CENTER for PRODUCE SAFETY

CPS 2020 RFP
FINAL PROJECT REPORT

Project Title
Digital farm-to-facility food safety testing optimization

Project Period
January 1, 2021 – December 31, 2022

Principal Investigator
Matthew Stasiewicz, PhD
University of Illinois at Urbana-Champaign
Department of Food Science and Human Nutrition
103 Agricultural Bioprocess Laboratory
Urbana, IL 61801-4714
T: 217-265-0963
E: mstasie@illinois.edu

Objectives

1. *Build a Field-to-Facility generic supply chain model of produce safety testing.*
2. *Adapt the supply chain and collect parameters to represent a variety of higher-risk commodities with distinct risk profiles and risk-management options.*
3. *Optimize testing across the supply chain of each commodity incorporating representative testing programs at primary production, harvesting, receiving, processing, and packing and assessing their impact to manage safety.*

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

FINAL REPORT

Abstract

Currently, there are inconsistencies in food safety tests and reject (sampling) plans. Buyers and importers may require product sampling for market access, or as part of a grower's food safety preventive controls to detect incoming contamination. These often focus on single points in the supply chain (preharvest or finished product). To better understand the impact of sampling, this study simulated representative sampling plans and processes for multiple commodities contaminated with relevant hazards (leafy greens & Shiga toxin–producing *E. coli* [STEC]; tomatoes & *Salmonella;* cilantro & *Cyclospora cayetanensis*).

A leafy greens process model (study 1) was developed to simulate seven processing systems containing food safety interventions: an optimal system (All-Interventions), a suboptimal system (No-Interventions), and five in-between scenarios, where single interventions were removed. A scenario analysis combined the seven processing systems, three contamination clustering spreads (widespread contamination covering 100% of the field, a randomly located 10% of field cluster, and a randomly located 1% of field cluster), and seven sampling plans (preharvest, harvest, receiving, finished product, and customer sampling) for 147 total scenarios. Concerning the performance of food safety interventions, the five interventions combined caused a 3.43 (3.33–3.56, 95% CI) log reduction to the total adulterant cells that reached the system endpoint (endpoint TACs). The most effective single interventions were washing, pre-washing, and preharvest holding, which caused a 1.32 (1.22–1.45, 95% CI), 1.27 (1.18–1.38, 95% CI), and 0.80 (0.73–0.90, 95% CI) log reduction to endpoint TACs, respectively. Concerning the performance of sampling plans, sampling plans before processing interventions occur (sampling at preharvest, harvest, and receiving) had a relatively higher probability of detecting contamination (4%–30%) and caused relatively larger reductions in endpoint TACs from rejecting lots that tested positive (reductions of 0.05–0.66 log). In contrast, sampling plans after processing (finished product, and customer sampling), had a lower probability of detecting contamination (0.1%–7%) and cause small reductions in endpoint TACs (0.04 log). Overall, these results support the idea that leafy green product testing is most beneficial to detect contamination early in the system, before implementing otherwise effective food safety interventions.

A 42-day tomato season (study 2) was simulated. Tomatoes were contaminated preharvest with four different contamination spreads (widespread contamination covering 100% of the tomatoes, 10% cluster, 1% cluster, and 0.1% cluster). Scenarios included 96 combinations of the four contamination spreads, four sampling locations (preharvest, harvest, receiving, packing), and six sampling plans (2, 6, 20, or 60 tomatoes, 20 or 60 tomato mash). The model suggests the best sampling location is dependent on the initial contamination cluster. For the widespread (100% cluster), 10%, and 1% clusters, sampling plans at harvest had the highest reductions to endpoint TACs, between 39.6% to 98.1%. For the very small, 0.1% cluster, sampling plans at packed product had the highest reduction, 12.3%. This study also demonstrates that sampling plans with higher composite mass (60 tomatoes and 20 tomatoes) yield higher detection power than those with lower mass for the widespread (100% cluster) and 10% clusters of contamination. For the 1% and 0.1% clusters, sampling plans with more sampling points (60 tomatoes and 60 tomato mash) had higher detection power.

A 45-day cilantro growing season (study 3) was simulated. The process represented *Cyclospora cayetanensis* dynamics in a harvest of cilantro (45 days) assuming contaminated water was used for irrigation (contaminated with a low level of 0.6 oocysts/L or a high level of 20 oocysts/L) of a 22,000 lb field, with contamination either on one random day of irrigation or over the whole harvest period. Data from two published studies were used to fit a logistic regression to the probability of detection based on oocysts present in a sample of 10 L of agricultural water or 25 g of produce.

The simulation showed that even a single 10-L water sample would reliably detect contamination >2 oocysts/L (median 100% detection). For product testing, a single 25-g sample would reliably detect contamination >0.5 oocyst/g (median 100% detection). In both cases, as the contamination level decreases, more water or produce sample mass or number are needed for a high probability of detection.

Overall, these studies demonstrate (i) that interventions play an important role in reducing the total adulterants that reach the system endpoint; (ii) once a system with good food safety interventions is in place, sampling can be optimized by sampling at preharvest, harvest, and receiving for relatively low-level, distributed contamination, as sampling after processing meaningfully reduces the power of sampling; and (iii) for highly clustered sampling, as was simulated in tomatoes, packed product sampling is likely more powerful, as processing increases cross-contamination, hence the chance a contamination portion is sampled.

Background

This study aims to address the industry and academic need to understand the role of product test and reject (sampling) plans within farm-to-facility systems for multiple commodities and multiple hazards. In the United States, from 2018–2020, 3 foodborne disease outbreaks have been linked to spinach and lettuce contaminated with *E. coli* O157:H7 [1-3]. Between 2005–2015, tomatoes were involved in 11 multistate *Salmonella* outbreaks [4]. Most recently in 2020, 2021, and 2022, *Cyclospora cayetanensis* has been linked to an increase in domestically acquired cases of cyclosporiasis. These repeated outbreaks have caused concern among producers, packers, buyers, and consumers, causing producers to increase efforts to manage risk through multiple pathways, including food safety interventions, and test and reject (sampling) plans.

Product sampling is a tool that is required many times by customers, buyers, and importers for market access [5-7]. Detection of pathogens by product sampling is challenging when prevalence and/or levels are low, as in the case of leafy greens [5, 8]. The power of sampling is dependent on the total mass collected, the number of grabs, and contamination patterns [9, 10]. Validated simulations have shown that many preharvest sampling plans do not reliably detect low-level or low-prevalence contamination [9, 10]. Therefore, there is still a need to evaluate the power of sampling beyond the preharvest stages and assess the stages where sampling provides the most value. While sampling is not a food safety intervention and does not directly reduce the contamination load, sampling may be a tool that can be used to detect incoming high levels of contamination that the system may not control, verify effective food safety practices, or monitor for novel risks. The need to mitigate risk and the gap in knowledge of the relative effects of sampling make us ask the following questions:

1. Where in the production chain can testing best manage particular microbial adulterants?
2. What is the effect of food safety testing in the context of systems with different food safety interventions?
3. How can sampling be optimized for different systems, commodities, and hazards?

To answer these questions this project proposed to develop simulations to represent a variety of higher commodities (leafy greens, tomatoes, and cilantro), and incorporate representative sampling programs at different stages of the farm-to-customer process for each commodity; and subsequently, measure the effect of food safety sampling at different stages of the farm-to-facility process on the total adulterant cells (TACs) reaching the endpoint.

Research Methods and Results

This project had three main objectives: (i) to build a Field-to-Facility generic supply chain model of produce safety testing; (ii) to adapt the supply chain and collect parameters to represent a variety of higher-risk commodities with distinct risk profiles and risk-management options; and (iii) to optimize testing across the supply chain of each commodity, incorporating representative testing programs at primary production, harvesting, receiving, processing, and packing, and then assessing their impact to manage safety. For this project, three studies were conducted, one per commodity (leafy greens, tomatoes, and cilantro); each study developed a process model and evaluated sampling plans to meet the project objectives. The methods and results will be categorized by study.

Study 1 Methods: Leafy Greens

Model Overview

A farm-to-customer process model was developed for leafy greens. The initial process flow, contamination scenarios, and sampling plans were obtained by expert elicitation and observation during Salinas, CA, site-visits by PI Matthew Stasiewicz, Gustavo Reyes, and Jorge Quintanilla. The process model, contamination scenarios, and sampling plans were modified/supplemented by using published literature, reports, and industry recommendations. A scenario analysis, **Figure 1**, was developed to represent industry-relevant processing systems, contamination scenarios, and sampling plans. Seven processing systems were developed: one optimal system that incorporated all interventions, one suboptimal system that did not incorporate interventions, and five in-between systems that removed each one of the interventions. Three contamination spreads were evaluated: (i) a random uniform (widespread 100% cluster) event covering the entire field, (ii) a cluster covering 10% of the field, and (iii) a cluster covering 1% of the field. Seven sampling plans from preharvest to the customer were incorporated into the system. The combination of these variables resulted in 147 total scenarios.

Product Flow and Processing Steps:

The initial mass was 100,000 lb of romaine lettuce, chosen to represent a mass reasonably harvested and processed in one day by a grower and packer [11]. The initial mass was split into 50-lb units to represent the mass as it moved throughout the process. These 50-lb units were aggregated, mixed, and partitioned as needed, to represent processing units such as the production rate, pallet load, and finished product packages. Partitioning and mixing processes were modeled, as described by Nauta [12].

Processing consisted of six processing modules: (i) preliminary spray wash, (ii) shredding, (iii) conveyor belt transportation, (iv) flume washing, (v) shaker table, and (vi) dewatering centrifuge as described by Pérez-Rodríguez et al. [13]. Inactivation and cross-contamination were the primary microbial dynamics during processing.

Preharvest Contamination

Contamination was introduced at preharvest in one of three contamination patterns, **Figure 2**, representing the uncertainty around the actual spread of a contamination event. (i) Random uniform (widespread 100% cluster coverage) contamination; this contamination spread covered the whole mass to be harvested and processed, representing an event such as a field irrigated with contaminated water [14, 15] or rainfall that splashed contaminated soil onto the leafy green leaves, leading to the widespread contamination of the field [14]. (ii) Large cluster contamination, where 10% (10,000 lb) of the product was contaminated with an adulterant concentration of 10 CFU/lb. This larger cluster represented a contamination event due to run-off of cattle feces from adjacent farms, contamination from dust containing the adulterant, or other field activities that

may have led to contamination in the form of a large cluster [16, 17]. (iii) Small cluster contamination, where 1% (1,000 lb) of the total mass was highly contaminated (100 CFU/lb). A small cluster contamination event represented contamination due to animal intrusion or due to fecal contamination from wild animals, such as bird droppings that contaminated the leaves or fecal pellets deposited in the field that could be transferred to the leaves due to irrigation or rainfall splash [18-20]. The hazard level was set at 100,000 cells or 1 CFU/lb to provide empiric scenarios. This was based on reverse engineering of the 2018 *E. coli* O157:H7 outbreak associated with romaine lettuce [21], where the average contamination that likely led to the outbreak was determined to be 0.81 CFU/lb.

Production Scenarios and Interventions

Five food safety interventions were adapted to represent interventions in a farm-to-customer system. The interventions were preharvest holding, pre-cooling, pre-wash, chlorinated wash, and processing line sanitation. A detailed description of the interventions is found in **Table 1**.
The goal of the processing systems scenarios was to represent the wide range of uncertainty regarding adherence to food safety practices in different operations. Seven systems were generated to address the wide range of uncertainty. An All-Intervention system, where all five interventions are applied, a No-Intervention system, where no food safety interventions are applied, and five in-between systems were developed, where one of the five food safety interventions was removed one at a time. The five in-between systems are: No Holding, No Washing, No Pre-wash, No Sanitation, and No Pre-cooling. This assesses the benefit of conducting product testing when a specific food safety intervention fails.

Sampling

Seven separate sampling plans were simulated at different processing stages, **Table 2**.

i. Preharvest 4 Days (PHS 4D) sampling occurred 4 days before harvest.
ii. Preharvest 4 Hours (PHS 4H) sampling occurred 4 hours before harvest.
iii. Preharvest Intense (PHS Int) sampling occurred immediately before harvest.
iv. Harvest Sampling (HS) occurred at harvest.
v. Receiving Sampling (RS) occurred after temporary storage at the facility.
vi. Finished Product Sampling (FPS) was performed as the shredded product was packed into 5-lb bags.
vii. Customer Sampling (CS) simulated sampling after transportation from a processing facility to a retail or food service customer.

The sampling plans were designed to match a 1,500-g total composite sample mass and 60 total grabs, **Table 2**. These sampling plan characteristics were selected to observe guidelines from the International Commission on Microbial Specification for Foods (ICMSF) sampling plan stringency (cases); in this case, ICMSF's case 15 for severe hazards for which growth may happen [8]. The sampling plans also adhered to recommendations from Western Growers' (WG) Appendix C, which guides growers and processors on preharvest product testing as specified in the Leafy Greens Marketing Agreement (LGMA) approved guidelines [22]. These documents recommend 60 total individual 25-g grabs to be tested, resulting in a 1,500-g composite mass. All sampling adhered to the 60 grabs, 1,500-g composite recommendations, except for preharvest sampling intense (PHS Int). For PHS Int, the composite sample mass and the number of grabs were increased 4-fold to observe recommendations made by WG's Appendix C under the intensified sampling recommendations. WG recommends that for intensified sampling, the maximum sampling area be 1 acre. Since our model simulated a total mass of 100,000 lb, this translated to approximately 4 acres of romaine lettuce harvested [11]. Therefore, four 1,500-g samples were taken under this scenario, each sample consisting of 60 25-g grabs.

The presence of STEC in an individual grab sample was calculated as in Jongenburger et al. [23] following the low-level heterogeneous contamination assumption since it allowed for obtaining the probability of detection of each grab.

$$P_{detect} = 1 - e^{-C \cdot M_{GrabSample}}, \quad (1)$$

where P_{detect} is the probability of detection, C is the concentration of the adulterant (adulterant cells/g) in the sampling unit (50 lb unit), $M_{GrabSample}$, is the mass of the individual grab sample (g).

Once P_{detect} was obtained, the presence or absence of the adulterant in the given sample was calculated by checking if a random number between 0 and 1 drawn from a uniform distribution was less than P_{detect} , meaning the adulterant is present in the grab sample and detected; otherwise, the adulterant is absent from the grab sample and not detected. At the end of each sampling process, if any of the grabs detected an adulterant cell, the product was rejected as part of an ICMSF 2-class attribute sampling plan [8].

Scenario Analysis Metrics

A total of 147 combinations were generated for analysis. The following metrics were used to assess the performance of both interventions and sampling plans across the scenarios:

Sampling power was a way to measure how well a sampling plan performed at detecting contamination. It is defined by the percentage of times that the sampling plan detected contamination out of the total (n= 10,000) number of iterations.

$$Sampling\ Power\ = \frac{Model\ Iterations\ where\ sampling\ plan\ detected\ pathogen}{Total\ iterations} * 100 \quad (2)$$

Relative Efficacy in reducing endpoint total adulterant cells (TACs) from reaching the system's endpoint was used to quantify how well interventions and sampling plans performed. The endpoint TAC relative efficacy between the 7 sampling plans was quantified across all 7 processing systems and compared to each processing system without sampling.

$$Relative\ Efficacy = 1 - \frac{Endpoint\ (\text{TAC}),\text{what if SCENARIO}}{Endpoint\ (\text{TAC}),\text{SYSTEM NO SAMPLING}} \quad (3)$$

Factor sensitivity analysis (FS) was used to compare interventions and sampling plans. Factor Sensitivity is the log reduction between endpoint TACs from each scenario and the system with no food safety interventions or sampling plans (No-Interventions) for 10,000 iterations. The greater the absolute FS, the greater effect that a specific scenario or condition had on total consumer reduction of endpoint TACs [24, 25].

$$FS\ =\ log \frac{Output(Intervention\ or\ sampling\ plan)}{Output\ (Baseline\ no\ Intervention\ or\ sampling\ plan)} \quad (4)$$

Study 1 Results: Leafy Greens

Individual interventions reduce contamination levels throughout the system; combined effective interventions achieve a greater reduction.

The main objective of this model was to determine the effect of sampling at different stages of the farm-to-customer process, in the context of the system with other intervention strategies: (i) preharvest holding, (ii) pre-cooling at receiving, (iii) pre-wash, (iv) chlorinated wash, and (v) processing line sanitation. Contamination in the system, total adulterant cells (TACs), and progression were tracked for the No-Intervention, All-Intervention systems and individual interventions, **Figure 3.**

When All-Interventions were applied to the system, the total reduction of adulterant cells at the endpoint compared to the No-Intervention system averaged 3.43 log TAC, (3.33–3.56, 95% CI). The two most effective interventions were the chlorinated wash and the pre-wash, which reduced the final TACs in the system by 1.32 log TAC (1.22–1.45, 95% CI) and 1.27 log TAC

(1.18–1.38, 95% CI), respectively. The third most effective intervention was the implementation of holding time at preharvest, providing an average reduction of 0.80 log TAC (0.73–0.90, 95% CI). Interventions that had minor effects were conducting sanitation of the processing lines and pre-cooling, providing an average reduction of 0.06 log TAC (0.02–0.15, 95% CI) and 0.01 log TAC (-0.10–0.11, 95% CI), respectively. The poor efficacy of sanitation was due to the low transfer coefficients between produce and surfaces as well as relatively low contamination during processing; therefore, contamination did not accumulate on the processing surfaces.

The power of sampling plans depends on the contamination levels at each sampling point.

Sampling plan power describes the ability of a given sampling plan to detect contamination at a given processing stage under the sampling conditions stated in **Table 2**. Sampling plan power refers to the number of iterations where the sampling plan was able to detect contamination over the total number of iterations (n= 10,000). Higher power means the sampling plan is more likely to detect contamination at a given processing stage.

The power to detect contamination of the 7 sampling plans (**Table 2**) at the 7 sampling stages and 3 contamination scenarios (147 total combinations) is summarized in **Figure 4**. The results indicate that the most powerful sampling plan was at preharvest sampling 4 days (PHS 4D), with powers ranging between 14.2% and 29.6% higher power observed when the holding intervention was in place. Opposite to all the other sampling plans, the power of the PHS 4D sampling plan increased when effective interventions were implemented, such as the preharvest holding intervention. This intervention limited the contamination window to 2–8 days before harvest, compared to the 0–8 days when it was not in place. With a narrower contamination window, the probability of a contamination event occurring before PHS 4D was higher, and contamination at the sampling point was higher, hence the higher sampling plan power. The second sampling plan with the most power was the preharvest intense (PHS Int) plan ranging between 4.0% and 15.1% due to its higher composite sample mass and the total number of grabs. As the system progressed, the power of sampling plans decreased.

Results showed that power was the lowest when sampling occurred after system processing. Effective processing interventions showed a power between 0.0% and 0.1% for finished product sampling (FPS). On the other hand, if no effective interventions are in place, the sampling plan power was higher (ranging between 7% to 7.3%). Similarly, customer sampling (CS) with effective processing interventions had lower power, between 0.0% and 0.1%, and 3.7% and 4.0% during no effective interventions.

The efficacy of sampling plans is dependent on interventions and sampling location

The relative efficacy achieved by incorporating each sampling plan was quantified for each of the 147 combinations, **Figure 5**. The main pattern suggests that product testing had a lower effect on reducing endpoint TAC when multiple effective interventions or all interventions were in place, except for PHS 4D. PHS 4D had the highest relative effect when the All-Intervention and those systems that included the holding intervention were in place, due to reasons explained in the earlier section. For PHS 4D, the relative efficacy in endpoint TACs ranged between 5% and 25% for systems that included the holding intervention. For those systems that did not include the holding intervention (No Intervention and No Holding), relative efficacy was lower, ranging between 3% to 8%. This result predicts that PHS 4D alone had at most an 8% relative reduction in the endpoint TAC, compared to the higher relative efficacy achieved, 5% to 25% when paired with the holding intervention.

The assessment predicted that product testing had a lower relative effect at reducing endpoint TAC when multiple effective interventions or all interventions were in place for all other sampling plans. The most effective sampling plan across all systems was preharvest sampling intense (PHS Int). PHS Int showed a relative efficacy in endpoint TACs between 68% and 78%

for systems that did not apply the holding intervention, while for those systems affected by the holding intervention, the relative efficacy in endpoint TACs was lower, between 24% and 37%. These results demonstrate that sampling provides greater relative reduction to the endpoint TACs when no effective or no interventions are in place compared to when effective interventions such as holding are in place. A similar pattern can be observed for other sampling plans later in the system. FPS and customer sampling (CS) may be affected by all 5 interventions. When No-Interventions were in place, the relative efficacy in endpoint TACs ranged between 47% and 54% for FPS and 32% to 38% for CS. When All-Interventions were in place, the relative efficacy in endpoint TACs was between 0% and 1%, predicting that sampling at these later stages provided negligible effects when the optimal system was in place. Similarly, for the other systems, when one intervention failed at a time, FPS and CS did not provide much value; the relative efficacies ranged between 0% and 3%. Therefore, for effective systems, it is more beneficial to conduct sampling earlier in the system rather than after multiple reduction steps occur.

Factor sensitivity shows that for optimal system sampling before processing interventions leads to greater TAC reductions compared to sampling after processing.

Factor sensitivity (FS) analysis, **Figure 6**, assessed the effect of sampling plans across each of the seven what-if processing systems. The FS provides information on which sampling plans would most efficiently reduce endpoint TAC if a specific intervention were to fail, as well as on how removing interventions decreases the efficacy of the All-Intervention system.

The most effective sampling plan across all systems was the preharvest intense (PHS Int) with added reductions between 0.13 and 0.66 log endpoint TAC. This is represented by the distance between the black line and the end of the bar in **Figure 6**. For the No-Intervention system, the second and third most effective sampling plans were the FPS and receiving sampling (RS) with added reduction of 0.26 log and 0.17, respectively. For the All-Intervention system, the second and third most effective plans had limited effects, the PHS 4D and RS, had an added log reduction of 0.053 and 0.054, respectively.

For the No Sanitation, No Pre-cooling, No Wash, and No Pre-wash systems, the second and third most effective sampling locations were the PHS 4D followed by preharvest sampling 4 hours (PHS 4H) with reductions between 0.084 to 0.11 log, and 0.054 to 0.092 log, respectively. For the No Holding system, the second and third most effective sampling plans were RS, and harvest sampling (HS) with reductions of 0.34 and 0.30, respectively.

In addition, for all the systems except for the No-Intervention system, the least effective sampling plans were FPS and CS with added reductions ranging between 0 to 0.037, and 0 to 0.031 log, respectively.

Study 2 Methods: Tomatoes

Process Model Development

A site visit to Immokalee, FL, was conducted by PI Stasiewicz, and graduate students Gustavo Reyes and Jiaying Wu, in November 2021 to understand the tomato farm-to-packinghouse process. The tomato farm-to-facility sampling model was developed in Python 3.9.12 [26]. The model simulates *Salmonella* microbial dynamics at different stages of the farm-to-packinghouse process, where contamination can enter the field based on specific contamination scenarios that will be described in further detail later in this document. Contamination can increase or decrease as a factor of processes such as transportation, in-field survival, and flume washing. Cross-contamination can occur between tomatoes, processing equipment, and the flume tank water. Ultimately contamination can be removed from the system by implementing sampling plans that reject the product if the pathogen is detected. An illustration of the model stages, process framework, and microbial dynamics that occur at each stage is found in **Figure 7**, where each

box represents a different module of the model. Modules are linked to each other, where contamination from one module feeds into the next one. Throughout the process, total adulterant cells (TACs) in the system and the proportion of contaminated tomatoes were tracked.

The tomato farm-to-packinghouse sampling model focuses on simulating one field of a medium/larger variety of tomatoes, such as round tomatoes, where the field is harvested 3 times per season. The season is 42 days long, and every 14 days the field is harvested. For the days the field is harvested, the packinghouse operations were simulated. At the end of the processes, the TAC at the system endpoint, and the power of the sampling plans were collected as outputs. Product testing was simulated at four process stages: (i) preharvest, (ii) harvest, (iii) receiving, and (iv) packed product. At each of these stages, product testing follows specific sampling plan characteristics; these sampling plans will be discussed in detail later.

Field Setup

The simulated field is a field that yields approximately 132,000 lb of tomatoes over 42 days; this equates to a total of 230,492 (260-g) individual tomatoes. As mentioned, the field will be harvested on 3 occasions, with each harvest yielding approximately a third of the total yield. Each tomato is assigned a (i) plant number and a (ii) pick number, as well as (iii) an initial contamination level at field creation. The total mass was split into individual tomatoes; each tomato weighs approximately 260 g. Splitting the total mass into individual tomatoes allowed us to track the contaminated sample and apply microbial dynamics to each tomato independently when needed.

Contamination Event

A contamination event introduces contamination into the system. Four contamination patterns were modeled: (i) a random uniform (widespread 100% cluster), 100% of the total area; (ii) a large cluster 10% of the total area; (iii) a small cluster, 1% of the total area; and (iv) a very small cluster, 0.1% of the total area. These four contamination spreads were designed to address the uncertainty around the contamination spread of the external source that may contaminate the product; a description of the contamination scenario is found in **Table 4**. Fecal contamination of tomatoes may be caused by direct contact with the fruit, contaminated water, or soil. Some contamination events may include irrigation with contaminated water, soil splash from rain, contaminated pesticides/fertilizers, animal intrusion, and an employee, among other environmental sources [27-29].

A contamination level of 1 CFU/lb was modeled. This initial contamination was chosen based on previous work done for leafy greens (1); 132,000 cells were spread across the field. This contamination level is representative of a previous foodborne disease outbreak. Since tomato is a multiple-harvest crop, a harvesting season was defined to consist of 3 picks, with 14 days in between picks. A contamination event occurs once per harvesting season. The contamination event occurs on a random day, between days 1 and 42.

In addition to the four contamination scenarios described previously, two additional contamination scenarios were included in the analysis. These two additional analyses were brought up in the Illinois team's site visit in November 2021. The additional contamination scenarios are more specific to address problems that may happen in traditional tomato production: (i) a harvesting bucket/harvester contamination that creates small clusters as a harvesting bucket/harvester moves through the harvested mass; and (ii) a harvesting bin, one 1,000-lb cluster is created at harvest, to simulate the contents of a contaminated harvesting bin. These two additional scenarios are found in **Table 5**.

Sampling Plans

Product testing was modeled to take place in four stages: (i) Preharvest: 3 days before every pick; (ii) harvests, at every pick; (iii) receiving, at packinghouse receiving; and (iv) Packed product, when the product is packed at the packinghouse.

Specific guidance for product testing of tomatoes is currently not available. Therefore, to determine the overall composite sample mass for the sampling plans, we considered the International Commission on Microbial Specification for Foods (ICMSF) attribute sampling plans. ICMSF recommends taking different numbers (n) of 25-g samples, depending on the severity of the hazard. In the case of this model, the hazard is *Salmonella*, which is a severe hazard by ICMSF. ICMSF recommends a 2-class attribute plan for serious hazards under cases 10, 11, and 12. Since tomatoes are consumed raw, and contamination may increase if conditions are appropriate, class 12 is an appropriate initial choice. Case 12 recommends taking 20 x 25-g samples; this is a composite 500-g sample. Although case 12 is recommended for *Salmonella* in produce, other commodities, such as leafy greens, are shifting to performing more rigorous sampling under case 15. Case 15 recommends taking 60 x 25-g samples, this adds up to a 1,500-g composite sample. For this study, we will attempt to match tomato product testing to comply with cases 12 and 15 total mass and grabs requirements.

A challenge for tomatoes as a food matrix is that they are a whole fruit commodity, meaning that obtaining a grab sample from a tomato in the field may not be logistically feasible. The current Food and Drug Administration's (FDA) Bacteriological Analytical Manual (BAM) Chapter 5 protocols for S*almonella* suggest whole tomatoes be classified as whole fruit [30]. For the preharvest, harvest, receiving, packed product, and repacked product processing stages, tomatoes remain as a whole fruit commodity, so taking 25 g of a whole tomato is not feasible. Since taking 25 g is not feasible, different sampling methods will need to be evaluated to match the ICMSF recommendation. Thus, we will evaluate four sampling plans that sample the fruit as a whole and two additional sampling plans that create a tomato mash that is later subsampled. A description of the sampling plans is found in **Table 6**.

Scenario Analysis

A scenario analysis was developed to assess the effect that (i) different contamination spreads, (ii) sampling at different sampling locations, and (iii) sampling plan design have on sampling plan power and the total adulterant cells (TACs) that reach the system endpoint. A total of 144 combinations were obtained from the scenario analysis. The framework of the scenario analysis is found in **Figure 8**.

Performance metrics were developed to assess and compare each combination of contamination spreads, sampling locations, and sampling plans. The performance metrics that will be used for the results are the following:

Sampling Plan Power: Refers to the number of iterations where the sampling plan detected the adulterant over the total number of iterations simulated. A sampling power of 100% means that the sampling plan detected contamination in every iteration. 0% sampling plan power means the sampling plan could not detect contamination in any run. This will address the power that sampling plans have at detecting an uncertain contamination event that could happen anytime during the harvesting season. So, in some cases, if contamination is not present the sampling plan will have a power of 0% since then that sampling plan could not detect a certain contamination event that would happen later in the season.

Endpoint TAC relative efficacy: Every sampling location and sampling plan combination will be normalized to conducting no sampling. A relative efficacy closer to 100% means that the sampling plan did an excellent job of reducing cells from reaching the endpoint. Relative efficacy of 0% means that the sampling plan did not provide any gains in reducing endpoint TAC.

Study 2 Results: Tomatoes

Mass, the number of grabs, and the processing stage drive power of sampling plans.

The power analysis suggests multiple findings (**Figure 9**). The first finding is that sampling plan power is highly dependent on the degree of clustering, as the size of the cluster decreases (a smaller percentage of the total mass is contaminated) the power of the sampling plan will decrease. For the random uniform (widespread 100% cluster) scenario, the most powerful sampling plans were Receiving-60 tomatoes and Harvest-60 tomatoes, with sampling plan powers of 0.30 (0.23-0.37, 5^{th}-95^{th} quantiles) and 0.30 (0.23-0.37, 5^{th}-95^{th} quantiles), respectively. For the 10% cluster, the power was lower than for the random uniform (widespread 100% cluster) contamination, the most powerful sampling plans were Receiving-60 tomatoes and Harvest-60 tomatoes, with sampling plan powers of 0.26 (0.19-0.32, 5^{th}-95^{th} quantiles) and 0.26 (0.19-0.32, 5^{th}-95^{th} quantiles), respectively. The two most effective sampling plans for the 1% cluster were Receiving-60 tomatoes and Harvest-60 tomatoes, with powers of 0.15 (0.08-0.21, 5^{th}-95^{th} quantiles) and 0.15 (0.08-0.20, 5^{th}-95^{th} quantiles), respectively. For the 0.1% cluster, the most powerful sampling plans were also the Harvest-60 tomatoes and the Receiving-60 tomatoes, with powers of 0.029 (0-0.06, 5^{th}-95^{th} quantiles) and 0.028 (0-0.06, 5^{th}-95^{th} quantiles), respectively. These results indicate that the Harvest and Receiving process stages are the best to conduct sampling regardless of the spread of contamination. Sampling plan power highly depends on the contamination event's spread (clustering level).

Results also indicate that the sampling plan power depends on the contamination event's spread. For larger contamination spreads (e.g., 100% and 10% clusters), sampling plans with higher composite sample mass (e.g., 60 and 20 tomatoes) had the greatest power for each sampling location. Conversely, for smaller contamination spread (e.g., 1% and 0.1%), sampling plans that collected more tomatoes from the field (e.g., 60 tomatoes and 60-tomato mash) were the most powerful for each sampling location. This is demonstrated in the Harvest stage, which was one of the most powerful sampling stages. For the random uniform (widespread 100% cluster), harvest sampling plans with higher composite sample mass, 60 and 20 tomatoes had the greatest power, 0.30 (0.23-037, 5^{th}-95^{th} quantiles) and 0.23 (0.17-0.30, 5^{th}-95^{th} quantiles), respectively. The same order is shared with the 10% cluster. Conversely, the 1% cluster and the 0.1% show that the number of tomatoes sampled drives power, the 60-tomato sampling plan had the power of 0.15 (0.08-0.20 5^{th}-95^{th} quantiles) and 0.029 (0-0.06, 5^{th}-95^{th} quantiles), respectively. For the 60-tomato mash, the powers were 0.10 (0.05-0.15 5^{th}-95^{th} quantiles) and 0.024 (0-0.06, 5^{th}-95^{th} quantiles) for the 1% and 0.1% spread, respectively.

The final power results indicate that sampling plan power decreases meaningfully after the packinghouse processes occur. During the packinghouse process reduction and cross-contamination, microbial dynamics may occur; dynamics like reduction make contamination harder to detect at later stages, while cross-contamination if contamination is highly concentrated, may make contamination easier to detect. For the random uniform (widespread 100% cluster), the 10% cluster, and the 1% cluster power is meaningfully lower for packed product sampling (PPS) compared to effective stages such as harvest and receiving, for the 60-tomato sampling plan, powers being on average 54% to 64% lower at PPS compared to harvest. For the 0.1% cluster, the difference in powers for the 60-tomato sampling plan was not as meaningful, at 12.6%. These results reiterate that sampling at earlier stages before any reduction steps will lead to a higher probability of detecting the hazard if present.

Sampling plan relative efficacy indicates best sampling plan location to mitigate cells reaching the system endpoint.

The relative efficacy for endpoint total adulterant cells (TACs) were used to evaluate the effectiveness of a sampling plan in each of the sampling stages and contamination spreads.

The results observed in **Figure 10** indicate that for the random uniform (widespread 100% cluster) event, the 10% cluster, and the 1% cluster, the best stages to conduct sampling are at Harvest and Receiving. For the random uniform (widespread 100% cluster), the most effective sampling plans were the 60 tomatoes and the 20 tomatoes, with relative efficacy of 98.1% at Receiving and 97.85% at Harvest for the 60 tomatoes, and 90% at Harvest, and 89% at Receiving for the 20 tomatoes. Less effective sampling plans such as the 2 and 6 tomatoes had relative efficacy ranging between 27.3% and 34%. For the 10% cluster, the most effective sampling plans were the 60 and 20 tomatoes at Harvest and Receiving. Results showed powers ranging between 96.3% to 96.5% for the 60-tomato plan, and 75.9% to 76.6% for the 20-tomato plan. The two most effective locations for the 1% cluster were Harvest and Receiving. However, in contrast to the widespread (100%) and 10% cluster contamination, the most effective plans were the 60-tomato and the 60-tomato mash plans. The 60-tomato plan had relative efficacy of 43.5% at both sampling locations, whereas the 60-tomato mash plan had relative efficacy ranging between 39.6% to 39.9%. The 0.1% cluster of all sampling locations had the lowest efficacy. The most effective plan was the Packed Product Sampling – 60-tomato plan had a 12.3% efficacy.

The efficacy of sampling plans depends on the contamination levels, and the contamination spread at each sampling point. Packed Product Sampling was the stage with the lowest contamination levels for all the contamination spreads. The contamination levels (total adulterant cells) were on average 317.8 TACs (0–1842, 5^{th}–95^{th} percentiles). Whereas for pre-processing stages such as Harvest and Receiving, the contamination levels were 4,486.2 TACs (0–27,692, 5^{th}–95^{th} percentiles), and 4,290.5 TACs (0–26,204, 5^{th}–95^{th} percentiles), respectively. As mentioned, contamination spread also influences the efficacy of sampling plans, the higher number of contaminated tomatoes, the more likely a grab is to take a contaminated tomato. The contamination spread for processes that started with higher contaminated tomatoes such as the random uniform (widespread 100% cluster), 10% cluster, and 1% cluster observed a reduction in the total number of contaminated tomatoes after processing. The percentage of 12 contaminated tomatoes at Receiving (before processing) was 4.89% (0–28.1%, 5^{th}–95^{th} percentiles) for the random uniform (widespread 100% cluster) event, 1.95% (0–9.6%, 5^{th}–95^{th} percentiles) for the 10% cluster contamination event, 0.4% (0–1.02%, 5^{th}–95^{th} percentiles) for the 1% cluster contamination event. After processing the percentage of contaminated tomatoes dropped to 0.4% (0–2.39%) for the random uniform (widespread 100% cluster), 0.3% (0–2.04%) for the 10% cluster, 0.17% (0–0.9%) for the 1% cluster. For the 0.1% cluster contamination event the percentage of contaminated tomatoes remained constant on average, but in the upper quantiles, the contamination spread increased after processing. At receiving the contamination spread was 0.053% (0%–0.11%) compared to a packed product of 0.05% (0%–0.19%).

These results indicate that if we have a robust sampling plan, such as the 60-tomatoes sampling plan, the best location to conduct sampling may be at Harvest or Receiving assuming the contamination spread is greater than 1%. If the grower suspects that contamination may create a very small cluster ~0.1% (approximately 200 tomatoes), the sampling plans analyzed in this document may be underpowered. In addition, due to cross-contamination at processing, the best sampling location is at packed product sampling.

Results from Specific Contamination Scenarios

Two additional contamination scenarios were included to assess the effect of sampling under contamination spreads that could happen at harvest. Harvesting bucket contamination refers to contamination that occurs after the product is harvested; this creates small clusters of contamination. The harvesting bin is a large cluster of contamination that creates a 1,000-lb cluster every time the product is harvested. This type of contamination event is interesting because preharvest and harvest sampling have no power to detect these two spreads. Therefore, the two sampling locations evaluated were receiving and packed product sampling.

Since the contamination event for this specific contamination scenario occurs once per pick, the sampling plan power metric will be excluded from the results since every pick will have a contamination event that occurs at harvest, so there is no uncertainty around when the contamination event will occur. Instead, the relative efficacy on endpoint TAC is used to evaluate the best location and sampling plan for these two specific contamination events, **Figure 11**.

The results indicate that the best location to sample for the harvesting bucket contamination event is packed product. At the packed product location, the 60 tomatoes sampling plan was able to reduce 83% of the contamination from reaching the system endpoint, the second most effective sampling plan was the 20 tomatoes sampling plan at 67% efficacy. For Receiving, the best sampling plan was the 60 tomatoes, with 80% relative efficacy; the second-best sampling plan was the 60-tomato mash, with 79% relative efficacy. For the Bin contamination, the best sampling location is at receiving, with the 60-tomato and the 60-tomato mash being the two most effective sampling plans at 83% and 82% relative efficacy, respectively. Contrasting the sampling plans with lower power, we can observe that these sampling plans perform better for the harvesting bucket contamination event than the harvesting bin, suggesting that sampling plans with lower efficacy better detect small clusters distributed throughout the total mass compared to a single larger cluster.

Study 3 Methods: Cilantro

A simulation was built in Python 3.9.12 to represent *Cyclospora cayetanensis* water and produce testing. The goals were to (i) identify and fit a model for both water and produce testing using available literature to predict the probability of detection based on the contamination levels and the total number of samples; and (ii) develop a preharvest-harvest process model that assesses water and product sampling, given two contamination scenarios (low and high), two contamination frequencies and 12 sampling plans. This will be described in detail later in this document.

Fitting models for agricultural water testing and product testing

Data from Durigan et al. [31] and FDA BAM Chapter 19C [32] was used to fit a logistic regression model to assess current water detection methods for *C. cayetanensis*. The FDA conducted a study to assess the efficacy of a dead-end ultrafiltration method for *Cyclospora* detection, in which they tested water samples seeded with 200, 100, 25, 12, 6 oocysts per 10-liter sample.

Data from Murphy et al. [34] and FDA BAM Chapter 19B [35] was used to fit a logistic regression model to assess current produce testing detection methods for *C. cayetanensis*. The study was conducted to assess the efficacy of the fresh produce sampling method for *Cyclospora cayetanensis* detection, in which they tested different cilantro samples seeded with 200, 10, 5, and 0 oocysts per 25-g samples.

The logistic regression model was fitted with the results (0 or 1) as the response variable and the contamination levels (oocyst per 10 liters or oocyst per 25 g) as the predictor. **Table 7** compares empirical data and the fitted data obtained from the models. The fit of the model was evaluated. The deviance between the null model and the model with the predictor was used to calculate the p-value using a chi-squared distribution. The p-value for both models was ($p < 0.001$), indicating goodness of fit for the logistic regression models. The formulas for the logistic regression model are below:

$$Pdetect_{produce} = \frac{e^{-3.419 + 0.454 \times Oo_{Psample}}}{1+e^{-3.419 + 0.454 \times Oo_{Psample}}} \quad (1)$$

$$Pdetect_{Water} = \frac{e^{-4.488 + 0.856 \times Oo_{Wsample}}}{1+e^{-4.488 + 0.856 \times Oo_{Wsample}}} \quad (2)$$

where $Pdetect$ is the probability of detection, $Oo_{Wsample}$ is the concentration of oocysts in the water sample (Oocysts per 10L), and $Oo_{Psample}$ is the concentration of oocysts in the produce sample (Oocysts per 25g).

Simulation framework to compare contamination levels and probability of detection

The contamination module simulated the contamination of bulk water, or a field irrigated with contaminated bulk water. *C. cayetanensis* oocysts were modeled as discrete variables (whole numbers) to correctly represent individual oocysts in bulk water or on the product as they were sampled.

To simulate sampling, a number of 25-g or 10-L samples were taken from the cilantro field or the bulk water. The number of samples was selected to be 1, 2, 4, 8, 16, and 32 to observe the relationship between doubling the number of samples and the probability of detection.

The system of equations used to simulate sampling is in **Table 8.** The sampling system of equations is made of 4 main steps.

1. First, it calculates the probability of collecting oocysts from bulk water ($\Pr_{OoSW}$) or the cilantro field ($\Pr_{OoSP}$). For bulk water, the ratio between the volume of the sample and the volume of bulk water was used as the probability of collecting *C. cayetanensis* oocysts from bulk water into the sample. For produce sampling, samples were taken using simple random sampling. For each sample, the probability of collecting oocysts into the sample was calculated as the ratio between weight of the sample (g) and weight of the partition (g).
2. The second step was the quantification of the total oocysts in the sample by using a binomial distribution. For water, the oocysts in bulk water were used as the number of attempts, and the probability of a sample having oocysts was used as the probability of success. For produce, the oocysts in the sampled partition were used as the number of attempts, and the probability of a sample having oocysts was used as the probability of success.
3. Third, the probability of detection $\Pr_{PCRDetect}$ was predicted. If 0 oocysts were collected in the sample, the probability of detection was chosen as 0 (due to a 0 false positive rate from the empirical data, as well as other studies that have evaluated the specificity of these protocols [36]). If the oocysts in the sample were ≥1, then the logistic fits from equations 1 and 2 were used to quantify the probability of detection.
4. Fourth, the test of the outcome was quantified as detected (1) or not detected (0), by randomizing a uniform distribution. If the randomized value was lower than the probability of detection *C. cayetanensis* oocysts were detected.

To calculate the probability of detection, this model framework is nested within a two-layer iteration process. In the first layer, the contamination levels in the field or the water are fixed, while contamination and the sampling modules are randomized based on the previously shown equations. This allows for the probability of acceptance to be defined as follows:

$$P_d = \frac{\sum_{i=1}^{nj} I_a}{n_1} \quad (3)$$

where P_d is the probability of detection, I_a is a binary indicator of detected (1) and not detected (0), n_1 is the number of iterations in the inner layer (n=100).

The outer layer reiterates the inner layer for k, (k = 100), thus allowing to obtain a vector of P_ds, which are used to calculate the mean, median, and variability intervals of the probability of detection as a total of 10,000 iterations.

Preharvest-Harvest Simulation Scenario Analysis

A 45-day simulation was developed to simulate a cilantro season of production. Parameters for the growing season water and fresh produce sampling were based on FDA recommendations for parasitic animals in foods and other available literature (**Table 9)**. A scenario analysis was developed to compare: (i) two contamination levels – high contamination levels, where irrigation water is contaminated with 20 oocysts per liter, and low contamination, where irrigation water is contaminated with 0.6 oocysts per liter; (ii) two contamination frequencies – frequency scenario 1, where irrigation water is simulated to be contaminated at preharvest every day of the harvest season, and frequency scenario 2, where irrigation water is contaminated randomly once during the harvest season (sporadic); and (iii) 12 agricultural water or produce testing sampling plans to assess the effect of different sampling frequencies and times along the harvest season. The 12 different sampling scenarios during harvest and preharvest points were simulated (10,000 iterations) and compared to a baseline scenario (no sampling of *C. cayetanensis*) to assess the efficacy of water and fresh produce sampling plans (**Table 10**). The relative efficacy of each scenario was calculated as follows:

$$\text{Relative efficacy} = 1 - \frac{Endpoint\ Oocyst\ Scenario}{Endpoint\ Oocyst\ Baseline} \quad (5)$$

Similar to other studies, the relative efficacy represents the benefit of a sampling plan on oocysts that reach the system endpoint compared to not conducting sampling at all. Where 100% means that the sampling plans prevented all oocysts from being harvested, while 0% means that no benefit was obtained compared to not conducting sampling.

Study 3 Results: Cilantro

Sampling Simulations

The probability of detection of agricultural water testing compared to the total concentration of *C. cayetanensis* (Oocysts per liter), stratified by the number of 10-L samples, is shown in **Figure 12A**. A higher number of samples yields a higher probability of detection only when contamination is lower than 1.58 oocysts per liter. When contamination levels are above 1.58 oocysts per liter, the probability of acceptance reaches 100%, meaning that only one 10-L sample is needed to reliably detect contamination.

The benefit of taking more samples for different contamination levels is found in **Figure 12B**. The difference in the median probability of detection between taking (n = 2, 4, 8, 16, and 32) and taking (n = 1) 10-L samples was plotted. As the number of 10-L samples increases, the difference in the probability of detection increases. Similarly, as the number of 10-L samples increases, taking more samples will also be able to detect contamination more effectively at lower contamination levels. As the contamination levels approach the detection limit (1.58 oocysts per liter), the overall benefit of taking more samples decreases.

The probability of detection of product testing compared to the total contamination levels in the field (oocysts per gram), and the total number of 25-g samples are shown in **Figure 13A**. Produce sampling has a negligible effect as in-field contamination levels approach 0 oocysts per gram. When contamination reaches 0.92 (oocysts per gram), all samples have between 98% to 100% chance of detecting the hazard. As the number of samples increases, produce testing can detect lower contamination levels in the field.

The difference in the median probability of detection between taking (n = 2, 4, 8, 16, and 32) and taking (n = 1) samples are plotted in **Figure 13B**. Like water testing, this plot shows the added benefit of taking more samples than taking one sample. As the number of samples increases, the overall benefit of the probability of detection increases. Taking more samples means that lower levels of contamination can be detected. However, as contamination levels get

closer to 0.92 (oocysts per gram), the benefit of taking more samples decreases until all samples can detect contamination.

Scenario Analysis Simulations

Results of the frequency scenarios, continuous irrigation with contaminated water, and one random irrigation contamination event with contaminated water are shown in **Figure 14**.

For the contamination frequency scenario 1 (daily irrigation with contaminated water), **Figure 14A**, the results demonstrate that when contamination levels in irrigation water were high (20 oocysts per liter), most scenarios showed high relative efficacy, ranging between 99.6% and 100%. The only plan that did not have high efficacy was Produce Testing 1 (produce testing on the first day of the season), where the relative efficacy was 2.5%. When contamination levels in irrigation water were low (0.6 oocysts per liter), produce testing showed the lowest relative efficacy across all scenarios, ranging between 0.12% to 4.7%, with higher efficacy later in the season. Water testing alone had higher relative efficacy ranging between 57.4% and 92.9%. Water and produce testing together had efficacies ranging between 58.4% and 92.6%. The highly effective sampling plans were daily water testing, and daily water and produce testing, with efficacies of 100%. Daily produce testing showed a relative efficacy of 54.3%.

Figure 14B shows the relative efficacy of sampling plans under the contamination frequency scenario 2 (irrigation water is randomly contaminated one day per season). The results show low relative efficacy across the board. At high contamination levels (20 oocysts per liter), the relative efficacy of *C. cayetanensis* detection was high for daily produce testing (34%), daily water testing (100%), and daily water and produce testing (100%). The remaining sampling plans ranged between 0.02% to 6.9% relative efficacy. When contamination levels were low (0.6 oocysts per liter), the most effective sampling plans were daily water testing (57.8%), and daily water and produce testing (58.15%). The rest of the sampling plans showed low efficacy ranging between 0% and 5.8%.

Data Availability for Studies

The code and supplemental documents are found in the following repositories for each of the three studies.
Study 1: Leafy Greens: https://github.com/foodsafetylab/Farm-to-Consumer-LG-Sim
Study 2: Tomatoes: https://github.com/foodsafetylab/Tomato-Sampling-Model
Study 3: Cilantro: https://github.com/foodsafetylab/Cilantro-Sampling-Model

Outcomes and Accomplishments

Study 1: Leafy Greens

- This study developed a farm-to-customer process model for leafy greens to address gaps in knowledge of the effect of product sampling paired with other food safety interventions.
- A scenario analysis was developed to assess 7 processing systems, 7 sampling plans, and 3 contamination patterns (147 scenarios). The scenario analysis was used to:
  (i) Compare the effect of individual interventions at reducing total endpoint adulterant cells (TACs).
  (ii) Assess the power and relative efficacy of the 7 sampling plans at detecting contamination.
  (iii) Compare the effect of sampling plans across different systems to find the best sampling plan for each system.
- The scenario analysis showed that sampling before effective processing interventions (preharvest, harvest, receiving) yields higher efficacy at reducing TACs than sampling after effective processing interventions occur, at finished product sampling and customer sampling. This result demonstrates that sampling can be used as a tool to detect a high level of incoming contamination, whereas sampling after effective interventions occur had negligible effects.

Study 2: Tomatoes

- The goal of this study was to identify the best sampling location and sampling plans for the detection of preharvest contamination in tomatoes. A farm-to-packinghouse simulation was developed for tomatoes. The simulation assessed 4 contamination spreads, 4 sampling locations, and 6 different sampling plans.
- The model demonstrates that when the initial contamination is relative, random uniform (widespread 100% cluster), in a 10% or 1% cluster, then sampling at harvest and receiving are the best sampling locations. When initial contamination is in a very concentrated cluster, 0.1%, the best sampling location is at the packed product.
- The model shows that for the random uniform (widespread 100% cluster) and the 10% cluster contaminations, the best sampling plans are those that test the larger amount of product (60 and 20 tomatoes). When contamination is highly clustered, in the 1% or 0.1% clusters, the best sampling plans are those that sample more individual tomatoes (60 tomatoes, and 60 tomato mash).

Study 3: Cilantro

- This study developed a generic model that utilizes current literature based on the U.S. Food and Drug Administration (FDA) recommended sampling methods for *Cyclospora cayetanensis* testing in cilantro to assess the efficacy of these methods in a harvest season and inform readers about the effectiveness of these sampling techniques.
- The project model simulated the probability of detection of water testing and produce testing in cilantro, based on FDA BAM Chapter 19B and 19C protocols. For water testing, the contamination simulation showed that when contamination is above 1.58 oocysts/liter, a single 10-L water sample will reliably detect contamination. For produce testing, results showed that above 0.92 oocysts/gram, a single test will reliably detect contamination.
- The project used the model to simulate and evaluate different *C. cayetanensis* sampling scenarios in water and fresh produce testing based on two contamination scenarios: continuous irrigation of contaminated water during harvest, and irrigation of contaminated water 1 time (randomly) during harvest season.

- For continuous irrigation with contaminated water, almost all testing scenarios showed a high relative efficacy (100%) at detecting *C. cayetanensis* for high contamination levels, except for produce testing at the beginning of the season. At lower contamination levels, relative efficacy was lower with produce (3%–12%), water (57%–93%), and water and produce (58%–93%).
- For irrigation with contaminated water 1 time (randomly), at high and low contamination, daily water testing (57.8%–100%), daily produce testing (58%–100%), and daily water and produce testing (100%) showed high efficacy of detection. Relative efficacy decreased for water and produce testing occurring once per season (0%–7%).

Summary of Findings and Recommendations

Study 1: Leafy Greens

This study indicates that sampling is less impactful at reducing the endpoint adulterant cells when effective systems-based interventions are in place. The model showed that the sampling plans in this study have limited power to detect contamination levels like the ones that caused a foodborne disease outbreak in 2018. The model showed that conducting sampling too early in the system may not be as beneficial compared to sampling closer to a contamination event, as shown by the Holding intervention. Sampling plans should focus on locations where contamination is likely to be high enough for powerful sampling, at preharvest, harvest, and receiving, as this would allow sampling to detect the incoming product that presents high levels of contamination. Effective interventions reduce contamination during processing, making finished product sampling and customer sampling have negligible effects. In addition, finished product sampling could better inform a contamination event if all interventions were to fail. This study suggests that interventions are effective at reducing incoming contamination. Producers and buyers should focus on implementing suitable food safety interventions as primary preventive controls. Once interventions are implemented, sampling plans can be used as a tool to detect high-level contamination or unappreciated/untreated sources of contamination.

Study 2: Tomatoes

This study builds upon the findings from the study of leafy greens. This study aimed to identify possible sampling locations for tomatoes along the farm-to-packinghouse process. In addition, we also wanted to implement sampling plans based on scientific recommendations that would be feasible for growers to implement. The findings suggest that the best sampling locations are at harvest and receiving, and this aligns with the findings for the leafy green study. Locations where contamination levels are predicted to be present consistently and at higher levels are the sampling locations that result in the higher detection power. When the clustering level is higher (0.1% cluster) the best sampling location was predicted to be at packed product, showing that when contamination is high and highly clustered, the processing steps provide enough cross-contamination to make detection more powerful towards the end of the system. However, the model shows that for the 0.15% cluster contamination event, relative efficacy was a maximum of 12.3% compared to the 98.1% reduction achieved when contamination was random uniform (widespread 100% cluster). The model shows that for random uniform (widespread 100%) and the 10% cluster the best sampling plans are those that test the larger amount of product (60 and 20 tomatoes), while when contamination is highly clustered (1% or 0.1% cluster), the best sampling plans are those that sample more individual tomatoes (60 tomatoes, and 60 tomato mash). While some very large sampling plans (60 tomatoes and 60 tomato mash) typically showed large reductions in cells reaching the system endpoint, these specific plans are likely too large to be practical. More feasible plans like the 2 and 6 tomatoes sampling plans, which match

total mass to 375 and 1,500 g masses collected with ICMSF plans, are underpowered, especially under highly clustered scenarios.

Study 3: Cilantro

Simulation results of FDA BAM Chapter 19B and 19C protocols showed that the number of samples matters for *C. cayetanensis* detection in cilantro. This study identified important contamination thresholds for *C. cayetanensis* detection. At levels above 1.58 oocysts per liter for water testing, and 0.92 oocysts per gram for produce testing, a single 10-L or 25-g sample will reliably detect *C. cayetanensis* contamination. This model provides the industry with information on the performance of the current testing methods for *C. cayetanensis* under different contamination levels, and testing scenarios.

Similar results were observed after scenario analysis. Results showed that increased number of samples and daily testing have higher efficacy at detecting *C. cayetanensis* during a harvest season, whereas a single sample during the harvest season will have a hard time detecting a random contamination event.

Overall analysis showed that testing both water and produce has a higher efficacy at detecting *C. cayetanensis* than only testing for water or produce. This is more evident when random irrigation with contaminated water occurs. Based on the simulation results, it would be ideal to increase the number of samples for both water and produce testing to increase the detection of *C. cayetanensis,* especially when lower contamination is expected. However, expenses associated with *C. cayetanensis* testing will limit the number of samples required for improved detection.

APPENDICES

Publications and Presentations

Publications:

- A manuscript discussing the finding of study 1: leafy greens is published.
    - Reyes, G.A., Quintanilla Portillo, J. & Stasiewicz, M.J. 2023. Leafy green farm-to-consumer process model predicts product testing is most effective at detecting contamination when conducted early in the system before effective interventions. *Applied and Environmental Microbiology* 89: e00347-23. https://journals.asm.org/doi/full/10.1128/aem.00347-23
- A research note for study 2: tomatoes is currently under development (as of July 2023). Target journal, Journal of Food Protection.
- A manuscript for study 3: cilantro has been submitted and is under revision (as of July 2023). Target journal, Journal of Food Protection

Presentations:

Reyes, G.A., Wu, J. & Stasiewicz, M.J. Evaluating Product Testing Combined with Other Strategies for Reducing Risks from Pre-Harvest Contamination of *E. coli* O157:H7 on Generic Leafy-Green Produce Using a Farm-to-Facility Simulation. *Presented at:* 2022 International Association for Food Protection Annual Meeting (Technical Session) Id: T1-10.

Reyes, G.A. & Stasiewicz, M.J. Evaluating the effectiveness of sampling plans and locations in multi-harvest commodities through the development of a farm-to-packinghouse simulation for tomatoes. 2023 International Association for Food Protection Annual Meeting. ABSTRACT SUBMITTED

Chavez, R.A., Reyes, G.A. & Stasiewicz, M.J. Development of cilantro pre-harvest and harvest model for *Cyclospora cayetanensis* testing. 2023 International Association for Food Protection Annual Meeting. ABSTRACT SUBMITTED

Budget Summary

A total of $222,598 was awarded to this project. We did have the necessary funds to fully implement the project. Some of the remaining funds will be used for publication and travel costs, but not all funds will be spent. We will complete the paper writing before the 2023 CPS Symposium and attend that symposium as planned.

Tables 1–10 and Figures 1–14 (see below)

Study 1 Tables, Leafy Greens:

Table 1: *Description of the Interventions used to represent good food safety practices.*

Intervention name	Description
Holding time	The product is harvested 2-8 days after the contamination event, in contrast to the No-Intervention system conditions where the product can be harvested 0-8 days after the contamination event.
Pre-cooling	At receiving, the product is rapidly cooled to 3-5 °C, the effect of pre-cooling reduces the temperature quickly, potentially reducing the growth of adulterants that may have been exposed to elevated temperature during transportation between the farm to the facility.
Spray Pre-Wash	As the first step of the production process, the produce is sprayed with a sanitizing solution that is predicted to reduce the microbial load with variable efficacy between (1.1 – 1.46 log adulterant cells).
Chlorinated Wash	Chlorinated wash is applied to produce after shredding, this reduces the degree of cross-contamination as well as reduction of adulterants. One of two wash systems is selected if washing is selected as an intervention. (i) the state-of-the-art washing system that constantly maintains FC levels inside the flume tank at 10 ppm or (ii) the washing system with a 2-minute dosing period every 12 minutes.
Processing Line sanitation	The processing facility surfaces are washed every 2500, 5000, or 7500 lb of production, with variable compliance (0, 0.25, 0.50, 0.75, 1) and variable efficacy (-1, -2, -3, -4 log cells).

Table 2**:** *Description of sampling plans for each sampling scenario.*

Sampling Plan	Total Number of Samples	Subsamples description	Total mass sampled (g)	Subsample Mass (g)	Grabs per subsample (#)	Total Grabs (#)	Weight per grab (g)	Rejection Rule
Preharvest 4 days (PHS 4D)	1	4, 375 g subsamples	1,500	375	15	60	25	"Total mass"
Preharvest 4 hours (PHS 4H)	1	4, 375 g subsamples	1,500	375	15	60	25	"Total mass"
Preharvest Sampling Intense (PHS Int)	4	4, 375 g subsamples	6,000	375	15	240	25	"Total mass"
Harvest Sampling (HS)	1	4, 375 g subsamples	1,500	375	15	60	25	"Total mass"
Receiving Sampling (RS)	1	1, 375g sample every 6-7 pallets	1,500	375	15	60	25	"Total mass"
Finished product Sampling (FPS*)	1	1 subsample every 3.5 hr. of production	1,500	375	15	60	25	"Total mass"
End Consumer Sampling (CS)	1	1, 375g sample every 15-16 pallets	1,500	375	15	60	25	"Total mass"

*Based on the day being 14 hours of production as per [11].

Table 3: *Endpoint total adulterant cells (TACs) comparison between the No-Interventions scenarios and the All-Intervention scenarios. Log10 change calculated by calculated as (log10 (All-Intervention/ No-Intervention)). Relative efficacy is calculated as (1 - (Endpoint TAC Interventions/ Endpoint TAC All Intervention)).*

Interventions	Random Uniform (Widespread 100% cluster)			10% Cluster			1% cluster		
	Endpoint TACs	Log Change	Relative efficacy	Endpoint TACs	Log Change	Relative efficacy	Endpoint TACs	Log Change	Relative efficacy
No-Intervention	15,851,671	-*	-	15,841,829	-	-	15,823,153	-	-
All-Intervention	5,939	-3.43	99.96%	6,139	-3.41	99.96%	5,903	-3.43	99.96%
No Holding	40,238	-2.60	99.75%	40,425	-2.59	99.74%	38,815	-2.61	99.75%
No Pre-cooling	6,171	-3.41	99.96%	6,482	-3.39	99.96%	6,618	-3.38	99.96%
No Pre-wash	105,908	-2.18	99.33%	105,674	-2.18	99.33%	105,561	-2.18	99.33%
No Washing	118,506	-2.13	99.25%	118,464	-2.13	99.25%	118,545	-2.13	99.25%
No Sanitation	5,606	-3.45	99.96%	5,668	-3.45	99.96%	5,576	-3.45	99.96%

**No-Intervention is the reference condition, so no change is recorded.*

Study 2 Tables: Tomatoes

Table 4: *Description of contamination scenarios. These address uncertainty around clustering in tomato mass harvested.*

	Random Uniform (Widespread 100% cluster)	10% Cluster*	1% Cluster	0.1% Cluster
Total Cells in System	132,000	132,000	132,000	132,000
Overall Field Concentration	1 CFU/lb	1 CFU/lb	1 CFU/lb	1 CFU/lb
Concentration in Contaminated Cluster Area	1 CFU/lb	10 CFU/ lb	100 CFU/lb	1,000 CFU/lb
Occurrence	Once between 1-42 days Preharvest contamination	Once between 1-42 days Preharvest contamination	Once between 1-42 days Preharvest contamination	Once between 1-42 days Preharvest contamination

***Clusters** are based on the total tomatoes initially in the area, regardless of the pick number. Each tomato is assigned a field location. The contamination event contaminates the field location regardless of how many tomatoes are in the field.

Table 5: *Description of the two scenarios to represent specific contamination spreads. Scenarios developed from a site visit to Florida, USA in November 2021.*

	Harvesting Bucket Contamination	Bin Contamination
Total Cells in System	132,000	132,000
Total Cells per pick	44,000	44,000
Overall Concentration	1 CFU/lb	1 CFU/lb
Concentration in Contaminated Areas	55 CFU/lb	44 CFU/lb
Pattern	25 – 32 lb clusters per pick (0.6%)	1-1,000 lb cluster per pick (2.27%)
	75 – 32 lb clusters total (1.8%)	3-1,000 lb clusters total (6.81%)
Occurrence	Occurs once per pick	Occurs once per pick

Table 6: *Description of sampling plans for whole tomatoes. These 6 sampling plans are applied at 4 stages: Preharvest (PHS), Harvest (HS), Receiving (RS), and Packed Product (PPS).*

	Sampling Plan 1	Sampling Plan 2	Sampling Plan 3	Sampling Plan 4	Sampling Plan 5	Sampling Plan 6
Total Tomatoes	2	6	20	60	20	60
Total mass (g)	520 g	1,560 g	5,200 g	15,600 g	5,200 g	15,600 g
Whole tomato Enrichment	Yes	Yes	Yes	Yes	-	-
Technical Replicates	-	-	-	-	20 x 25 g	60 x 25g
Replicates Mass	-	-	-	-	500 g	1,500 g
ICMSF Rationale	Matches case 12 mass	Matches case 15 mass	Matches case 12 grabs	Matches case 15 grabs	Matches case 12 grabs & use a total of 500g mass for enrichment	Matches case 15 grabs & use a total of 1,500g mass for enrichment

Study 3 Tables: Cilantro

Table 7. *Summary of detection rates from Durigan et al. (2020) and Murphy et al. (2018) and model detection rates from the logistic regression fits. The empirical data was used to fit the detection based on seeding levels, oocyst per 10 L for agricultural water and oocyst per 25 g for produce samples.*

Contamination Oocysts/ 10 L	Data from Durigan et al. (2020) for the detection of *C. cayetanensis* oocysts in agricultural water			Model detection rate (%)
	Sample tested	Positive samples qPCR	Detection rate (%)	
0	12	0	0	0*
1	-	-	-	2.6
2	-	-	-	5.9
3	-	-	-	12.7
4	-	-	-	25.6
5	-	-	-	44.7
6	12	8	66.6	65.6
7	-	-	-	81.8
8	-	-	-	91.4
9	-	-	-	96.1
10	-	-	-	99.9
11	-	-	-	99.9
12	3	3	100	100
13	-	-	-	100
14	-	-	-	100
15	-	-	-	100
16	-	-	-	100
17	-	-	-	100
18	-	-	-	100
19	-	-	-	100
20	-	-	-	100
.	.	.	.	.
25	6	6	100	100
100	3	3	100	100
200	6	6	100	100

Contamination Oocysts/ 25 g	Data from Murphy et al. (2018) for the detection of *C. cayetanensis* oocysts in fresh produce			Model detection rate (%)
	Sample tested	Positive samples nPCR	Detection rate (%)	
0	80	0	0	0*
1	-	-	-	4.9
2	-	-	-	7.8
3	-	-	-	12.3
4	-	-	-	18.9
5	80	27	33.7	27.9
6	-	-	-	39.1
7	-	-	-	51.5
8	-	-	-	63.8
9	-	-	-	74.5
10	80	64	80	82.8
11	-	-	-	88.9
12	-	-	-	93.0
13	-	-	-	95.6
14	-	-	-	97.3
15	-	-	-	98.4
16	-	-	-	99.0
17	-	-	-	99.4
18	-	-	-	99.6
19	-	-	-	99.8
20	-	-	-	99.9
.	.	.	.	.
25	-	-	-	100
100	-	-	-	100
200	80	80	100	100

***** Represents that the model value was changed from 1.1% to 0% for water testing and from 3.3% to 0% for product testing, as the observed false positive rate for the data was 0%.

Table 8. *Equations used to simulate sampling.* Where $\Pr_{OoSW}$ is the probability of collecting *C. cayetanensis* oocysts from bulk water into the water sample. $\Pr_{OoSP}$ is the probability of collecting *C. cayetanensis* oocysts from a sampled field partition into the produce sample. $Oo_{Wsample}$ is the number of *C. cayetanensis* oocysts in the water sample. Oo_{BW} are the total of *C. cayetanensis* oocysts in bulk water. $Oo_{Psample}$ is the number of *C. cayetanensis* oocysts in the produce sample. $Oo_{Portion}$ are the total of *C. cayetanensis* oocysts in the sampled partition of cilantro. $\Pr_{PCRDetect}$ is the probability of detection based on the logistic fits. $DetectYN$ is a Boolean variable, where 1 means *C. cayetanensis* oocyst(s) was detected, and 0 means no detection.

Step No.	Step	Water Sampling	Produce Sampling
1	Probability of collecting oocyst from bulk water or produce field into sample	$\Pr_{OoSW} = \frac{Volume\ (L)\ sample}{Volume\ \ (L) Bulk\ water}$	$\Pr_{OoSP} = \frac{W\ (g)_{sample}}{W\ (g)_{Portion}}$
2	Total oocyst in the 10L sample or the 25g produce sample	$Oo_{Wsample} = binomial\ (\ Oo_{BW}\ , \Pr_{OoSW})$	$Oo_{Psample} = binomial\ \left(Oo_{Portion}, \Pr_{OoSP}\right)$
3	Probability of detection	$\Pr_{PCRDetect} = fit(\ Oo_{Wsample})$	$\Pr_{PCRDetect} = fit(\ Oo_{Psample})$
4	Was *C. cayetanensis* detected?	$DetectYN = \ ifelse\ (Uniform\ (0,1) < \Pr_{PCRDetect}, Detected, Not\ Detected)$	

Table 9. *Parameters used for C. cayetanensis sampling simulation in Cilantro.*

Description	Value	Units	Reference
Farm Size	1	acre	[37]
Seed per Acre	100	seeds	[38]
Bed Size	80	inches	[38]
Irrigation	Once a day	no units	[39]
Total water during harvest period	12	inches	[39]
Mature time of cilantro (season length)	45	days	[40]
Production outcome	22,000	lb	[41] [37]
Cilantro plant weight	1	lb	[40]
Cyclospora growth	no growth	no units	[42]
Water sample volume	10	L	[31]
Product sample weight	25	g	[34]

Table 10. *Contamination scenarios for C. cayetanensis sampling simulation in Cilantro.*

Codes	Scenario	Scenario Description
Contamination Level Scenarios		
CS1	Contamination scenario 1 (high)	Bulk water is contaminated with 20 oocysts per liter
CS2	Contamination scenario 2 (low)	Bulk water is contaminated with 0.6 oocysts per liter
Contamination Frequency Scenarios		
FS1	Frequency Scenario 1	Irrigation of contaminated water at preharvest every day of a harvest season
FS2	Frequency Scenario 2	Random, irrigation of contaminated water at preharvest 1 time (randomly) during the harvest season
Sampling Scenarios		
Daily Water Testing	Daily testing of water	Daily testing of water
Daily Produce Testing	Daily testing of produce	Daily testing of Produce at preharvest
Daily Water & Produce Testing	Daily testing of both (water and produce)	Daily testing of water and daily testing of produce at preharvest
Water & Produce 1	Water and produce testing scenario 1	Produce and water testing at harvest
Water & Produce 2	Water and produce testing scenario 2	Produce and water testing at harvest and 1 time at pre-harvest at the start of the season
Water & Produce 3	Water and produce testing scenario 3	Produce and water testing at harvest and 2 times at pre-harvest (1 at the start of preharvest, and 1 at the day of harvest)
Water Testing 1	Water testing scenario 1	Water testing 1 time per harvest season, on day 1 of preharvest
Water Testing 2	Water testing scenario 2	Water testing 2 times per harvest season on day 1 of preharvest and the day of harvest
Water Testing 3	Water testing scenario 3	Water testing 3 times per season on day 1 of preharvest, at mid-season (day 22), and on the day of harvest
Produce Testing 1	Produce testing scenario 1	Produce testing on the day of harvest
Produce Testing 2	Produce testing scenario 2	Produce testing on day 1 of preharvest and on the day of harvest
Produce Testing 3	Produce testing scenario 3	Produce 2 times at pre-harvest (1 at the start of preseason, at mid of preseason), and on the day of harvest

Study 1 Figures: Leafy Greens

Figure 1: Scenario Analysis Framework for leafy greens

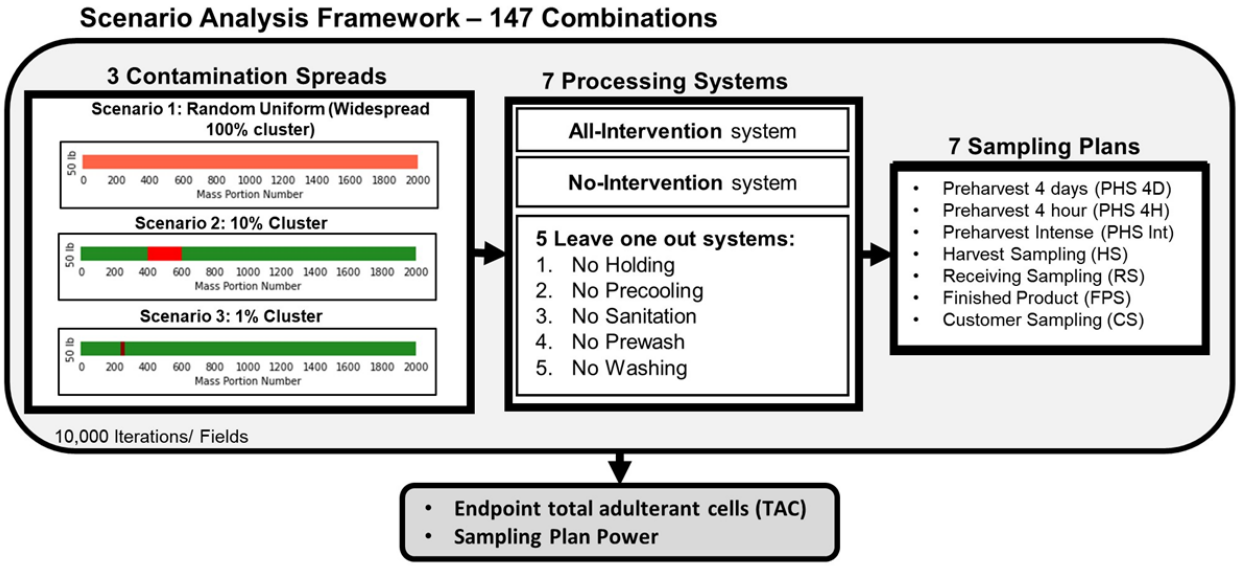
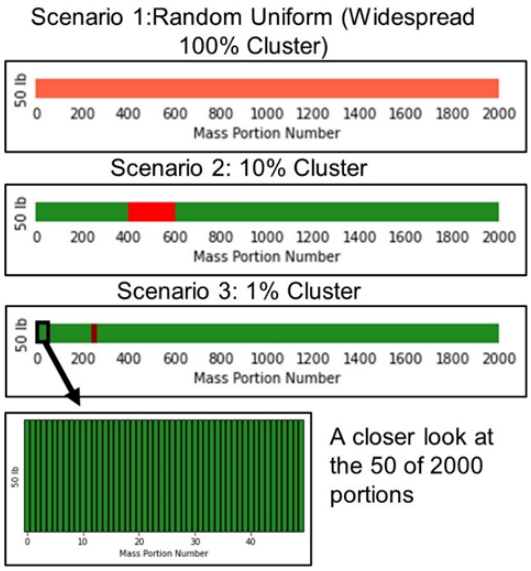


Figure 1: The scope of the scenarios included. 3 contamination patterns, 7 processing systems, and 7 sampling plans. Each was evaluated individually to predict sampling plan power and the effect on total adulterant cells (TACs) that could reach the system endpoint.

Figure 2: Contamination scenarios leafy greens



Total mass Contaminated	Concentration in contaminated mass	Total Adulterant Cells in processing mass
100,000 lb.	1 Adulterant Cell/lb.	100,000 Cells
10,000 lb.	10 Adulterant Cell/lb.	100,000 Cells
1,000 lb.	100 Adulterant Cell/lb.	100,000 Cells

Figure 2: Representation of the 3 contamination scenarios in a 100,000 lb generic mass to be harvested and packed. The red-shaded area represents contaminated mass, whereas the darker shade shading represents a greater concentration of adulterant cells.

Figure 3: Contamination progression

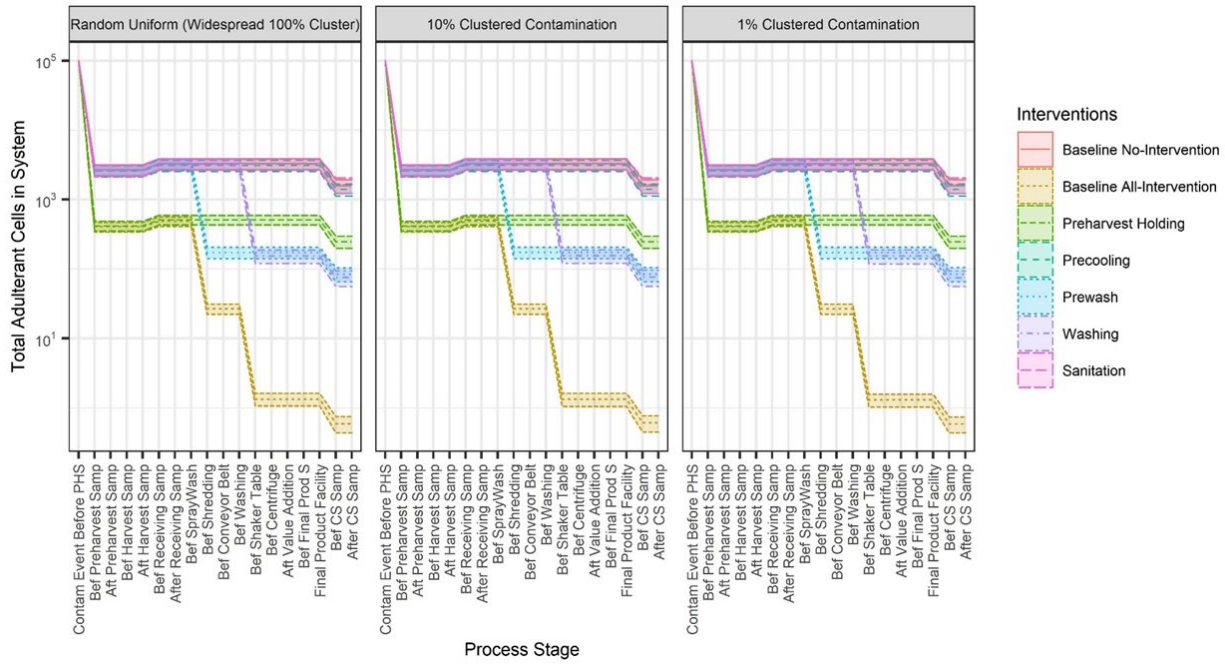


Figure 3: Contamination progression. Each panel represents a contamination scenario. The x-axis represents processing steps, the y-axis represents the Total Adulterant Cells (TAC) present in a system. The yellow bar is All-Interventions, and the red bar is No-Interventions. The other colored lines represent different systems with different individual interventions. The lines represent the mean, and the shaded areas represent the 95% CI of the mean. A drop in the TACs compared to the No-Intervention system demonstrates the effect of each or all interventions.

Figure 4: Sampling plan power results

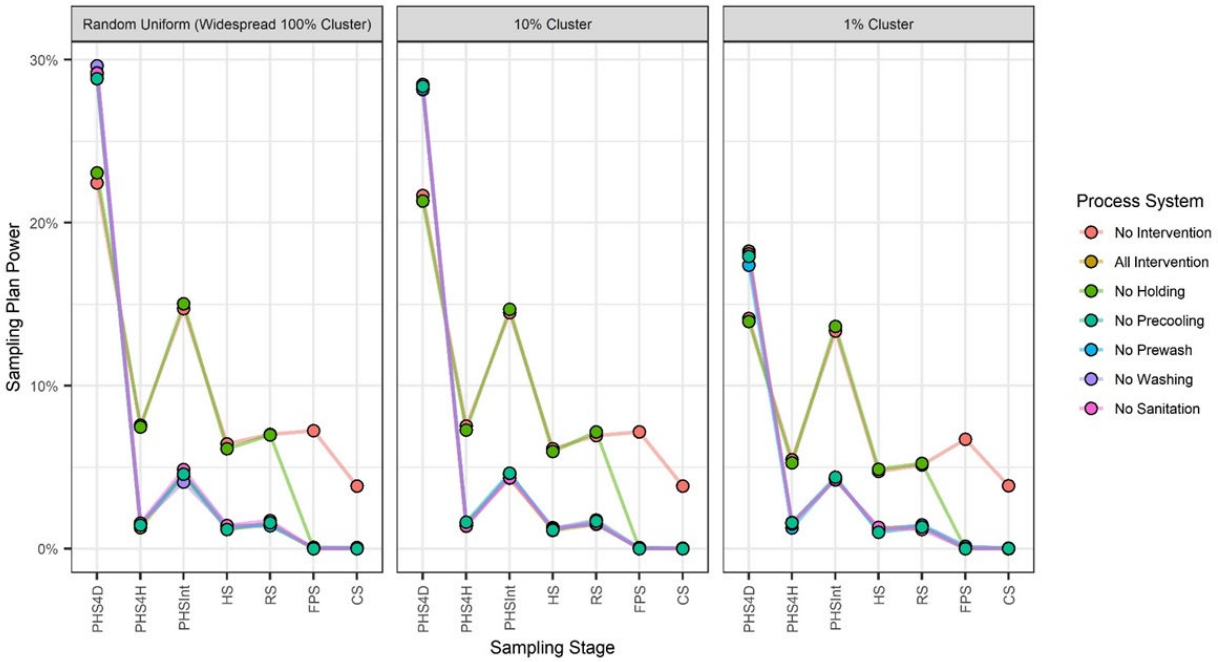


Figure 4: Sampling plan powers stratified by clustering levels and processing scenario. Sampling plan power is defined as the total number of iterations where the sampling plan detected the contamination/ Total # of iterations (n=10,000) The results indicate that preharvest sampling 4 days (PHS 4D) was the most powerful sampling plan with power ranging between 14.1% and 29.6%. Sampling plans at later stages such as finished product sampling (FPS) and customer sampling (CS) show lower sampling plans power especially when are implemented.

Figure 5: Sampling relative efficacy results

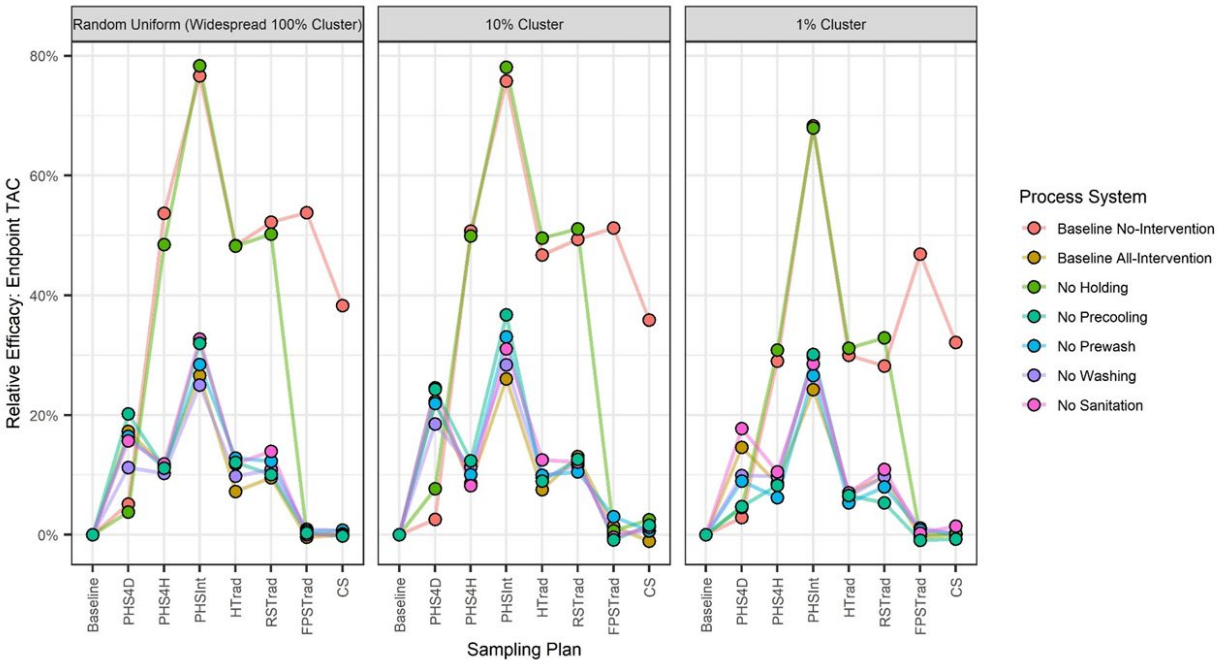


Figure 5: Final relative endpoint total adulterant cells (TACs) from 10,000 simulated contaminated iterations. Represents the relative endpoint TACs compared to each processing scenario. A lower relative difference of endpoint TACs means that the sampling plan was more effective compared to its given processing scenario. The results indicate that the sampling plan that most effectively reduced endpoint TACs is the preharvest intense (PHS Int) sampling plan. Sampling plans showed to be less effective when all interventions were in place except for preharvest sampling 4 days (PHS 4D).

Figure 6: Factor Sensitivity Analysis

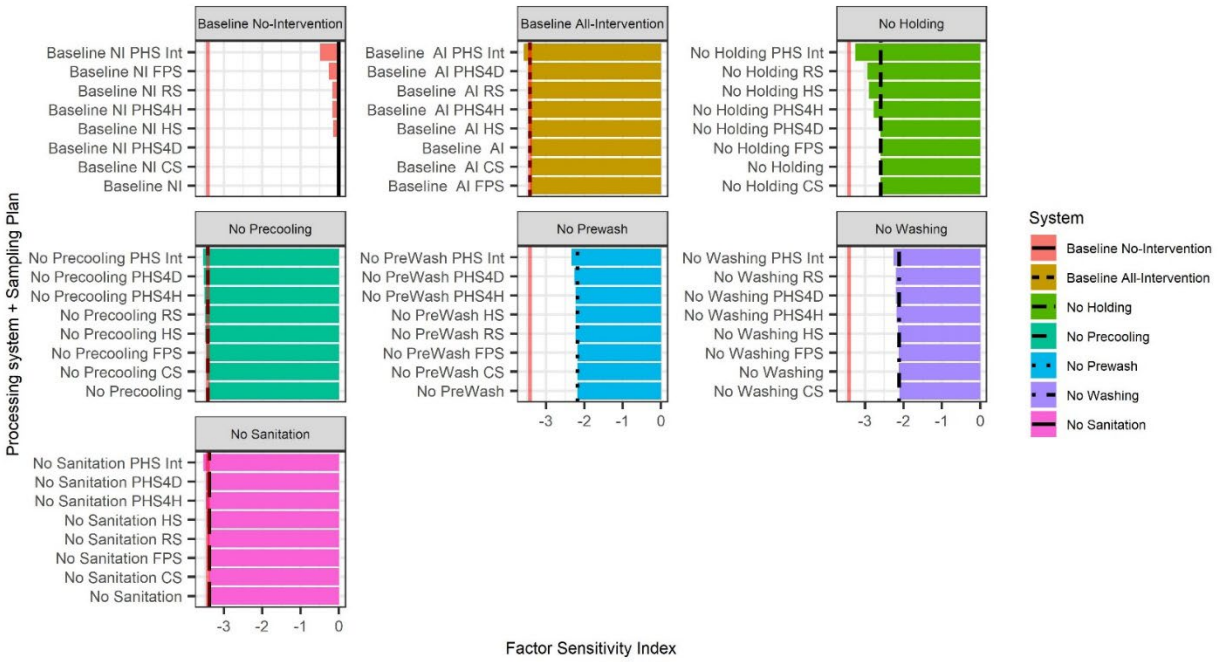


Figure 6: Results of the factor sensitivity analysis. Factor Sensitivity (FS) is the log reduction between the endpoint from each scenario and the system with no food safety interventions [Log10(Intervention/No-Interventions), for 10,000 iterations]. The greater the absolute value of FS, the more effect that specific scenario or condition has on endpoint TAC. The red line represents the effect of the All-Intervention system; the black line represents the effect of each system without sampling. The difference between the red and black lines represents the efficacy lost by removing interventions. The goal of the sampling plans is to take the bars beyond the black lines, the greater the distance between the black line and the bar, the greater the effect of a sampling plan in that specific system. Interventions have greater factor sensitivity than the sampling plans. When the All-Intervention system is in place, sampling plans show small to no effect.

Study 2 Figures: Tomatoes

Figure 7: Tomato process flow to farm-to-packinghouse model

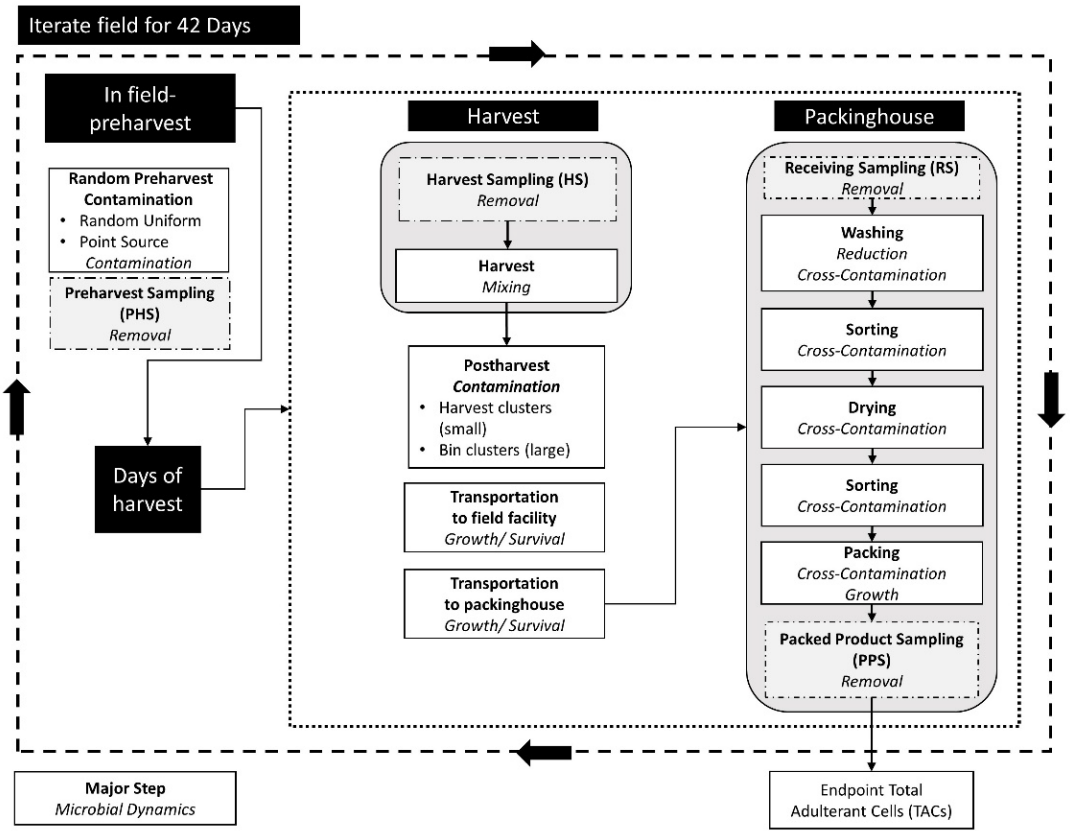


Figure 7: *Process flow for the tomato process, contamination event is simulated to occur in the field. Subsequent steps affect microbial dynamics from the farm to the packinghouse.*

Figure 8: Scenario analysis for the tomato farm-to-packinghouse analysis

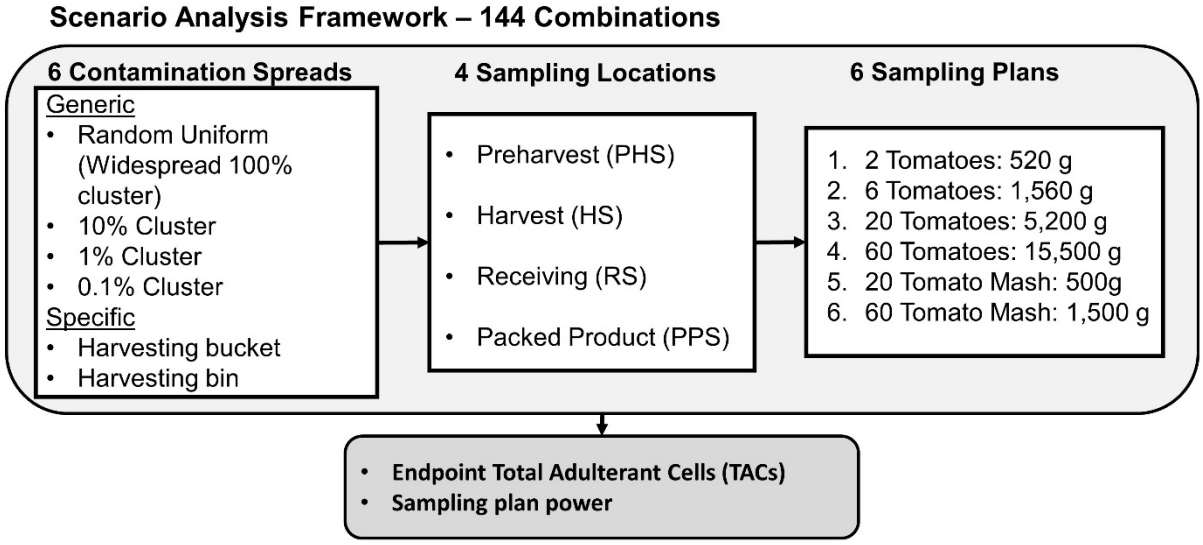


Figure 8: Scenario analysis framework. Contamination a total of 4 generic contamination spreads, 2 specific contamination spreads, 4 sampling locations, and 6 total sampling plans to obtain the relative efficacy and sampling plan power.

Figure 9: Sampling plan power results

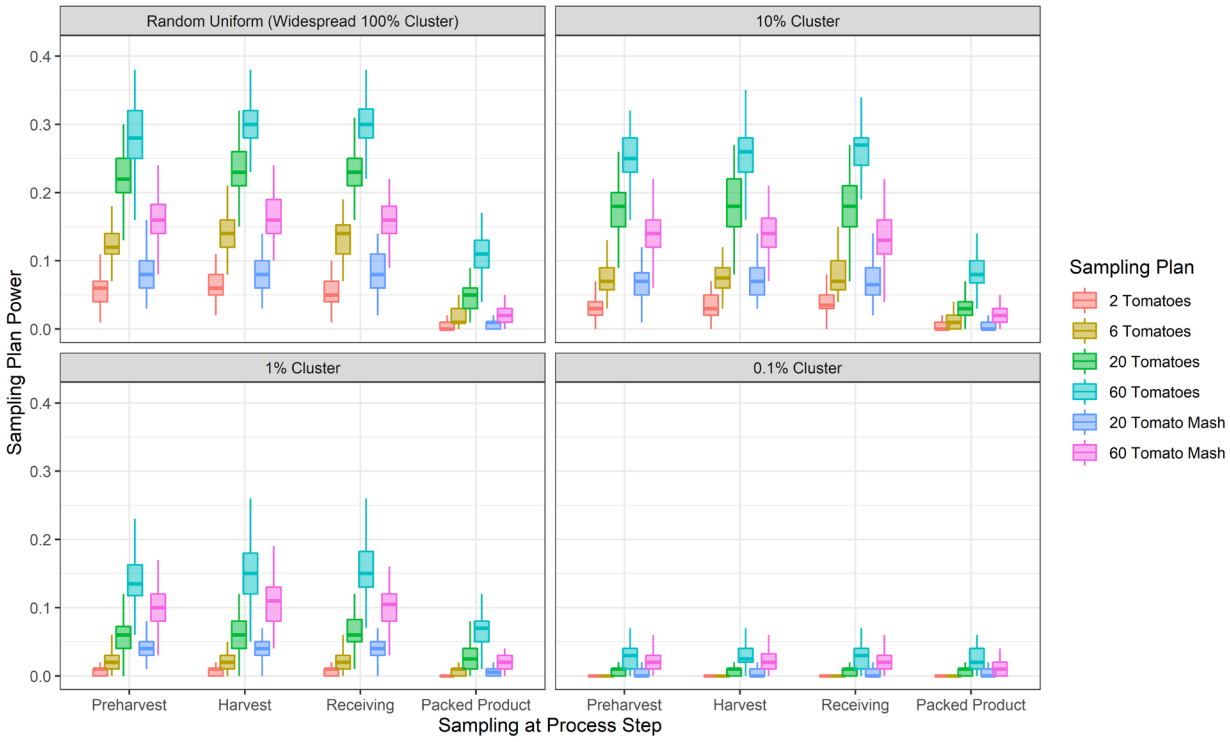


Figure 9: Summary of sampling plan powers. For a description of the sampling plans, reference Table 6. The dashed line represents the sampling plan power depends on the degree of clustering; as the cluster's size decreases, the power of the sampling plan will decrease. The red boxes represent that as the clusters decrease in size but increase contamination concentration, the total number of tomatoes sampled is more critical than the total mass tested. Whereas for larger clusters or random uniform (100% cluster) total mass sampled drives power.

Figure 10: Sampling plans relative efficacy on endpoint total adulterant cells (TACs)

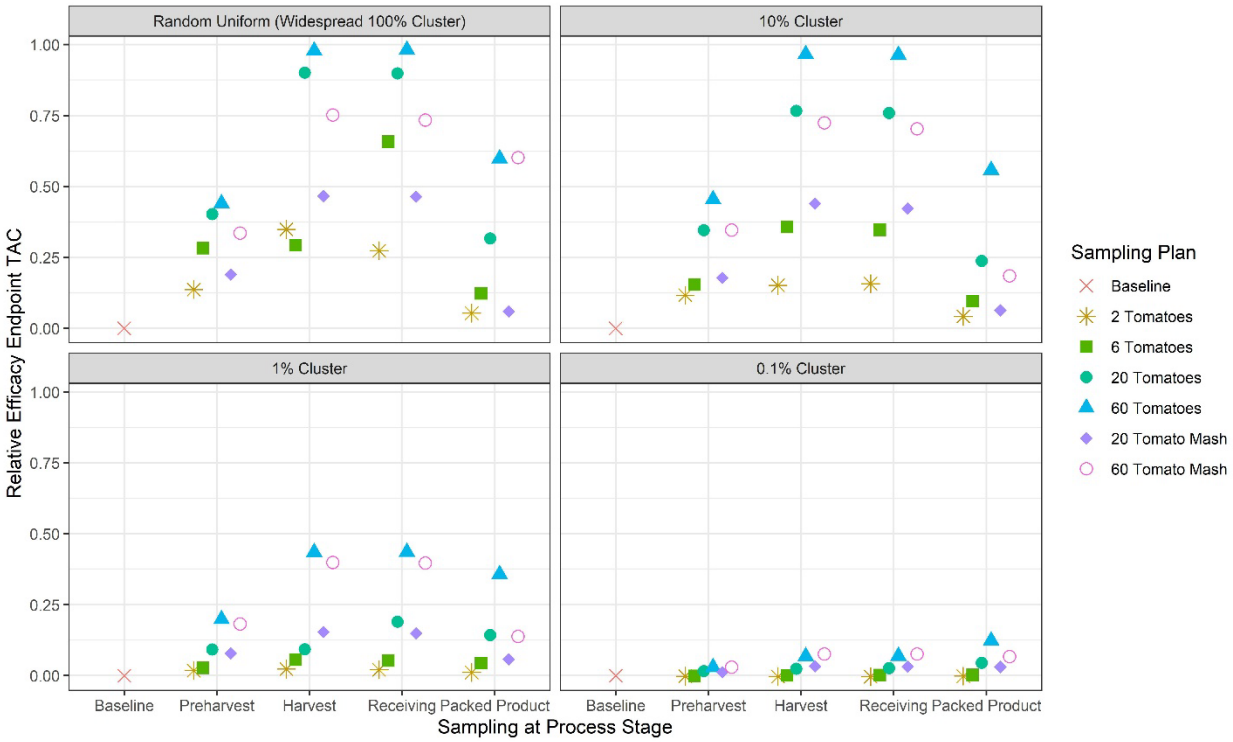


Figure 10: Summary of sampling plan relative efficacy. Stratified by contamination spread. The highest relative efficacy means that the sampling plan performed better at reducing cells from reaching the system endpoint. The blue boxes and text represent the best sampling location based on the contamination spread. The results suggest that for the random uniform (100% cluster) to 1% cluster the best sampling locations are at Harvest and Receiving, while for the 1% cluster the best sampling location is at packed product.

Figure 11: Relative efficacy for two additional contamination scenarios from the 2021 site visit

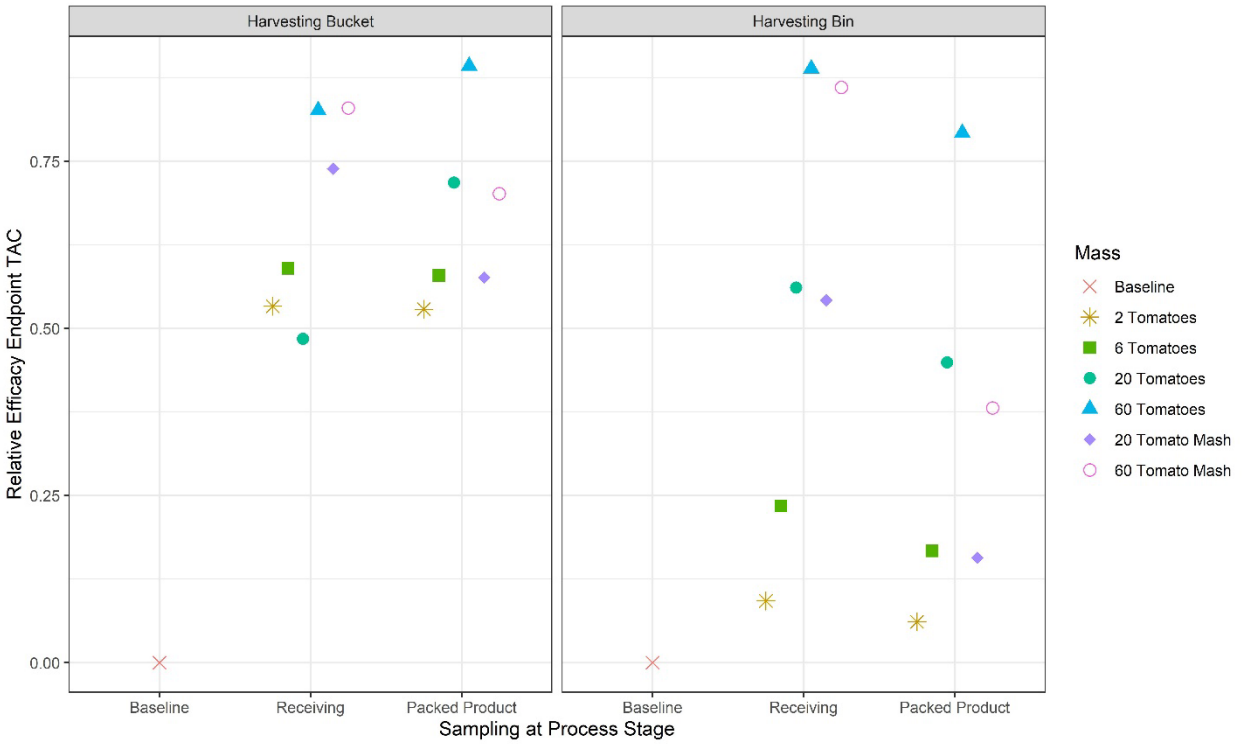


Figure 11: Two additional contamination scenarios. The relative efficacy plot shows the best sampling location best of the contamination event, as well as how well different sampling plans perform at detecting these contamination events.

Study 3 Figures: Cilantro

Figure 12: Probability of detection of C.cayetanensis for agricultural water testing

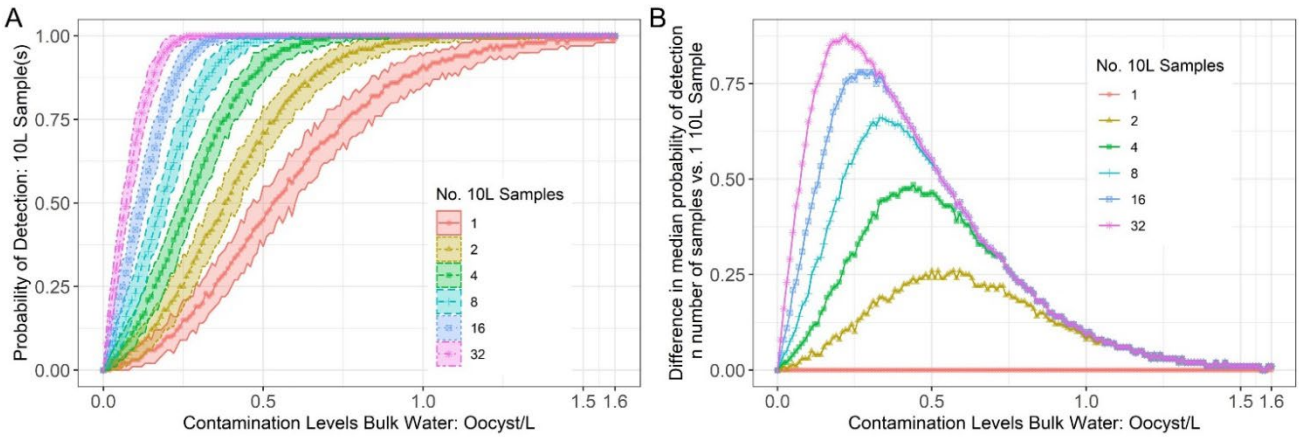


Figure 12: A) The probability of water testing detection compared to the total contamination levels in bulk water, and the total number of 10-liter samples. As the number of samples is doubled the probability of detection increases. B) The difference between the median probability of detection of the number of samples (n = 2, 4, 8, 16, 32) vs an (n = 1) 10-L sample. The area under the curve represents the added benefits on the probability of detection for taking more than 1 sample.

Figure 13: Probability of detection of C.cayetanensis for produce water testing

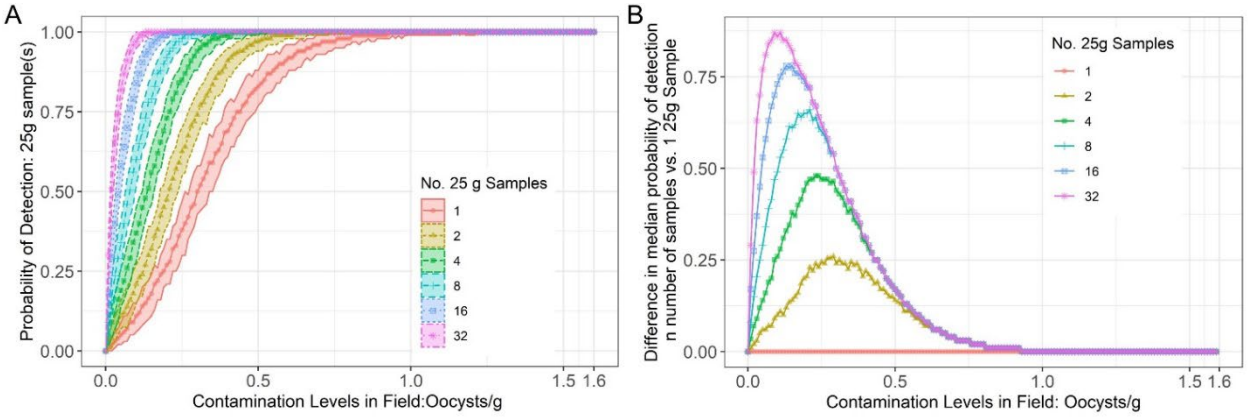


Figure 13: A) The probability of detection of product testing compared to the total contamination levels in the field (oocysts/g), and the total number of 25-g produce samples. As the number of samples is doubled the probability of detection increases. B) The difference between the median probability of detection of the number of samples (n = 2, 4, 8, 16, 32) vs an (n = 1) 25-g sample. The area under the curve represents the added benefits on the probability of detection for taking more than 1 sample.

Figure 14: Outcome of simulation scenarios after harvesting season of Baseline 1 simulation and scenario comparison

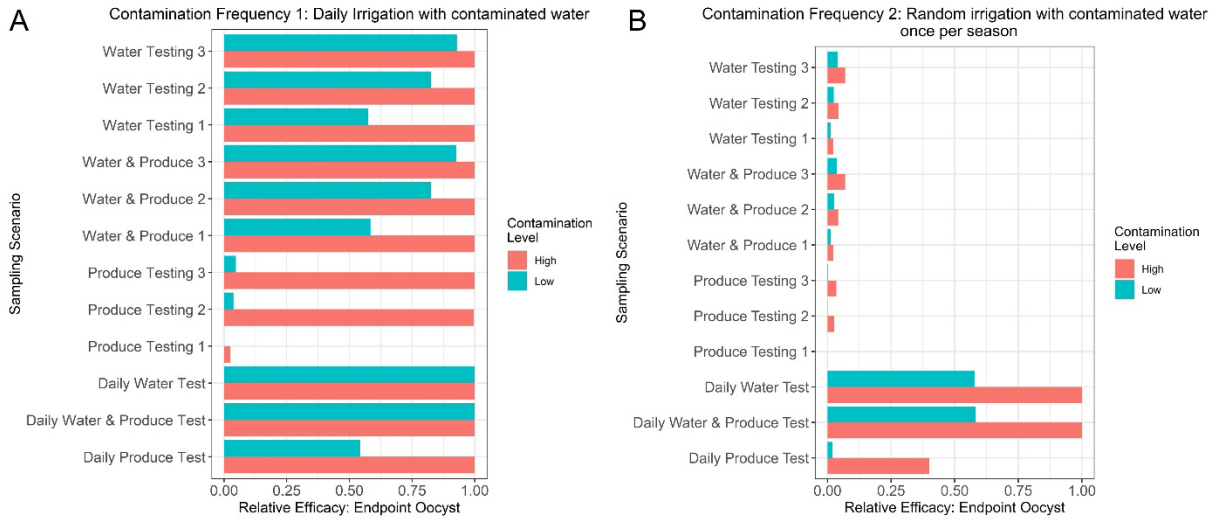


Figure 14: Relative efficacy on endpoint oocysts (1- endpoint oocysts sampling scenario/ scenario no sampling). Where 1 means that the sampling plan detected contamination and rejected the product, and 0 means contamination was not detected. A) Contamination frequency 1, where the irrigation water is contaminated throughout the season. B) Contamination frequency 2, random contamination, where irrigation water is contaminated once randomly per season.

References cited

1. U.S. Food and Drug Administration (FDA), *Factors potentially contributing to the contamination of leafy greens implicated in the fall 2020 outbreak of E. coli O157:H7.*, U.S.F.a.D. Administration, Editor. 2021.
2. U.S. Food and Drug Administration (FDA). *FDA Investigated Multistate Outbreak of E. coli O157:H7 Infections Linked to Romaine Lettuce from Yuma Growing Region*. 2018 [cited 2021 October 24]; Available from: https://www.fda.gov/food/outbreaks-foodborne-illness/fda-investigated-multistate-outbreak-e-coli-o157h7-infections-linked-romaine-lettuce-yuma-growing.
3. Centers of Disease Control and Prevention (CDC), *E. coli Outbreak Linked to Baby Spinach.* 2021.
4. M, V.A. *Outbreaks of Foodborne Diseases Associated with Tomatoes1*. 2021.
5. D'Lima, C.B. and T.V. Suslow, *Comparative Evaluation of Practical Functionality of Rapid Test Format Kits for Detection of Escherichia coli O157:H7 on Lettuce and Leafy Greens.* Journal of Food Protection, 2009. **72**(12): p. 2461-2470.
6. United Fresh. *Microbiological Testing of Fresh Produce: A White Paper on Considerations in Developing and Using Microbiological Sampling and Testing Procedures if Used as Part of a Food Safety Program for Fresh Fruit and Vegetable Products.* 2010 [cited 2022 January 26]; Available from: http://www.unitedfresh.org/newsviews/food_safety_resource_center.
7. (ERS), U.E.R.S., *Food Safety Practices and Costs Under the California Leafy Greens Marketing Agreement*. 2017.
8. International Commission on Microbiological Specifications for Foods (ICMSF), *Microorganisms in Foods 7 Microbiological Testing in Food Safety Management*. 2018, Boston, MA: Springer International Publishing.
9. Xu, A. and R.L. Buchanan, *Evaluation of sampling methods for the detection of pathogenic bacteria on pre-harvest leafy greens.* Food Microbiology, 2019. **77**: p. 137-145.
10. Quintanilla Portillo, J., et al., *A Validated Preharvest Sampling Simulation Shows that Sampling Plans with a Larger Number of Randomly Located Samples Perform Better than Typical Sampling Plans in Detecting Representative Point-Source and Widespread Hazards in Leafy Green Fields.* Applied and Environmental Microbiology. **88**(23): p. e01015-22.
11. Mokhtari, A., et al., *Evaluation of Potential Impacts of Free Chlorine during Washing of Fresh-Cut Leafy Greens on Escherichia coli O157:H7 Cross-Contamination and Risk of Illness.* Risk Analysis, 2021. **42**(5): p. 966-988.
12. Nauta, M.J., *The Modular Process Risk Model (MPRM): a Structured Approach to Food Chain Exposure Assessment*, in *Microbial Risk Analysis of Foods*, W. Schaffner Donald and M.P. Doyle, Editors. 2007, ASM Press: Washington, D.C. p. 99-136.
13. Pérez Rodríguez, F., et al., *A mathematical risk model for Escherichia coli O157:H7 cross-contamination of lettuce during processing.* Food Microbiology, 2011. **28**(4): p. 694-701.
14. Allende, A., et al., *Quantitative microbial exposure modelling as a tool to evaluate the impact of contamination level of surface irrigation water and seasonality on fecal hygiene indicator E. coli in leafy green production.* Food Microbiology, 2018. **75**: p. 82-89.
15. Pang, H., et al., *Quantitative Microbial Risk Assessment for Escherichia coli O157:H7 in Fresh-Cut Lettuce.* Journal of Food Protection, 2017. **80**(2): p. 302-311.
16. Yanamala, S., et al., *Potential for microbial contamination of spinach through feedyard air/dust growing in close proximity to cattle feedyard operations.* Journal of Food Safety, 2011. **31**(4): p. 525-529.

17. Berry Elaine, D., et al., *Effect of Proximity to a Cattle Feedlot on Escherichia coli O157:H7 Contamination of Leafy Greens and Evaluation of the Potential for Airborne Transmission.* Applied and Environmental Microbiology, 2015. **81**(3): p. 1101-1110.
18. Fonseca, J.M., et al., *Assessing the Food Safety Risk Posed by Birds Entering Leafy Greens Fields in the US Southwest.* International Journal of Environmental Research and Public Health, 2020. **17**(23): p. 8711.
19. Mishra, A., et al., *A System Model for Understanding the Role of Animal Feces as a Route of Contamination of Leafy Greens before Harvest.* Applied and Environmental Microbiology. **83**(2): p. e02775-16.
20. Weller, D.L., et al., *A Conceptual Framework for Developing Recommendations for No-Harvest Buffers around In-Field Feces.* Journal of Food Protection, 2019. **82**:1052-1060.
21. Brennan, J., et al. *Would my sampling plan have detected contamination levels that resulted in an outbreak? A Thought Experiment*. 2021.
22. Western Growers. *Apendix C Pre-harvest Product Sampling and Testing Protocol*. 2022 June 8 [cited 2022 October 8]; Available from: https://www.wga.com/sites/default/files/resource/files/Appendix%20C%20-%20Product%20Testing%20protocol%20Version%206-8-2022%20FINAL.pdf.
23. Jongenburger, I., H.M.W. Den Besten, and M.H. Zwietering, *Statistical Aspects of Food Safety Sampling.* Annual Review of Food Science and Technology, 2015. **6**(1): 479-503.
24. Zwietering, M.H. and S.J.C. van Gerwen, *Sensitivity analysis in quantitative microbial risk assessment.* International Journal of Food Microbiology, 2000. **58**(3): p. 213-221.
25. Miranda, R.C. and D.W. Schaffner, *Farm to fork quantitative microbial risk assessment for norovirus on frozen strawberries.* Microbial Risk Analysis, 2018. **10**: p. 44-53.
26. Python Software Foundation. 2001-2022.
27. Castillo, A. and M.O. Rodríguez-García, *Bacterial hazards in fresh and Fresh-Cut produce: Sources and control.* Preharvest and postharvest food safety: contemporary issues and future directions, 2004: p. 43-58.
28. Barak, J.D. and A.S. Liang, *Role of Soil, Crop Debris, and a Plant Pathogen in Salmonella enterica Contamination of Tomato Plants.* PLoS ONE, 2008. **3**(2): p. e1657.
29. GURTLER, J.B., et al., *Salmonella enterica Contamination of Market Fresh Tomatoes: A Review.* Journal of Food Protection, 2018. **81**(7): p. 1193-1213.
30. U.S Food and Drug Administration (FDA). *BAM Chapter 5: Salmonella*. 2022 [cited 2022 December 7]; Available from: https://www.fda.gov/food/laboratory-methods-food/bam-chapter-5-salmonella.
31. Durigan, M.A.-O., H.R. Murphy, and A.J. da Silva, *Dead-End Ultrafiltration and DNA-Based Methods for Detection of Cyclospora cayetanensis in Agricultural Water. LID - 10.1128/AEM.01595-20 [doi] LID - e01595-20.* (1098-5336 (Electronic)).
32. U.S Food and Drug Administration (FDA). *BAM 19c: Dead-end Ultrafiltration for the Detection of Cyclospora cayetanensis from Agricultural Water*. 2020 [cited 2022 October 17]; Available from: https://www.fda.gov/media/140309/download.
33. Durigan, M.A.-O., H.R. Murphy, and A.J. da Silva, *Dead-End Ultrafiltration and DNA-Based Methods for Detection of Cyclospora cayetanensis in Agricultural Water*. LID - 10.1128/AEM.01595-20 [doi] LID - e01595-20. (1098-5336 (Electronic)).
34. Murphy, H.R., et al., *Interlaboratory validation of an improved method for detection of Cyclospora cayetanensis in produce using a real-time PCR assay*. 2018(1095-9998 (Electronic)).
35. U.S Food and Drug Administration (FDA). *BAM Chapter 19b: Molecular Detection of Cyclospora cayetanensis in Fresh Produce Using Real-Time PCR*. 2022 [cited 2022 October 17]; Available from: https://www.fda.gov/food/laboratory-methods-food/bam-chapter-19b-molecular-detection-cyclospora-cayetanensis-fresh-produce-using-real-time-pcr.

36. Lalonde, L., J. Oakley, and P. Fries, *Verification and Use of the US-FDA BAM 19b Method for Detection of Cyclospora cayetanensis in a Survey of Fresh Produce by CFIA Laboratory*. Microorganisms, 2022. **10**(3): p. 559.
37. California Department of Food and Agriculture, *County agricultural commissioners' reports crop year 2019-2020*. 2022.
38. Western Institute for Food Safety & Security (WIFSS). *Cilantro.* 2016; Available from: https://wifss.ucdavis.edu/wp-content/uploads/2016/10/Cilantro_PDF.pdf.
39. Smith, R., et al., *Cilantro Production in California*. University of California, Agriculture and Natural Resources, 2011.
40. Mangan, F. *Cilantro, Coriandrum sativum*. 2018 [cited 2022 December 5].
41. *Cilantro | Food Source Information*. 2021; Available from: https://fsi.colostate.edu/cilantro/.
42. Ortega, Y.R. and L.J. Robertson, *Cyclospora cayetanensis as a foodborne pathogen*, ed. Springer. 2017.