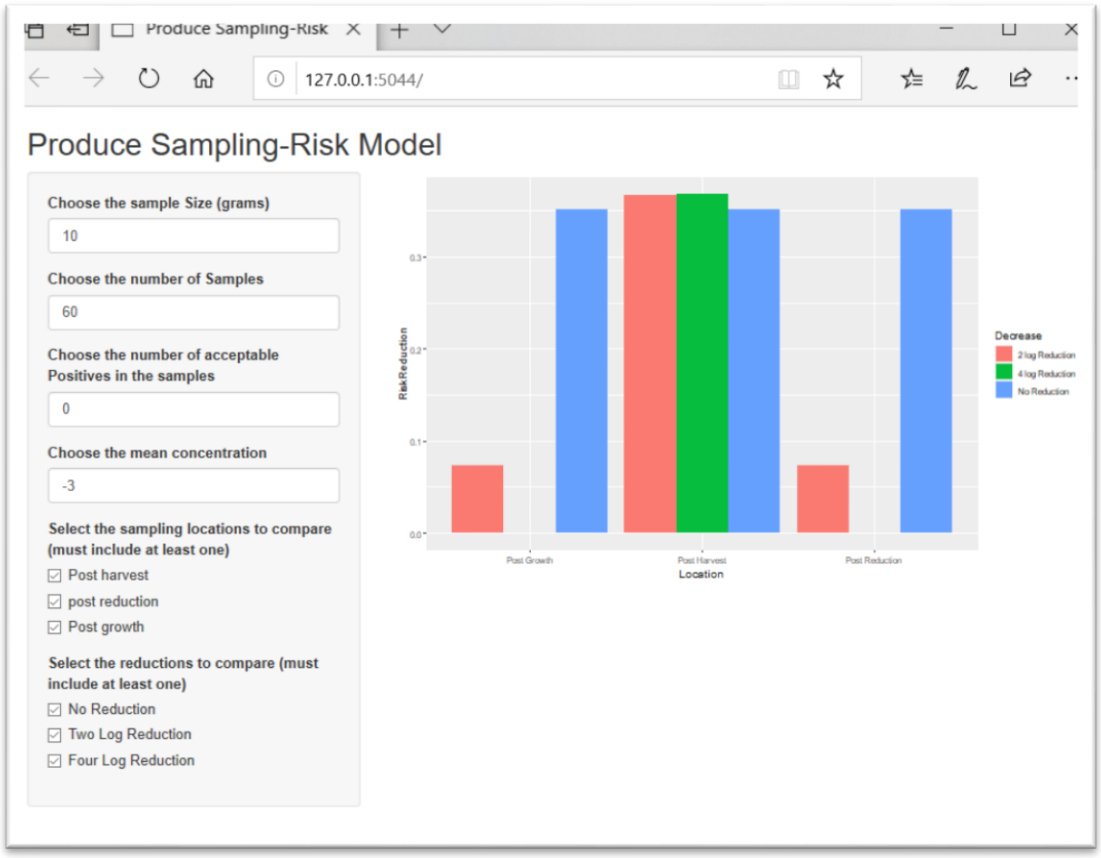


Figure A5: Screenshot of the tool interface as displayed on a browser. As selections are made, the results presented in the graph are updated. See, for example, Figure A6. (Note that at this time, sample size, number of samples, and number of acceptable positives is for display only; all other inputs are 'live' and will adjust the graph when different selections are made).

Figure A6: Illustration of the capacity to select which sampling locations are included in the analysis.

1.3 Key findings

An early proof-of-concept interface has been developed that demonstrates the feasibility of further expansion of the Sampling-Risk model to an app-type product that enables industry to independently explore the relationship between sampling and consumer risk.

1.4 Next steps

It has been demonstrated that an app-type tool is feasible and could be a reasonable next step for this work. A key choice would be the exact platform, with the following requirements:

- the core mathematical functionality of the Sampling-Risk model must be reproducible in the environment selected,
- provides compatibility with a range of platforms likely to be used by industry (iOS, Windows, Android etc.),
- ideally provides future sustainability of the tool (for example, it should not be reliant upon software that may become obsolete in the next 2 or 5 years).

The overall process would require sufficient time/budget to allow for implementation of the model within the app environment in addition to the coding of the interface itself. Ideally, a consultation process would be undertaken to establish if there are additional features/capabilities that would be useful to industry to ensure the final product would be useful in risk management activities. A pilot-test process should also be undertaken to ensure stability and usability of the system.