Risk Assessment of *Escherichia coli* O157:H7 along the Farm-to-Fork Fresh-Cut Romaine Lettuce Supply Chain

Names and e-mail addresses for all authors

Ece Bulut[1*] eb643@cornell.edu
Sarah I. Murphy[1] sim39@cornell.edu
Laura K. Strawn[2] laurakstrawn@vt.edu
Michelle D. Danyluk[3] mddanyluk@ufl.edu
Martin Wiedmann[4] martin.wiedmann@cornell.edu
Renata Ivanek[1] ri25@cornell.edu

Author affiliation(s)

[1]Department of Population Medicine and Diagnostic Sciences, Cornell University, Ithaca, NY 14853, USA
[2]Department of Food Science and Technology, Virginia Tech, Blacksburg, VA 24061, USA
[3]Food Science and Human Nutrition Department, University of Florida, Lake Alfred, FL 33850, USA
[4]Department of Food Science, Cornell University, Ithaca, NY 14853, USA

Contact information for Corresponding author

eb643@cornell.edu
602 Tower Rd, Ithaca, NY 14853

Abstract

Outbreaks of *Escherichia coli* O157:H7 (ECO157) in romaine lettuce remain an ongoing public health concern. Quantitative microbial risk assessment (QMRA) models are key tools for identifying control measures to mitigate foodborne diseases. Here, we introduce a comprehensive QMRA framework along the farm-to-fork fresh-cut romaine lettuce chain, including a novel preharvest difference equation model, to predict annual ECO157 illness cases in the United States and evaluate control strategies. We demonstrated the importance of managing irrigation contamination at preharvest to control illness cases. Wildlife intrusions pose lower health risk, followed by runoffs and biological soil amendments of animal origin. When preharvest contamination persists and combines with time-temperature abuses at postharvest, the predicted ECO157 illness cases rise considerably. We showed a broad range of interventions targeting both preharvest and postharvest stages can effectively improve the microbial safety of fresh-cut romaine. The comprehensive practices and interventions explored in this study will aid decision-makers in establishing/enhancing food safety best management practices.

Keywords: Quantitative microbial risk assessment, Escherichia coli O157:H7, lettuce, farm to fork

Main

The number of detected foodborne illness outbreaks linked to leafy greens has increased in the United States (USA) between 1996 and 2016, with *Escherichia coli* O157:H7 (ECO157)-contaminated romaine lettuce emerging as the leading source *[1]*. Between 2015 and 2021, ECO157-contaminated romaine has been implicated in 7 outbreaks in the USA, 6 of which were multistate incidents *[2,3]*. These outbreaks resulted in 4,274 laboratory-confirmed illnesses, leading to 766 hospitalizations and 11 deaths.

Contamination of leafy greens, including romaine, can occur at any point during preharvest, harvest and postharvest, involving various risk factors *[4]*. Preharvest risk factors may involve domestic animals *[5,6]*, wild animals *[4,7-9]*, biological soil amendments of animal origin (BSAAO) amended soil *[10]*, irrigation water *[11]* and runoff from animal operations *[12,13]*. Harvest and postharvest risk factors may include humans *[14]*, harvesting/production equipment *[15,16]* and production water *[17]*. In addition, the lack of processing steps that ensure effective removal or inactivation of ECO157 before consumption *[14]*, coupled with the recent increase in romaine consumption *[18]*, further elevates the risk.

In the USA, 90% of the leafy green growers in California and Arizona are members of the California or Arizona Leafy Green Products Handler Marketing Agreements (LGMA) *[19,20]*. These agreements have established guidelines to minimize contamination during the production and harvesting of leafy greens. The guidelines cover best practices in areas such as personnel training, agricultural water management, soil amendments, harvest equipment cleaning, packaging, transportation, produce testing, and prevention of animal intrusion. For example, both LGMAs prohibit the use of Type B agricultural waters—potentially contaminated with fecal indicators—for overhead irrigation of leafy greens within 21 days of harvest. If Type B water is used, it must be treated to meet Type A standards (agricultural water unlikely to contain fecal contamination) and must pass microbiological testing criteria. Additionally, the use of untreated, un-composted, incompletely composted or non-thermally treated BSAAOs is prohibited. Leafy greens cannot be produced from fields where such BSAAOs have been applied until one year has passed. If wildlife intrusion is detected, affected crops must be isolated within a three-foot radius. Furthermore, the position of animal feeding operations near leafy green farms must be assessed for any issues that might impact safety. The FDA Food Safety Modernization Act (FSMA) Produce Safety Rule (PSR) mandates similar standards for agricultural water, soil amendments, animal management, personnel training and equipment cleaning *[21]*. While FSMA PSR does not require microbiological testing of agricultural water *[22]*, it mandates routine inspections to ensure agricultural water safety and proper maintenance *[22]*. The application of raw BSAAOs is regulated by the '90-120-day' rule, which requires that raw BSAAOs be incorporated into the soil at least 120 days before harvest *[23]*.

Stabilized compost must meet a ECO157 count of < 0.3 MPN per gram of BSAAO *[24]*. Produce likely to be contaminated by animal intrusion should not be harvested *[25]*. The impact of nearby animal feeding operations should be evaluated based on their effects on agricultural water, such as drainage or discharge of untreated water *[26]*.

Quantitative microbial risk assessment (QMRA) models are commonly used to quantify the risk of microbial contamination or illness associated with exposure to microbial hazards. The outcome of these models provides valuable information for decision makers for setting regulatory standards, designing interventions to reduce microbial risks, or assessing the effectiveness of control measures *[27]*. Over the years, numerous QMRA models have been developed for leafy greens contaminated with *Escherichia coli* strains, including ECO157 *[28-41]*. These risk assessments have investigated numerous pathways of *E. coli* contamination along the farm-to-fork leafy greens supply chain. Nevertheless, an updated risk assessment tool is needed to thoroughly evaluate the potential health implications of a broader range of factors related to ECO157 exposure and spread.

We developed a comprehensive QMRA aiming to model ECO157 contamination along the fresh-cut romaine lettuce production and consumption chain in the USA. Our model incorporates a broad set of factors, utilizes a novel difference equation modeling approach at the preharvest stage, and accounts for ECO157 exposure and spread throughout the farm-to-fork chain, as informed by existing literature and expert elicitation. This QMRA was designed as a decision support tool for managing various risk factors of ECO157 contamination in the USA and is publicly available via a GitHub repository (https://github.com/IvanekLab/Romaine_ECO157_QMRA).

Methods

Overview of the model

In this study, we adapted elements from QMRAs developed for leafy greens in Spain *[27]*, Belgium *[39]* and USA *[32,36,38]*. We integrated these QMRAs with a modeling method for partitioning and mixing processes for microbial organisms in the food supply chain *[42]*, as well as the concept of non-persister and persister cell state dynamics *[43]* for modeling biphasic decay of ECO157 on romaine and in soil. We also used the FSMA PSR, California and Arizona LGMA guidelines, and an expert elicitation approach *[44]* to capture practices specific to the USA. This resulted in a comprehensive risk assessment model that simulates various field and processing/handling conditions. We predicted the number of annual illness cases in the USA attributed to consuming ECO157-contaminated fresh-cut romaine, which is our main model output. We integrated two models into an exposure assessment: the preharvest model (including preharvest and harvest sub-models) and postharvest model (including processing-

retail, and consumer sub-models). In the preharvest model, we used a difference equation model to track the population dynamics of ECO157 in soil, irrigation water and romaine prior to harvest. Figure 1 provided an overview of the QMRA model's structure and Tables 1, 2 and 3 described the parameters used in the model. The unit of analysis in the preharvest model was a (romaine) 'batch', but it becomes a 'package' and then a 'serving' in the postharvest model. We defined a 'lot' as all romaine packages produced in a day between two cleaning and sanitation shifts. We assumed that a lot was harvested from a single 1-acre field and processed in a day. Based on this assumption, each lot contained 12 romaine batches or 26,000 romaine plants. To quantify the variability and uncertainty of the model outcomes, we performed a Monte Carlo simulation using Latin Hypercube Sampling with 100,000 iterations. This was executed in Microsoft Excel 2016 (Microsoft Corporation, WA USA), utilizing the @RISK add-on software (version 8.5.0, Lumivero, CO USA). Model validation was carried out by comparing the predicted annual illness cases with the data from existing surveillance studies *[45,46]*. Expert elicitation, along with sensitivity and uncertainty analyses were conducted to guide the selection of promising risk management strategies and informed parameter choices in scenario analyses. Regression trees were used in the preharvest and postharvest models to identify interaction of factors that lead to ECO157 contamination in romaine and illness cases. The following sections detailed the baseline model structure.

Preharvest model

The preharvest model included growing and harvesting romaine in an open field. We considered a square field measuring approximately 63 m by 63 m (equivalent to 1-acre) for the cultivation of romaine. A total of 26,000 romaine plants were harvested from the modeled field in a day. Each plant was assumed to measure 25cm in width, length and height, and to weigh 300g. The allocated area for each individual plant within the field was 30cm by 50cm *[36]*.

We assumed continuous romaine production throughout the year, with daily harvests, occurring in one of three distinct geographical locations: (i) the Coastal Region and Central Valley in California (CC), (ii) the Desert Region (Imperial) in California (DR), and (iii) the Yuma Region in Arizona (Y). These locations were selected since they are the top leafy green producers in the USA *[47-49]*. The choice of location affected several model parameters, such as the time from seed planting to harvest, and the type of irrigation systems used in the fields (Table 1). Additionally, the month of romaine harvest affected the parameters related to weather conditions and survival of ECO157 (Table 3).

Contamination of romaine in the field before harvest was attributed to four sources based on the available information in the literature and the expert elicitation approach: (i) runoff from a neighboring dairy farm, (ii) intrusion by feral swine, (iii) irrigation water, and (iv) the application

of BSAAO (i.e., treated/composted/processed dairy manure, which may include incompletely treated/composted/processed dairy manure). All calculations were conducted for a randomly selected 1/12 portion of the modeled field (i.e., 1/12 lot), which we refer to as the 'batch area', containing 2,125 romaine plants that weigh 637.5kg at harvest. The plants were grown and harvested as a batch unit and population dynamics of ECO157 were tracked for the batch. The size of the batch area enabled the development of a more realistic model by accounting for the risk factors associated with the exposure and spread of ECO157 throughout the farm-to-fork chain. For example, we assumed that only the batch area closest to the edge of the field (and thus to a hypothetical neighboring dairy farm) would be impacted by dairy farm runoff. Due to the rapid decay of ECO157 on field-grown romaine, we adopted a 14-day 'countdown' period and introduced ECO157 to the romaine only in the final 14 days before harvest *[39]*. The duration of this period was chosen assuming that furrow irrigation happened every 10 days on average and that an average waiting time of 4 days was used before harvest. Throughout this countdown period, we tracked ECO157 counts on day *t* on both the romaine ($C_{L_{t+1}}$; CFU/romaine batch) and in the top 1cm layer of soil of the batch area (referred to as 'soil batch'; $C_{S_{t+1}}$; CFU/soil batch) using a system of coupled difference equations (1) - (6) coded within the QMRA in @RISK (notations were explained in the following paragraphs and also in Supplementary Tables S1 and S2):

$$LN_{t+1} = LN_t + Irr_{L_t} + Pig_{L_t} + Runoff_{L_t} + \left(R_{splash_t} + I_{splash_t}\right) \times (SN_t + SP_t) - LNtoLP_t - LNdecay_t \quad (1)$$

$$LP_{t+1} = LP_t + LNtoLP_t - LPdecay_t \quad (2)$$

$$C_{L_{t+1}} = LN_{t+1} + LP_{t+1} \quad (3)$$

$$SN_{t+1} = SN_t + Irr_{S_t} + Pig_{S_t} + Runoff_{S_t} - \left(R_{splash_t} + I_{splash_t}\right) \times SN_t - SNtoSP_t - SNdecay_t \quad (4)$$

$$SP_{t+1} = SP_t + SNtoSP_t - \left(R_{splash_t} + I_{splash_t}\right) \times SP_t - SPdecay_t \quad (5)$$

$$C_{S_{t+1}} = SN_{t+1} + SP_{t+1} \quad (6)$$

where $t = 0,1,2,3,\ldots,13,14$ were days of the countdown period (i.e., at t=0 there were 14 days until harvest). If in a time step, the calculated LN_{t+1}, LP_{t+1}, SN_{t+1} or $SP_{t+1} < 0$, we set the corresponding LN_{t+1}, LP_{t+1}, SN_{t+1} or SP_{t+1} to 0.

We adopted the concept of non-persister and persister cell state dynamics to provide a more accurate die-off mechanism for ECO157 *[50]*, reflecting the well-known biphasic decay phenomenon (i.e., fast initial decay followed by slower decay) of the ECO157 cells on leafy green and in soil environments *[51-54]*. In Equation (1), LN_{t+1} represented non-persister

ECO157 count in romaine batches on day t+1. It was determined by the previous day's count in romaine batches (LN_t); accumulation through irrigation (Irr_{L_t}), runoff ($Runoff_{L_t}$), intrusion by feral swine (Pig_{L_t}), rain splash ($R_{splash_t} \times (SN_t + SP_t)$) and irrigation splash ($I_{splash_t} \times (SN_t + SP_t)$); removal through cells that switched from non-persister to persister state ($LNtoLP_t$) and daily reduction on romaine ($LNdecay_t$). Equation (2) describes the persister ECO157 count in romaine batches on day t+1 (LP_{t+1}). Total ECO157 count (non-persister and persister cells) in romaine batches on day t+1 ($C_{L_{t+1}}$) was calculated in Equation (3). Equations (4), (5) and (6) estimated the total ECO157 count (non-persister and persister) in soil batches on day t+1 ($C_{S_{t+1}}$) and were similar to the calculations for romaine batches. It was assumed that the cross-contamination of ECO157 cells from soil to romaine via splash events triggers an adaptation process, resulting in a transformation of any transferred cells from a persister state to a non-persister state when they land on romaine leaves *[43]*.

The ECO157 counts in romaine batches ($C_{L_{14}}$) and soil batches ($C_{S_{14}}$) at *t*=14 were used in a harvest sub-model (explained below) to calculate the preharvest outcome: the ECO157 count in harvested romaine batches (C_{L_h}). From there, the ECO157 prevalence (P_{Lh_f}) and the ECO157 count in harvested romaine batches (C_{Lh_f}) were calculated in one simulation (100,000 iterations) and used as inputs in the postharvest model. Specifically, P_{Lh_f} was derived by counting the iterations out of 100,000 where the ECO157 count on harvested romaine batches was at least one (where $C_{L_h} \geq 1$ CFU/romaine batch). The P_{Lh_f} was established as a Triangular distribution, where the mode/mean of the distribution corresponded to the predicted preharvest prevalence in one simulation. The distribution's minimum and maximum were set at half and twice of the predicted prevalence, respectively. The C_{Lh_f} was derived by fitting a distribution around the predicted ECO157 counts in contaminated harvested romaine batches (where $C_{L_h} \geq 1$ CFU/romaine batch) using @RISK software. The distribution that best fit the data was determined using Akaike Information Criterion, Bayesian Information Criterion and Average Log-Likelihood methods (Table 2).

Accumulation and loss of ECO157 contamination in soil batch (C_S)

We assumed that irrigation water, BSAAOs, swine, and cattle feces were the only sources of ECO157 contamination for the modeled field, explained in the following sections.

BSAAO Application (accumulation in soil): We modeled ECO157 prevalence (P_{man}) and count (C_{man}) in BSAAO by accounting for the reports of ECO157 in stacked/piled manure from dairy farms in the USA *[55]*. The maximum ECO157 count in BSAAO was the maximum ECO157 count found on raw dairy feces *[36]*. This represented an unintentional application of an incompletely treated BSAAO, though the application probability of such BSAAO was low. The timing of

BSAAO application (D_{man}) was set on the same day of planting romaine seeds. The number of days from planting romaine seeds to harvest ($Daytogrow$) was set between 65 and 130 days depending on the field location (CC, DR or Y regions) *[44]*. We considered a constant daily decay of ECO157 among both non-persister and persister cells in BSAAO (R_{dman} and R_{dman_P}), from the day of BSAAO application to the time of harvest *[56,57]*. A total of 10% of the simulated fields received BSAAO application (Pr_{man}), assuming that those were organic operations that used BSAAO as nitrogen source. The BSAAO application (M_{man}) ranged from 5 to 15 tons per acre *[88]*. When applying BSAAO, we incorporated a dilution factor (DiF_{man}) assuming a uniform distribution of BSAAO across the tilled column of the soil batch. This approach suggested that a gram of BSAAO has a volume (Rh_{con}) of 2.5cm^3 *[58]* and was uniformly distributed across the topsoil (Dh_{con}) up to 15 to 20cm depth *[59]*. In this and other instances of using a dilution factor in the soil column in our model, the assumption of a uniform distribution of ECO157 within the top layer of soil batch was justified since the unit of analysis of romaine contamination was the whole batch. For example, transfer of 100 ECO157 cells from one contaminated spot of soil onto a single romaine plant (which becomes contaminated with 100 ECO157) and transfer of 1 ECO157 cell from each of 100 different contaminated soil spots resulting in 100 plants contaminated with 1 ECO157 each, resulted in the same contamination at the batch level (i.e., 100 ECO157 in a given romaine batch). The estimation of the initial ECO157 count in soil batch (C_{S_0}) was provided in Supplementary Table S1.

Runoff from a neighboring dairy farm (accumulation in soil) ($\boldsymbol{Runoff_{S_t}}$): We adapted the runoff contamination component from Mishra et al.'s (2016) QMRA model *[36]*. Runoff events were considered to occur on rainy days only (at a probability $1 - Bs_t$; a day could either be sunny or rainy). They lead to contamination of batch area at the edge of the field, determined randomly by the '$edge$' parameter. We assumed that 5 to 15kg of dairy cattle feces (M_{cow}) could enter the batch area during a runoff event, where the quantity of fecal material entering the batch area was proportional to the monthly average precipitation of the field location ($Prec_{norm_m}$). We incorporated ECO157 count (C_{cow}) and month-specific ECO157 prevalence (P_{cow_m}, Table 3) in dairy cattle feces from Cooley et al. (2013) *[8]*. The spread of runoff across the $edge$ batch area affected both romaine and soil, determined by the 'Pr_{LtoS}' parameter. This parameter was a proportion describing the area a romaine plant occupies relative to the soil area allocated for plant growth. The runoff in soil was uniformly distributed and spread into the soil. The ECO157 in runoff was spread to a soil depth of 1cm by a dilution factor (DiF_{cow}). The estimation of the term $Runoff_{S_t}$ in the model (1) - (6); i.e., the ECO157 count in soil batch added via runoff on day t, was provided in Supplementary Table S2.

Intrusion by feral swine (accumulation in soil) ($\boldsymbol{Pig_{S_t}}$): Wildlife intrusion in our model was integrated using the risk model developed by Mishra et al. (2016). A contamination event was

dependent on the probability of feral swine intruding the field ($Pr_{pig_{countt}}$), derived from reported feral swine population density in Salinas Valley, CA (~ 5 swine/km^2) *[6]*. In the event of an intrusion, we assumed that between 1-5 adult swine would enter the field and defecate, as they could be found individually and in family groups in CA *[60]*. In $Pr_{pig_{countt}}$, the probability of intrusion by one animal was the highest, and the probability of intrusion by more animals decreased progressively. The fecal quantity found on the field after the defecations (M_{pig}) was estimated by the daily fecal production reported per adult swine ($Mdef_{pig}$) *[61]*. In addition, we factored in a defecation frequency of 4 times/swine/day ($Tdef_{pig}$) *[62]* and assumed at least one of those defecation events occurred in the modeled field ($Ndef_{pig}$). Feral swine could defecate on romaine plants or field soil, determined by the 'Pr_{LtoS}' parameter. When feces fall on soil, we assumed the ECO157 in fecal material was spread evenly within the top 1cm of the soil batch, which we modeled using a dilution factor (DiF_{pig}). The ECO157 count (C_{pig}) and the ECO157 prevalence (P_{pig_m}) in swine feces were modeled based on the data provided by Cooley et al. (2013) *[8]*. The term Pig_{S_t} in the model (1) - (6); i.e., ECO157 count in soil batch added via feral swine on day t, was estimated as shown in Supplementary Table S2.

Irrigation (accumulation in soil) ($\boldsymbol{Irr_{s_t}}$): Contamination of soil by irrigation occurred through contaminated water that was delivered to the field and not captured by the romaine plants on the field. The ECO157 prevalence ($P_{irr} = P_{irr_W}$ or P_{irr_S}) and count ($C_{irr_t} = C_{irr_W}$ or C_{irr_S}) in the irrigation water depended on the water source: surface or well (ground) waters. Water source was selected based on its respective frequency of use in field location [*63]*. We assumed the value for the ECO157 prevalence in well water (P_{irr_W}) as 0.1%, given the limited literature on well waters, and considering that well waters were generally free of ECO157. The ECO157 prevalence in surface water (P_{irr_S}) was based on the ECO157 survey on surface waters in Monterey County, CA *[8]*. The ECO157 counts for both surface and well waters included two parts: (i) the generic *E. coli* counts found in both water sources and (ii) a factor that converts the generic *E. coli* counts to the ECO157 counts *[34]*. The generic *E. coli* count in well water was the LGMA testing limit of generic *E. coli* for Type A agricultural waters, set at 10MPN/100ml *[19,20]*. For consistency across model calculations, we assumed that the most probable number (MPN) was equivalent to colony forming units (CFU) *[64]*. The generic *E. coli* count in surface water (C_{irr_S}) was represented as a PERT distribution, with the minimum and most likely values taken from the standard testing limits set by the LGMAs for Type B agricultural waters used in drip and furrow irrigation systems *[19,20]*. The maximum value of the PERT distribution was based on the highest generic *E. coli* count observed in surface waters from CA region *[65]*. Starting in 2023, the LGMAs implemented more stringent testing limits for overhead irrigated surface waters, setting the threshold at 235MPN/100ml generic *E. coli* for any single sample. However, we applied the less stringent testing limit historically used by LGMA (576MPN/100ml)

to account for agricultural waters in poorer conditions and to reflect the practices more relevant to the period for which the model validation data were available. However, our scenario analysis in Supplementary Table S3 predicted that adopting the more stringent testing limit led to 31.1% reduction in predicted illness risk compared to our baseline model, suggesting that the 2023 Arizona and California LGMAs may have a positive impact on leafy green safety. Our model included different irrigation systems (f_{irr} = overhead spray (O), furrow (F), or drip (D)) based on the field location (Table 1). Each system delivered varying volumes of irrigation water per irrigation event (g_{irr}), considering that most of the water (i.e., 51% of the total volume required for romaine growth) was applied in the last 30 days before harvest (W_O, W_F, W_D) *[44]*. The volume of water captured by romaine batches (v_{irr}) also differed for each system *[27,66]*. Overhead spray and furrow irrigation only occurred on sunny days (Bs_t) *[67]*, and differed in frequency: Overhead spray, applied once every 2-3 days, was reflected in probability Pr_O ($fO_t = Binomial(1, Pr_O)$) and furrow was applied about once every 10 days, reflected in probability Pr_F ($fF_t = Binomial(1, Pr_F)$). Drip irrigation occurred every day ($Pr_{D_t} = 1$) regardless of sunny or rainy weather. We used a dilution factor (DiF_{irr}) to simulate an even distribution of ECO157 across the field and spread of cells to a depth of 1cm. To align with typical harvest practices in the USA *[68]*, we introduced a waiting period ranging from 2 to 8 days before harvest ($BwaitP_t$) and terminated overhead spray and furrow irrigation during the selected waiting time. Shallowly buried drip lines were assumed for drip-irrigated fields, making them unaffected by the waiting period. The term Irr_{s_t} in the model (1) - (6), the ECO157 count in soil batch added via irrigation water on day *t*, was provided in Supplementary Table S2.

Soil splash via rain or irrigation (i.e., loss from soil) ($\boldsymbol{R_{splash_t}}, \boldsymbol{I_{splash_t}}$): The effect of irrigation and rain splash in the field was based on the QMRA model developed by Allende et al. (2018) *[27]*. Accordingly, splash of soil (and ECO157) to romaine would be facilitated either via irrigation water on a sunny day or rainwater on a rainy day. The maximum quantity of soil splashed by rain was assumed to be equal to that splashed by overhead spray irrigation. However, the quantity of soil splashed by rain was also proportional to the standardized monthly precipitation level ($Prec_{norm_m}$), reflecting the intensity of rain. We assumed that quantity of soil splashed on sunny days varied by different irrigation systems, with overhead spray irrigation causing the most soil to splash and drip irrigation causing the least (M_{irr}). The amount of soil splashed during irrigation was also affected by the intensity of irrigation system, where each system assumed to mimic an intermediate rainfall, equal to the median of the standardized precipitation $\left(Prec_{norm_m}\right)$. In case of a splash, a fraction of ECO157 in soil would transfer onto the romaine (LsR during rain or LsI during irrigation, Supplementary Table S1) via

transfer of contaminated soil. The terms for modeling rain (R_{splash_t}) and irrigation splash (I_{splash_t}) in the model (1) - (6) were derived as shown in Supplementary Table S2.

Daily decay of ECO157 (loss from soil) ($\boldsymbol{SNdecay_t, SPdecay_t}$): Daily inactivation of persister and non-persister cells in soil was constant. We used four different decay rates to represent non-persister *[36,56]* and persisters' *[57]* decay in BSAAO-amended soil and soil with no BSAAO application. We used a slower inactivation in BSAAO amended soil (R_{dman} for non-persisters and R_{dman_P} for persisters) and faster inactivation in soil with no BSAAO application (R_{ds} for non-persisters and R_{ds_P} for persisters). Persister cells decayed at a rate 10 times slower than that of their non-persister counterparts, based on the decay curve provided in Chen (2020) for ECO157 persisters in soil *[57]*. Decay parameters in soil ($SNdecay_t, SPdecay_t$) in the model (1) - (6) were explained in Supplementary Table S2.

Switch to persister cells (no change in the cell count in soil) ($\boldsymbol{SNtoSP_t}$): A small fraction of non-persister ECO157 cells were transitioned to persister cells on each day *[43,69,70]*. Persisters coexisted with non-persisters and comprised up to 0.01% of the population in soil environments *[69,70]*. The cells could also switch between the persister and non-persister states [*71]*. Although the characteristics and the formation mechanisms of persisters were previously reported as complex *[71]*, our model focused only on their enhanced ability to survive, which facilitated modeling biphasic decay. While the direct estimates for the switch rates in BSAAO-amended soil were unavailable, we used the study by Hofstrenge et al. (2013) *[70]*, which indicated a strong one-to-one correlation between switch rates ($Pmax_{man}, Pmax_S$) and the fraction of persisters ($NtoP_{man}, NtoP_S$). We assumed that the persister state was more prevalent in BSAAO-amended soil batches ($NtoP_{man}$). This assumption was based on the observation that pathogens regrew and recovered to a culturable state months after being introduced into soils through dairy BSAAO application *[72]*. This delayed recovery might suggest the initial presence of persister cells within the BSAAO that was applied on the field *[73]*. Details on the switch terms ($SNtoSP_t$) in the model (1) - (6) were provided in Supplementary Table S2.

Accumulation and loss of ECO157 in romaine batches

Given the rapid decay of ECO157 on field romaine *[27,39,74]*, we assumed that the initial ECO157 count on the field romaine was negligible one day prior to the start of the 14-day countdown period ($C_{L_0} = 0$). However, the sensitivity of model predictions to this assumption was tested in scenario analysis (see Supplementary Table S4).

Irrigation (accumulation on romaine): Contamination of romaine through irrigation followed a similar mechanism to that described for soil contamination. The volume of irrigation water captured by the romaine batches (v_{irr}) transferred a fraction of ECO157 cells from water to

the romaine plants (Tr_{ecirr}). The total volume of water delivered to the batch area (g_{irr}) was either captured by romaine or spread into the soil. Thus, we modelled the fraction of irrigation water delivered to the soil $\left(1 - \frac{v_{irr}}{g_{irr}}\right)$ and romaine ($\frac{v_{irr}}{g_{irr}}$). The term (Irr_{L_t}) describing the irrigation contamination in the model (1) - (6) was explained in Supplementary Table S2.

Runoff from a neighboring dairy farm and intrusion by feral swine (accumulation on romaine) ($\boldsymbol{Pig_{L_t}}$and $\boldsymbol{Runoff_{L_t}}$): The effect of wildlife and runoff contamination on romaine was modelled mechanistically similar to their effect in soil detailed previously. The difference was the proportion of feces/runoff that was spread in either the soil or on romaine, as determined by the parameter 'Pr_{LtoS}' (for soil: $1 - Pr_{pigLtoS}$, and for romaine: Pr_{LtoS}). The feces/runoff was modelled as spreading uniformly through the romaine batches. The terms $Pig_{L_t}$and $Runoff_{L_t}$ describing the ECO157 count in romaine batches via feral swine intrusion and runoff events in the model (1) - (6) were explained in Supplementary Table S2.

Soil splash via rain or irrigation (accumulation on romaine)($\boldsymbol{R_{splash_t}}$, $\boldsymbol{I_{splash_t}}$): The daily loss of ECO157 from the soil due to rain or irrigation splash was applied to a given romaine batch, resulting in ECO157 accumulation on romaine batches. The transferred ECO157 cells were distributed uniformly in a given romaine batch (Supplementary Table S2).

Daily decay of ECO157 (loss from romaine) ($\boldsymbol{LNdecay_t}$, $\boldsymbol{LPdecay_t}$): For non-persister ECO157 decay on romaine ($LNdecay_t$, Supplementary Table S2), the daily decay rate (R_{dL}) followed the month-specific solar decay pattern (R_{sun_m}) throughout the year *[27]*. Specifically, a slower decay rate, reflecting winter conditions, was applied from December to April (for production in Y and DR regions), while a faster decay rate, reflecting summer conditions, was implemented from May to October (for production in Y region). Additionally, the parameter 'h_{sun_m}', which indicated the average daily sunlight hours in a given month, was determined based on weather data specific to both field location and the time of romaine harvest. The decay for persisters ($LPdecay_t$, Supplementary Table S2) was assumed 10 times slower than the non-persisters *[43,57]*.

Switch to persister cells (no change in the cell count on romaine) ($\boldsymbol{LNtoLP}$): The switch rate from non-persister to persister cells on romaine ($LNtoLP_t$, Supplementary Table S2) was constant and established based on the data from the study by Munther et al. (2020) *[43]*.

Harvest ($\boldsymbol{CC_{blade}}$): After modeling the population dynamics of ECO157 in soil and on romaine plants during the 14-day countdown period (equations (1) - (6)), we proceeded with modeling

the harvesting process, which accounted for cross-contamination of romaine from soil caused by contaminated coring knife blades *[15,38]*.

Postharvest model

The bacterial population dynamics outlined by Pang et al. (2017), Danyluk and Schaffner (2011) and Nauta (2005) served as the baseline for the postharvest practices (i.e., cold storage, processing (chlorine wash), retail, transportation and consumer stages) *[32,38,42]*.

Cold storage before chlorine wash ($\boldsymbol{CC_{coldst}}$): The harvested romaine was transported to cold storage, where it was kept for a period ranging from 1 to 48 hours (t_{coldst}). In this stage, we used 5°C as the minimum temperature for ECO157 growth. Time spent (t_{coldst}) above 5°C increased ECO157 counts (growth rate, G_{coldst}) and time spent ≤ 5°C decreased ECO157 counts (decay rate, D_{coldst}) *[38,75]*. Cold storage temperature was assumed to follow a triangular distribution centered around the standard refrigeration temperature, allowing inadequate storage ($T_{endcoldst}$). We assumed that the temperature of romaine at the time of harvest ($T_{harvest}$) was equivalent to the air temperature during the harvesting process, under the assumption that harvesting occurs early in the morning *[76]*. Accordingly, we used field location- and month-specific weather data for determining the harvesting temperature. CC_{coldst} in Table 2 represented the change in ECO157 count on romaine batches during cold storage.

Chlorine wash ($\boldsymbol{CC_{w_{agent}}}$): The washing step was assumed to reduce ECO157 count by 0.52 Log_{10}CFU/gram romaine on average, with a range of reduction between 0.01 Log_{10}CFU/gram romaine and 1 Log_{10}CFU/gram romaine (Table 2) *[77]*.

Partitioning of the romaine batches into packages ($\boldsymbol{Sc_{clstrb}}, \boldsymbol{Cluster_{parb}}$): The partitioning of a chlorine-washed romaine batches (C_{wash}) into individual packages included potential clustering of ECO157 cells (or aggregation of cells) *[42,78]*. To estimate the ECO157 count within a contaminated romaine package (C_{pack}), we used the following equations:

$$C_{pack} = Log_{10}\left(Binomial\left(C_{wash},\ P_{contpack}\right) + 1\right) \tag{7}$$

where

$$P_{contpack} = \frac{1-Sc_{clstrb}}{A_{pack}} + Sc_{clstrb} \times Beta(Cluster_{parb}, Cluster_{parb} \times (A_{pack} - 1)))$$

$$P0_{pack} = EXP\left(\frac{\Gamma\left(Cluster_{parb} \times A_{pack}\right) \times \Gamma\left(Cluster_{parb} \times \left(A_{pack} - 1\right) + C_{wash}\right)}{\Gamma\left(Cluster_{parb} \times \left(A_{pack} - 1\right)\right) \times \Gamma\left(Cluster_{parb} \times A_{pack} + C_{wash}\right)}\right) \tag{8}$$

Here, $P_{contpack}$ was an intermediary parameter used to determine the ECO157 count within a given package, accounting for the presence of clustering (Sc_{clstrb}; set to 0 when clustering was absent and 1 when present) and the degree of clustering when it occurred (with maximum clustering when $Cluster_{parb} = 0$, decreasing as $Cluster_{parb}$ approaches positive infinity). In the baseline model, we assumed no clustering ($Sc_{clstrb} = 0$), and as a result, the ECO157 cells from the romaine batches were evenly distributed to each package ($\frac{1}{A_{pack}}$). The '$P0_{pack}$' parameter in Table 2 (equation (8)) determined the probability of an uncontaminated package *[42]*. We derived the probability of a package being contaminated as $(1 - P0_{pack})$.

Retail storage, transport from retail to home, and consumer ($\boldsymbol{CC_{retail}}$, $\boldsymbol{CC_{trspt}}$, $\boldsymbol{CC_{home}}$): We used the same growth and decay approach implemented in cold storage process to model retail storage, transport from retail to home, and consumer steps (i.e., home storage) albeit with stage specific time and temperature parameters *[38,75]*. The values of the time and temperature parameters were previously parameterized by Pang et al. (2017) *[38]*.

Partitioning of the romaine package into servings ($\boldsymbol{Sc_{clstrp}}$, $\boldsymbol{Cluster_{parp}}$): We adopted a serving size of 85g for romaine (Ser) based on reference serving size provided by USDA (2015) *[98]*. To distribute ECO157 cells from romaine packages to servings, we followed the same approach used for partitioning romaine batches to packages (equations (7) - (8)) *[42]*.

Dose-response

Following the partitioning process, the ECO157 count (C_{serW}) and prevalence (P_{ser}) in romaine servings were estimated. The C_{serW} was used in a Beta-Poisson dose-response model ($Pill$) *[32,79]* (Table 2). We estimated our main outcome, the predicted number of illness cases per year in the USA from consuming ECO157-contaminated romaine servings ($Ncase_{US}$), by using the dose-response model along with the number of contaminated servings consumed annually in the USA ($Ncsc_{US}$).

Expert elicitation: Several parameters across the farm-to-fork chain, as well as specific interventions for the scenario analysis, were determined by using the input from study authors and industry experts who served on an advisory council for the USDA-NIFA-SCRI grant #2019-51181-30016. Three informal interviews were conducted with the advisory council members, who were from produce production and processing, distribution and retail, and sanitation industries. Additionally, an online survey was created in Qualtrics (Provo, UT, USA) and administered to the advisory council (12 respondents) to identify the intervention strategies of greatest interest to industry. Based on the feedback, (i) cattle vaccination, (ii) transition from overhead and furrow to drip irrigation, and (iii) transition from overhead to furrow irrigation,

were the most requested intervention strategies and were included in the scenario analysis. Further, in another online survey with the council conducted via Qualtrics, we presented the scope of the risk assessment model, including contamination sources and interventions.

Model validation: The main model outcome was validated with the information from the CDC for the year 2011 (most recent data). Since the most recent data was from 2011, we used the 2011 USA population data ($USpop$) *[80]* along with 2011 USA fresh romaine and leaf lettuce consumption data (consumer weight without accounting for the food loss at consumer level) *[81]*. The validation estimated the USA foodborne diseases associated with leafy greens contaminated with ECO157 *[45]*. Accordingly, the estimated number of foodborne diseases due to ECO157 contamination in the USA averages 63,153 cases annually, with 90% credible interval from 17,587 to 149,631. Of these, 19.3% - 31.5% are attributed to leafy greens *[46]*, resulting in an average of 12,189 – 19,893 cases each year in the USA linked to consuming contaminated leafy greens. Considering the 90% credible interval, the true range of illness cases attributable to consumption of romaine could be between 3,394 and 47,134.

Sensitivity and uncertainty analysis: We evaluated the effect of variable and uncertain parameters in the baseline model to determine the key factors influencing the main outcome (the predicted number of illnesses in the USA). We used @RISK and R (*dplyr* package) for the sensitivity and uncertainty analyses (for preharvest and postharvest models) and generated tornado plots of Spearman rank-order correlation coefficients (SRCCs) *[82]*. The SRCC analysis for the baseline preharvest model was based on a set of core and derived parameters, where derived parameters were calculated from core parameters. However, no parameter was included in the analysis more than once. This approach was taken to ensure that the results were more interpretable and informative. More information about the selection of parameters was provided in Supplementary Table S5. Parameters with an absolute value of the SRCC $|\rho| < 0.05$ were considered negligible and were excluded from the tornado plots. Likewise, only parameters with statistically significant SRCC were presented in graphs, adhering to a Bonferroni-corrected alpha threshold of 0.05 divided by the number of evaluated parameters (21 for preharvest, 18 for postharvest and 11 for uncertainty analysis in preharvest). The absolute value of the SRCC, or $|\rho|$, was interpreted according to its common usage in the literature (ρ=0 to <0.1, none; ρ=0.1 to <0.3, poor; ρ=0.3 to <0.6, fair; ρ=0.6 to <0.8, moderate; ρ=0.8 to 1.0 very strong correlation) *[83,84]*. In the sensitivity analysis of variable parameters, all uncertain parameters were set to their fixed deterministic values. A parameter was classified as an uncertainty parameter when no information was available in the literature, requiring assumptions. In the uncertainty analysis of the preharvest model, we fixed the variable parameters to the median values of their respective distributions and created a distribution for each of the uncertain parameters by establishing a triangular distribution. The minimum of

each triangular distribution was set as a 50% reduction from the baseline, and the maximum was set as a 50% increase from the baseline for all uncertain parameters, with the original baseline value as the mean/mode. In the preharvest model, this definition applied to the following parameters: (i) volume of irrigation water captured by romaine through furrow or drip irrigation per irrigation event (V_F, V_D); (ii) fraction of soil batch transferred to romaine batches via splash caused by furrow or drip irrigation (Tr_{man_F}, Tr_{man_D}); (iii) dilution factors used for the distribution of irrigation water, feces and runoff in soil batch ($DiF_{irr}, DiF_{pig}, DiF_{cow}$); (iv) switch rate from non-persister to persister cells in BSAAO-amended soil ($NtoP_{man}$); (v) ECO157 prevalence (P_{irr_w}) in well water; and (vi) transfer of ECO157 cells from irrigation water to romaine batches in an irrigation event (Tr_{ecirr}).

Regression trees: Tree analyses were conducted for both preharvest and postharvest models using the *rpart* package in R *[85]*. All analyses were implemented at default settings in *rpart.control* function. To prevent tree overfitting, we used 10-fold cross-validation *[78]*. Similar to the SRCC analysis, the regression tree for the baseline preharvest model was based on the set of the core and derived parameters previously determined (Supplementary Table S5).

Scenario analysis of risk management strategies: Based on the sensitivity and uncertainty analyses, as well as the results of expert elicitation, we identified nine risk management strategies for evaluation in the scenario analysis (Table 4). These scenarios are (i) cattle vaccination, (ii - iv) treatment of surface water, used for overhead spray irrigation, by either chlorine, peracetic acid or ultraviolet radiation, (v) transition from overhead to furrow irrigation system, (vi) transition from overhead and furrow to drip irrigation, (vii) chlorine wash alternative during postharvest washing, (viii) temperature reduction during retail, and (ix) consumer wash (the modifications to the parameters to apply to these scenarios were provided in Table 4). In addition, we evaluated the effect of parameters related to clustering of ECO157 cells during the partitioning process (i.e., romaine batch to package and package to serving) as an additional scenario (Table 4). The effect of each risk management strategy or clustering was assessed based on their impact on the main outcome.

Table 1. Parameters in the preharvest *Escherichia coli* O157:H7 (ECO157) model

Parameter	Description (units)	Distribution/Value	Mean (5th; 95th percentile)	Source
A_{field}	Area of the field (approximately 1-acre field) (cm^2)	$6245 \times 6245 = 39000025$	-	*[36]*
A_{batch}	Area of a field that grows a romaine batch ("batch area", 1/12 of the field area) (cm^2)	$6245 \times 6245/12 = 3247400$	-	*[36]*
N_{plant}	Number of plants harvested from the field (romaine plants)	26,000	-	*[36]*

$Plant_{width}$	The area that a single mature romaine plant occupied on the field, length x width of one plant (cm^2)	$25 \times 25 = 625$	-	*[86]*
$Plant_{space}$	The area allocated for one romaine plant in the field (cm^2)	$30 \times 50 = 1500$	-	*[36]*
$Plant_{barea}$	Number of plants grown in a batch area (romaine plants)	$17 \times 125 = 2167$	-	*[36]*
$Plant_{wght}$	Weight of one mature romaine plant (g)	300	-	Assumption [c]
f_{loc}	Select geographic location where the romaine batches were grown. Notations: CC: Coastal Region and Central Valley, CA DR: Desert (Imperial) Region, CA Y: Yuma Region, AZ	$Discrete(\{CC, DR, Y\}, \{0.6,0.1,0.3\})$	-	*[47,48,49]*
f_{month}	Select the month of harvest for the given geographic location. Notations: January: 1, February: 2, March: 3, April: 4, May: 5, June: 6, July: 7, August: 8, September: 9, October: 10, November: 11, December: 12	$f_{month}(f_{loc} = CC) = DUniform(\{4,5,6,7,8,9,10,11\})$ $f_{month}(f_{loc} = DR) = DUniform(\{1,2,3,12\})$ $f_{month}(f_{loc} = Y) = DUniform(\{1,2,3,12\})$	-	*[47,48]*
f_w	Select the source of irrigation water. Notations: S: Surface water W: Well water	$Discrete(\{s, w\}, \{0.11,0.89\})$	-	*[63]*
f_{irr}	Select the irrigation type for the given geographic location. Notations: O: Overhead spray irrigation F: Furrow irrigation D: Drip irrigation	$f_{irr}(f_{loc} = CC) = Discrete(\{O, F, D\}, \{0.4,0,0.6\})$ $f_{irr}(f_{loc} = DR) = Discrete(\{O, F, D\}, \{0.2,0.5,0.3\})$ $f_{irr}(f_{loc} = Y) = Discrete(\{O, F, D\}, \{0.08,0.92,0\})$	-	Author opinion [c]
$Modeldays$	Countdown period before harvest, modeled via a set of difference equations (days)	14	-	*[39]*
$edge$	Batch area was at the edge of the field and thus could be contaminated by runoff: yes=1, no=0	$Binomial\left(1, \frac{1}{12}\right)$	0.08 (0; 1)	Calculated [c]
C_{man}	ECO157 count per g of BSAAO (CFU/g)	$10^{Uniform(3.1,8.4)}$	2.0×10^7 (2.0×10^3; 1.3×10^8)	*[36,55]*
Sc_{vac}	Cattle vaccination intervention against ECO157: no=0 (default), yes=1	0	-	*[87]*
Rh_{con}	Volume of 1g BSAAO/swine feces/cattle feces (cm^3/g)	2.5	-	*[58]*
Dh_{con}	Depth of tiling on the field (cm)	$Uniform(15,20)$	17.50 (15.25; 19.75)	*[59]*
M_{man}	Mass of BSAAO applied on the batch area (g)	$\frac{Uniform(5000000,15000000)}{12}$	8.3×10^5 (4.6×10^5; 1.2×10^6)	*[88]*
DiF_{man}	Dilution of ECO157 cells in BSAAO in the volume of the tilled soil batch layer (unitless)	$\frac{M_{man} \times Rh_{con}}{A_{batch} \times Dh_{con}}$	-	Calculated [c]
Pr_{man}	BSAAO was used: yes=1, no=0	0.1	-	Author opinion [c]
P_{man}	BSAAO applied on a field was contaminated with ECO157: yes=1, no=0	$Binomial(1, Pert(0,0.17,0.21))$	0.15 (0; 1)	*[55]*

C_{man_P}	ECO157 count in any BSAAO applied soil, which returns 0 CFU for a non-contaminated soil and C_{man} for a contaminated soil (CFU/g)	$P_{man} \times Pr_{man} \times C_{irr}$	-	Calculated [c]
P_{irr_S}	ECO157 prevalence in surface water (unitless)	$Pert(0,0.08,0.21)$	$0.09\ (0.03; 0.16)$	*[8]*
P_{irr_W}	ECO157 prevalence in well water (unitless)	0.001	-	Assumption [c]
P_{irr}	Irrigation water contaminated with ECO157 applied during the countdown period: yes=1, no=0	$Binomial(1, P_{irr})$, where $P_{irr}(f_w = W) = P_{irr_W}$ $P_{irr}(f_w = S) = P_{irr_S}$	$0.04\ (0; 0)$	Calculated [c]
Sc_{wt}	Water treatment intervention applied prior to the overhead spray irrigation (O) using surface water (S): None=0 (default), Ultra-Violet (UV)=1, Peracetic Acid (PAA)=2, Chlorine=3	0	-	Selected [c]
C_{irr_S}	ECO157 count per ml of contaminated surface water post treatment (CFU/ml)	$\frac{10^{Pert(0.1,0.76,1.77)+\ Normal(-1.9,0.6,Truncate(,0))}}{10^{CC}}$ where $CC(Sc_{wt} = 0) = 0$ $CC(Sc_{wt} = 1) = CC_{UV}$ $CC(Sc_{wt} = 2) = CC_{PPA}$ $CC(Sc_{wt} = 3) = CC_{Ch}$	$0.08\ (0.007; 1.09)$	*[19,20,65]*
C_{irr_W}	ECO157 count per ml of contaminated well water (CFU/ml)	$10^{0.05}$	-	*[19,20,34]*
C_{irr}	ECO157 count per ml of contaminated irrigation water applied on day t (CFU/ml)	$C_{irr}(f_w = W) = C_{irr_W}$ $C_{irr}(f_w = S) = C_{irr_S}$	-	Calculated [c]
C_{irr_P}	ECO157 count on any irrigation water, which returns 0 CFU for a non-contaminated water and C_{irr} for a contaminated water (CFU/ml)	$P_{irr} \times C_{irr}$	-	Calculated [c]
$type_{irr}$	Probability the selected irrigation type was occurring on a given day (unitless)	$type_{irr}(f_{irr} = O) = Prp_O$ $type_{irr}(f_{irr} = F) = Prp_F$ $type_{irr}(f_{irr} = D) = Prp_D$	-	Calculated [c]
Pr_O	Probability of overhead irrigation on a given day (unitless)	0.4	-	*[44]*
Pr_F	Probability of furrow irrigation on a given day (unitless)	0.1	-	Author opinion [c]
Pr_D	Probability of drip irrigation on a given day (unitless)	1	-	Author opinion [c]
W_O	Volume of water applied through overhead irrigation to a romaine batch per an irrigation event (ml/day)	$\frac{Uniform(1850222756,2466963675) \times 0.51}{30 \times Pr_O \times 12}$	$7.7 \times 10^6\ (6.7 \times 10^6; 8.6 \times 10^6)$	*[44]*
W_F	Volume of water applied through furrow irrigation to a romaine batch per an irrigation event (ml/day)	$\frac{Uniform(2466963675,3083704593) \times 0.51}{30 \times Pr_F \times 12}$	$3.9 \times 10^7\ (3.5 \times 10^7; 4.3 \times 10^7)$	*[44]*
W_D	Volume of water applied through drip irrigation to a romaine batch per an irrigation event (ml/day)	$\frac{Uniform(1233481837,1850222756) \times 0.51}{30 \times Pr_D \times 12}$	$2.2 \times 10^6\ (1.8 \times 10^6; 2.6 \times 10^6)$	*[44]*
$Daytogrow$	Days from seeding to harvest (days)	$Daytogrow(f_{loc} = CC) = Uniform(65,80)$ $Daytogrow(f_{loc} = DR) = 130$ $Daytogrow(f_{loc} = Y) = 130$	$72\ (65; 79)$ - -	*[44]*
D_{man}	Number of days before BSAAO application (days)	$Daytogrow$	-	*[19,20]*

g_{irr}	Volume of water applied per an irrigation event in a batch area (ml/day)	$g_{irr}(f_{irr} = O) = W_O$ $g_{irr}(f_{irr} = F) = W_F$ $g_{irr}(f_{irr} = D) = W_D$	-	Calculated [c]
V_O	Volume of irrigation water captured per g of romaine per an overhead spray irrigation event (ml/g/day)	$Uniform(1.8,21.6)$	11.7 (2.8; 20.6)	*[27]*
V_F	Volume of irrigation water captured per g of romaine per a furrow irrigation event (ml/day)	100	-	Assumption [c]
V_D	Volume of irrigation water captured per g of romaine per a drip irrigation event (ml/day)	10	-	Assumption [c]
v_{irr}	Volume of irrigation water captured by romaine batch per an irrigation event (ml/day)	$v_{irr}(f_{irr} = O) = V_O \times Plant_{barea} \times Plant_{wght}$ $v_{irr}(f_{irr} = F) = V_F$ $v_{irr}(f_{irr} = D) = V_D$	-	Calculated [c]
Tr_{man_O}	Fraction of the top 1cm layer of soil batch transferred to a romaine batch via splash caused by overhead irrigation per irrigation event (unitless)	$BetaGeneral(0.4,0.8,0.05,16.4) \times Plant_{barea} \times \frac{Plant_{wght}}{5} \times \frac{Rh_{con}}{A_{batch} \times 1}$	0.05 (0.0005; 0.13)	*[67]*
Tr_{man_F}	Fraction of the top 1cm layer of soil batch transferred to a romaine batch via splash caused by furrow irrigation per irrigation event (unitless)	0.0005	-	Assumption [c]
Tr_{man_D}	Fraction of the top 1cm layer of soil batch transferred to a romaine batch via splash caused by drip irrigation per irrigation event (unitless)	0.0001	-	Assumption [c]
M_{irr}	Mass of soil transferred to a romaine batch via splash per irrigation event (g/day)	$M_{irr}(f_{irr} = O) = Tr_{man_O}$ $M_{irr}(f_{irr} = F) = Tr_{man_F}$ $M_{irr}(f_{irr} = D) = Tr_{man_D}$	-	Calculated [c]
Tr_{ecirr}	Fraction of ECO157 cells transferred from irrigation water to romaine (unitless)	0.1	-	Assumption [c]
DiF_{irr}	Dilution factor for ECO157 cells in irrigation water in the top 1cm layer of soil batch (unitless)	$\frac{g_{irr} \times \left(1 - \frac{v_{irr}}{g_{irr}}\right)}{(A_{batch} \times 1)}$	-	Calculated [c]
LN_0	ECO157 count in a romaine batch at the start of the 14-day countdown period (t=0) (CFU)	$LN_0(f_{irr} = O) = 0$ $LN_0(f_{irr} = F) = 0$ $LN_0(f_{irr} = D) = 0$	-	Assumption [c]
CC_{UV}	Log_{10} ECO157 count reduction of surface water contamination by UV treatment before overhead irrigation (Log_{10}CFU)	$Uniform(1.30,2.91)$	2.10 (1.38; 2.83)	*[89,90]*
CC_{PAA}	Log_{10} ECO157 count reduction of surface water contamination by PAA treatment before overhead irrigation (Log_{10}CFU)	$Uniform(2.48,3.09)$	2.78 (2.51; 3.06)	*[90,91]*
CC_{Ch}	Log_{10} ECO157 count reduction of surface water contamination by Chlorine treatment before overhead irrigation (Log_{10}CFU)	$Uniform(1.00,2.78)$	1.89 (1.09; 2.69)	*[90]*
$C_{cow_{prep}}$	Log_{10} ECO157 count per g of fresh cattle feces (Log_{10}CFU/g)	$Pert(0.89,3.08,8.4)$	3.60 (1.58; 6.04)	*[36]*

$C_{cow_{tot}}$	Log_{10} ECO157 count per g of fresh cattle feces accounting for the effect of ECO157 vaccine intervention, if implemented (Log_{10}CFU/g)	$C_{cow_{tot}} = Sc_{vac} \times Max(C_{cow} - 1,0) + (1 - Sc_{vac}) \times C_{cow}$	-	Calculated [c]
C_{cow}	ECO157 count per g of fresh cattle feces (CFU/g)	$10^{C_{cow_{tot}}}$	-	Calculated [c]
C_{cow_P}	ECO157 count on any cattle feces, which returns 0 CFU for non-contaminated feces and C_{cow} for contaminated feces (CFU/g)	$C_{cow} \times P_{cow_m}$	-	Calculated [c]
C_{pig}	Log_{10} ECO157 count per g of fresh swine feces (Log_{10}CFU/g)	$Pert(0.89,3.59,7.00)$	3.71 (1.86; 5.65)	*[36]*
C_{pig_P}	ECO157 count on any swine feces, which returns 0 CFU for non-contaminated feces and C_{pig} for contaminated feces (CFU/g)	$C_{pig} \times P_{pig_m}$	-	Calculated [c]
Pr_{pig}	Probability of feral swine entering the field in a day (unitless)	0.00405	-	*[6]*
$Pr_{pig_{countt}}$	Wildlife intrusion occurred on day t: yes=1, no=0	$Binomial\ (1, Pr_{pig})$	0 (0; 1)	Calculated [c]
$Pr_{pigCountSu}$	Total number of days during the 14-day countdown when wildlife intrusion occurred	$\sum_t Pr_{pig_{countt}}$	-	Calculated [c]
$Mdef_{pig}$	Mass of feces excreted by a feral swine per day (g swine feces/day)	4,265	-	*[36]*
$Tdef_{pig}$	Number of times a feral swine defecated per day (times/day)	4	-	*[62]*
$Ndef_{pig}$	Number of times feral swine (between 1-5 swine) defecated on the field (times/day)	$Poisson(1, Truncate(1,20))$	1.58 (1; 3)	Assumption [c]
M_{pig}	Mass of feces excreted by feral swine intruding the batch area (g)	$\frac{Mdef_{pig}}{Tdef_{pig}} \times Ndef_{pig}$	-	Calculated [c]
DiF_{pig}	Dilution factor for ECO157 cells in feral swine feces in soil batch associated with plants near which the swine defecated (unitless)	$\frac{M_{pig} \times Rh_{con}}{A_{batch} \times 1}$	-	Calculated [c]
M_{cow}	Mass of cattle runoff (carrying fresh cattle feces only) introduced into the edge batch area soil per runoff event (g/day)	$Triang(5000,10000,15000)$	10000 (6581; 13418)	*[36]*
DiF_{cow}	Dilution factor for ECO157 cells in cattle runoff introduced into in the top 1cm layer of soil batch (unitless)	$\frac{M_{cow} \times Rh_{con}}{A_{batch} \times 1}$	-	Calculated [c]
Pr_{LtoS}	Proportion of fecal matter (swine and runoff) landing on romaine rather than in soil (unitless)	$\frac{Plant_{width}}{Plant_{space} - Plant_{width}}$	-	*[36]*
Tr_{ecsoil}	Proportion of ECO157 cells in soil transferring to romaine during a rain or irrigation event (unitless)	$Uniform(0.00004,0.63)$	0.315 (0.032; 0.600)	*[92]*
R_{dman}	ECO157 decay rate for non-persisters in BSAAO amended soil (1 Log_{10}/day)[a]	0.038	-	*[56]*
R_{dL}	ECO157 decay rate for non-persisters on romaine in field (1/day) [a,b]	$1 - 10^{h_{sun_m} \times \frac{R_{sun_m}}{24}}$	-	*[27,34]*
R_{ds}	ECO157 decay rate for non-persisters in soil (1/day) [a]	$1 - 10^{-0.1744}$	-	*[36]*

R_{dman_P}	ECO157 decay rate for persisters in BSAAO amended soil (1 Log_{10}/day) [a]	$\frac{R_{dman}}{10}$	-	*[56,57]*
R_{dL_P}	ECO157 decay rate for persisters on romaine in field (1/day) [a]	$\frac{1-10^{h_{sun_m}\times\frac{R_{sun_m}}{24}}}{10}$	-	*[43]*
R_{ds_P}	ECO157 decay rate in soil for persisters (1/day) [a]	$\frac{R_{ds}}{10}$	-	*[57]*
$NtoP_{man}$	Switch rate from normal to persister in BSAAO amended soil (1/day) [a]	0.001	-	*[69,70]*
$NtoP_L$	Switch rate from normal to persister on romaine (1/day) [a]	0.0004	-	*[43]*
$NtoP_S$	Switch rate from normal to persister in soil (1/day) [a]	0.0001	-	*[69,70]*
$Pmax_{man}$	Maximum fraction of persisters in BSAAO amended soil (unitless)	$\frac{NtoP_{man}}{10^{R_{dman}}}$	-	*[43]*
$BwaitP$	Waiting time: The period before harvest during which irrigation was stopped (days)	$Triang(2,4,8)$	$4\ (2;6)$	*[68]*
t	Day of the countdown period in the difference equation model (1)-(6)	$1 \leq\ t\ \leq\ Modeldays$	-	Calculated [c]
Tr_{blade}	Average amount of soil attached to a harvesting blade after contact with soil in one cross contamination event (g soil/blade)	10.22	-	*[15]*
$Tr_{ecblade}$	Proportion of ECO157 cells transferred from a harvesting blade to a romaine plant in one cross contamination event (unitless)	0.0013	-	*[15]*
N_{blade}	ECO157 count in soil attached on blades (CFU/soil batch x soil batch/blade=CFU/blade)	$\frac{C_{S14}\times Tr_{blade}\times Rh_{con}}{1\ \times\ A_{batch}}$	-	Calculated [c]
CC_{blade}	ECO157 cells transferred from harvesting blades to a romaine batch (CFU/romaine batch)	$Tr_{ecblade}\times N_{blade}\times Plant_{barea}\times Plant_{wght}\times\frac{1}{3}$	-	Calculated [c]
C_{L_h}	Preharvest outcome: ECO157 count on harvested romaine batch (CFU)	$C_{L_{14}}+\ CC_{blade}$	-	Calculated [c]

[a] In the difference equation model the rate parameter indicates the proportionate change in ECO157 population per time step.
[b] See Table 3 for more details.
[c] Assumption: Parameter value accepted as true with little to no information; Author opinion: Parameter value accepted as true based on authors' experience; Calculated: Parameter value calculated based on other parameters; Selected: Used in scenario analysis (See Supplementary Table S6 for details for their use.)

Table 2. Parameters in the postharvest *Escherichia coli* O157:H7 (ECO157) model

Parameter	Description (units)	Distribution/Value	Mean (5th-95th percentile)	Source
C_{Lh_f}	*E. coli* O157:H7 (ECO157) count in a contaminated harvested romaine batch (Log_{10}CFU/romaine batch)	$Triang(-0.66,3.81,6.36)$	$3.17\ (0.59;5.41)$	Output from preharvest model fitted in @RISK

P_{Lh_f}	ECO157 prevalence in a harvested romaine batch (unitless)	$Triang(0.005,0.01,0.02)$	$0.01\ (0.007; 0.02)$	Output from preharvest model fitted in @RISK
t_{coldst}	Cold storage time (h)	$Triang(1,24,72)$	$32\ (10; 59)$	Author opinion [a]
$T_{endcoldst}$	Temperature after cold storage (°C)	$Triang\left(0,4.44,T_{harvest_m}\right)$	$5\ (4.3; 5.7)$	*[33]*
T_{coldst}	Temperature during cold storage (°C)	$\frac{1}{2} \times \left(T_{endcoldst} + T_{harvest_m}\right)$	-	*[38]*
G_{coldst}	Cold storage growth rate (Log_{10}CFU/h)	$\frac{\left(0.023 \times (T_{coldst} - 1.2)\right)^2}{2.303}$	-	*[38]*
D_{coldst}	Cold storage die-off rate (Log_{10}CFU/h)	$\frac{Lognorm(0.013,0.001,Shift(0.001))}{2.303}$	$0.006\ (0.005; 0.007)$	*[38]*
$GorD_{coldst}$	Growth (G) or die off (D) based on cold storage temperature (Log_{10}CFU/h)	$G, if\ T_{coldst} > 5,$ $D, if\ T_{coldst} \leq 5$	-	*[38]*
CC_{coldst}	Change in ECO157 count during cold storage (Log_{10}CFU)	$CC_{coldst}(GorD_{coldst} = G) = G_{coldst} \times t_{coldst}$ $CC_{coldst}(GorD_{coldst} = D) = -D_{coldst} \times t_{coldst})$	-	Calculated [a]
$C_{endcoldst}$	ECO157 count on a contaminated romaine batch after cold storage, going to processing (CFU/romaine batch)	$\text{Log}_{10}\left(ROUNDDOWN\left(10^{Max\{C_{Lh_f} + CC_{coldst},0\}}, 0\right)\right)$	-	Calculated [a]
B	Romaine batch size (g)	$Plant_{barea} \times Plant_{wght}$	-	*[93]*
Sc_{wash}	1 or 2 based on the scenarios below, (default=1)	1	-	Selected [a]
CC_{ch20}	Log_{10}CFU ECO157 reduction with chlorine washing (Log_{10}CFU/g romaine)	$Pert(0.01,0.52,1.03)$	$0.52\ (0.20; 0.84)$	*[77]*
CC_{ch30}	Chlorine wash alternative with a median of 3.22 Log_{10}CFU reduction instead of 0.52 in the baseline chlorine wash (Log_{10}CFU)	$Pert(3.04,3.22,3.4)$	$3.22\ (3.11; 3.33)$	*[94]*
$CC_{w_{agent}}$	ECO157 reduction associated with the scenario selected for the washing step (Log_{10}CFU)	$CC_{w_{agent}}(Sc_{wash} = 1) = CC_{ch20}$ $CC_{w_{agent}}(Sc_{wash} = 2) = CC_{ch30}$	-	Selected [a]
C_{wash}	ECO157 count on a contaminated romaine batch after washing (CFU/romaine batch)	$\text{Log}_{10}(ROUNDDOWN(\ Max\{10^{C_{endcoldst} - CC_{w_{agent}}},0\},0))$	-	Calculated [a]
$P0_{pack}$	Probability a non-contaminated package is partitioned from a contaminated romaine batch (unitless)	$EXP\left(\frac{\Gamma(Clstr_{par} \times A_{pack}) \times \Gamma(Clstr_{par} \times (A_{pack} - 1) + C_{wash})}{\Gamma\left(Clstr_{par} \times (A_{pack} - 1)\right) \times \Gamma(Clstr_{par} \times A_{pack} + C_{wash})}\right)$	-	*[42]*
P_{pack}	Probability of ECO157 contamination on romaine in a package (unitless)	$(1 - P0_{pack}) \times P_{Lh_f}$	$0.003\ (2.5 \times 10^{-5}; 0.016)$	Calculated [a]
Sc_{clstrb}	0 if no clustering in romaine batch (default), 1 if clustering	0	-	Selected [a]
$Clstr_{parb}$	Clustering parameter	1	-	*[42]*
$W_{lettucepack}$	Weight of shredded romaine per package (8 oz) (g)	225	-	

A_{pack}	Number of packages in a romaine batch (packages)	$\frac{B}{W_{lettucepack}}$	-	Calculated [a]
$P_{contpack_{dist}}$	Distribute ECO157 cells to a package when clustering	$Beta\left(Clstr_{parb}, Clstr_{parb} \times \left(A_{pack} - 1\right)\right)$	3.5×10^{-4} (1.8 $\times 10^{-5}$; 0.001)	Calculated [a]
$P_{contpack}$	Proportion of ECO157 cells from a contaminated romaine batch in a package (unitless)	$(1 - Sc_{clstrb}) \times \left(\frac{1}{A_{pack}}\right) + Sc_{clstrb} \times P_{contpack_{dist}}$	-	*[42]*
C_{pack}	ECO157 count in a contaminated package (Log_{10}CFU/pack romaine)	$Log_{10}\left(Binomial\left(C_{wash}, P_{contpack}\right) + 1\right)$	-	Calculated [a]
t_{retail}	Retail storage time (h)	$Triang(0.5,4,7) \times 24$	92.0 (37.6; 144.3)	*[38]*
T_{retail}	Retail storage temperature (°C)	$Normal\left(4.4441,2.9642, Truncate(0,20.56)\right)$	4.9 (0.9; 9.4)	*[75,95]*
G_{retail}	Retail growth rate (Log_{10}CFU/g romaine/h)	$\frac{\left(0.023 \times \left(T_{trnspt} - 1.2\right)\right)^2}{2.303}$	-	*[38]*
D_{retail}	Retail die-off rate (Log_{10}CFU/g romaine/h)	$\frac{Lognorm\left(0.013,0.001, Shift(0.001)\right)}{2.303}$	0.006 (0.005; 0.007)	*[75]*
$GorD_{retail}$	Growth (G) or die off (D) based on retail temperature (Log_{10}CFU/h)	$G, if\ T_{retail} > 5,$ $D, if\ T_{retail} \leq 5$	-	*[38]*
CC_{retail}	Change in ECO157 count during retail storage (Log_{10}CFU)	$CC_{retail}(GorD_{retail} = G) = G_{retail} \times t_{retail}$ $CC_{retail}(GorD_{retail} = D) = -D_{retail} \times t_{retail}$	-	*[38]*
C_{retail}	ECO157 count in a contaminated package after retail storage (Log_{10}CFU/package romaine)	$\text{Log}_{10}\left(ROUNDDOWN\left(10^{Max\{C_{pack} + CC_{retail}, 0\}}, 0\right)\right)$	-	Calculated [a]
$t_{transport}$	Transportation time (h)	$Normal(1.421,0.46478, Truncate(0.1833,3.8667), Shift(-0.24609))$	1.18 (0.43; 1.94)	*[95]*
T_{trhom}	Package temperature before putting into home refrigerator (°C)	$Normal\left(8.386,3.831, Truncate(0,20)\right)$	8.51 (2.54; 14.67)	*[95]*
T_{trnspt}	Temperature during transportation (°C)	$\frac{1}{2} \times (T_{retail} + T_{home})$	-	*[96]*
G_{trnspt}	Transportation growth rate (Log_{10}CFU/)	$\frac{\left(0.023 \times \left(T_{trnspt} - 1.2\right)\right)^2}{2.303}$	-	*[38]*
D_{trnspt}	Transportation die-off rate (Log_{10}CFU/h)	$\frac{Lognorm\left(0.013,0.001, Shift(0.001)\right)}{2.303}$	0.006 (0.005; 0.007)	*[75]*
$GorD_{trnspt}$	Growth (G) or die off (D) based on transport temperature	$G, if\ T_{trnspt} > 5,$ $D, if\ T_{trnspt} \leq 5$	-	*[38]*
CC_{trspt}	Change in ECO157 count during transportation (Log_{10}CFU)	$CC_{trnspt}\left(GorD_{trnspt} = G\right) = G_{trnspt} \times t_{trnspt}$ $CC_{trnspt}\left(GorD_{trnspt} = D\right) = -D_{trnspt} \times t_{trnspt}$	-	*[38]*
C_{trnspt}	ECO157 count in a contaminated package after transportation (Log_{10}CFU/romaine package)	$\text{Log}_{10}(ROUNDDOWN\left(10^{Max\{C_{retail} + CC_{trnspt}, 0\}},\right.$	-	Calculated [a]
t_{fcons}	Time to first consumption (h)	$Weibull(1.13,2.84) \times 24$	65.21 (4.92; 179.97)	*[97]*
t_{lcons}	Time to last consumption (h)	$Weibull(1.7,7.96) \times 24$	170.45 (33.29; 364.25)	*[97]*
t_{home}	Home storage time (h)	$\frac{1}{2} \times \left(t_{fcons} + t_{lcons}\right)$	-	Calculated [a]

T_{home}	Home storage temperature (°C)	$Normal\left(\begin{matrix}3.4517,2.4442,\\ RiskTruncate(-5,17.22)\end{matrix}\right)$	$3.45\ (-0.56; 7.47)$	*[38]*
G_{home}	Home growth rate (Log_{10}CFU/h)	$\frac{\left(0.023\ \times (T_{home}-\ 1.2)\right)^2}{2.303}$	-	*[38]*
D_{home}	Home die-off rate (Log_{10}CFU/h)	$\frac{Lognorm\left(0.013,0.001,Shift(0.001)\right)}{2.303}$	$0.006\ (0.005; 0.007)$	*[75]*
$GorD_{home}$	Growth (G) or die off (D) based on home storage temperature	$G, if\ T_{home} > 5,$ $D, if\ T_{home}\ \leq\ 5$	-	*[38]*
CC_{home}	Change in ECO157 count during home storage (Log_{10}CFU)	$CC_{home}(GorD_{home} = G) =\ G_{home} \times t_{home}$ $CC_{home}(GorD_{home} = D) =\ -D_{home} \times t_{home}$	-	*[38]*
C_{home}	ECO157 count in a contaminated package after home storage (Log_{10}CFU/romaine package)	$\mathrm{Log}_{10}(ROUNDDOWN\left(10^{Max\{C_{trnspt}+\ CC_{home},0\}},(\right.$	-	Calculated [a]
$P0_{ser}$	Probability a non-contaminated serving was partitioned from a contaminated romaine package	$EXP\left(\frac{\Gamma(Clstr_{par}\ \times\ Ser_{pack})\times\Gamma(Clstr_{par}\times(Ser_{pack}-1)+C_{home})}{\Gamma\left(Clstr_{par}\times(Ser_{pack}-1)\right)\times\Gamma(Clstr_{par}\times Ser_{pack}+C_{home})}\right)$	-	Calculated [a]
P_{ser}	Probability of ECO157 contamination in a serving	$(1-P0_{ser})\times P_{pack}\times P_{Lh_f}$	$0.002\ (1.8 \times 10^{-6}; 0.009)$	Calculated [a]
Sc_{clstrp}	0 if no clustering in package (default), 1 if clustering	0	-	Selected [a]
$Clstr_{parp}$	Clustering parameter	1	-	*[42]*
Ser	Serving size (g romaine/ romaine serving)	85	-	*[98]*
Ser_{pack}	Number of servings in a package (romaine servings)	$\frac{W_{lettucepack}}{Ser}$	-	Calculated [a]
$P_{contser_{prep}}$	Distribute ECO157 cells to a serving when clustering (unitless)	$Beta\left(Clstr_{parp}, Clstr_{parp}\times\left(Ser_{pack}\ -\ 1\right)\right)$	$0.38\ (0.031; 0.838)$	Calculated [a]
$P_{contser}$	Proportion of ECO157 cells from contaminated package in a serving (unitless)	$\left(1\ -\ Sc_{clstrp}\right)\times\left(\frac{1}{Ser_{pack}}\right)$ $+\ Sc_{clstrp}$ $\times P_{contser_{prep}}$	-	*[42]*
C_{serH}	ECO157 count in a contaminated serving after home storage (Log_{10}CFU/serving romaine)	$Log_{10}(Binomial(10^{C_{home}}, P_{contser})+\ 1)$	-	Calculated [a]
Sc_{conw}	0 if consumer does not wash romaine before consumption (default), 1 if consumer does	0	-	Selected [a]
P_{conw}	Probability a USA consumer washes the prepacked romaine (unitless)	0.62	-	*[99]*
$CC_{conw_{prep}}$	ECO157 reduction by consumer wash with water prep cell (Log_{10}CFU)	$Triang(0.65,0.99,0.99)\times P_{conw}$	$0.54\ (0.45; 0.61)$	*[39]*
CC_{conw}	Log_{10}CFU ECO157 reduction by consumer wash with water (Log_{10}CFU)	$Sc_{conw}\times CC_{conw_{prep}}$	-	Calculated [a]
C_{serW}	ECO157 count in a contaminated serving after consumer wash (CFU/romaine serving)	$10^{Max\{C_{serH}-CC_{conw},0\}}$	-	Calculated [a]
$alpha$	Dose-response parameter alpha (unitless)	0.267	-	*[79]*
$beta$	Dose-response parameter beta (unitless)	229.2928	-	*[79]*

$Pill$	Probability of illness as a result of consumption of a contaminated romaine serving (unitless)	$1 - \left(1 + \frac{C_{serW}}{beta}\right)^{-alpha}$	-	*[32]*
$USpop$	USA population (person)	311,556,874	-	*[80]*
A	Annual romaine consumption in the USA (g/consumer/year)	3,946	-	*[81]*
Np	Number of romaine servings consumed per person per year (romaine serving/consumer/year)	$\frac{A}{Ser}$	-	Calculated [a]
Nsc_{US}	Total number of romaine servings consumed per year in the USA (romaine serving/year)	$USpop \times Np$	-	Calculated [a]
$Ncsc_{US}$	Number of contaminated romaine servings consumed per year in the USA (romaine serving/year)	$Nsc_{US} \times P_{ser}$	-	Calculated [a]
$Ncase_{US}$	Number of illness cases per year in the USA (person/year)	$Ncsc_{US} \times Pill$	-	Calculated [a]

[a] Assumption: Parameter value accepted as true with little to no information; Author opinion: Parameter value accepted as true based on authors' experience; Calculated: Parameter value calculated based on other parameters; Selected: Used in scenario analysis (See Supplementary Table S6 for details for their use.)

Table 3. Distributions and values varying by the calendar month of romaine production

Parameter	P_{sun_m}	P_{cow_m}	P_{pig_m}	h_{sun_m}	R_{sun_m}	$Prec_{norm_m}$	$Prec_{actual_m}$	$T_{harvest_m}$
Description (unit)	Probability of a sunny day	ECO157 prevalence in fresh cattle feces	ECO157 prevalence in fresh feral swine feces	Average sunny hours per day (h)[e]	Solar decay rate (1/day)	Precipitation level normalized	Actual (average) level of precipitation during the year (mm)	Temperature at harvest (°C) [d]
Source	*[100]*	*[8]*	*[8]*	*[27,34]*	*[27,34]*	*[100]*	*[100]*	*[100]*
Harvest month [a]	**Value/Distribution**							
January	0.97	$Binomial(1,0.01)$	0 [c]	$DR = 10.30$ $Y = 9.00$	-0.48	0.21	11	$DR = 7.20$ $Y = 8.90$
February	0.93	$Binomial(1,0.02)$	0 [c]	$DR = 11.00$ $Y = 9.70$	-0.48	0.31	15	$DR = 7.20$ $Y = 9.90$
March	0.97	$Binomial(1,0.06)$ [c]	0 [c]	$DR = 12.00$ $Y = 10.70$	-0.48	0.12	7	$DR = 12.70$ $Y = 12.90$
April	0.87	$Binomial(1,0.09)$	0 [c]	$CC = 8.90$	-0.48	0.81	36	$CC = 8.20$
May	0.90	$Binomial(1,0.04)$	0 [c]	$CC = 9.10$	-0.48	0.41	19	$CC = 9.70$
June	0.97	$Binomial(1,0.13)^b$	0 [c]	$CC = 9.80$	-0.52	0.07	5	$CC = 11.40$
July	1.00	$Binomial(1,0.21)$	$Binomial(1,0.16)$	$CC = 9.60$	-0.52	0.00	2	$CC = 13.00$
August	1.00	$Binomial(1,0.10)$	$Binomial(1,0.13)$	$CC = 9.10$	-0.52	0.02	3	$CC = 13.30$
September	0.97	$Binomial(1,0.08)$	$Binomial(1,0.11)$	$CC = 9.10$	-0.52	0.02	3	$CC = 13.10$
October	0.93	$Binomial(1,0.17)$	$Binomial(1,0.08)$	$CC = 8.50$	-0.52	0.41	19	$CC = 11.20$

November	0.87	$Binomial(1,0.03)$	$Binomial(1,0.10)$	$CC = 7.60$	-0.52	1.00	44	$CC = 8.00$
December	0.97	$Binomial(1,0.02)$	$Binomial(1,0.12)$	$DR = 10.00$ $Y = 8.80$	-0.48	0.21	11	$DR = 7.20$ $Y = 8.30$

[a] Selected month is based on the location of the romaine production.

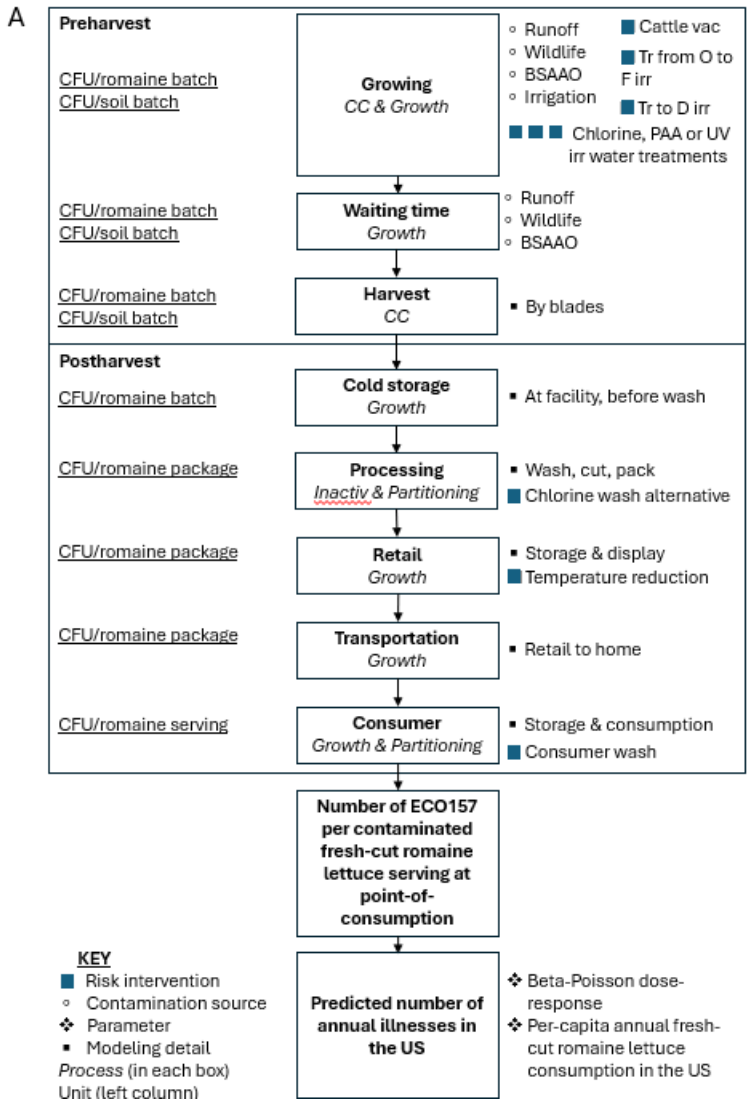
[b] Empirical data was 0. However, the value considered here was the average of estimates from the prior and the following months.

[c] Empirical data was 0.

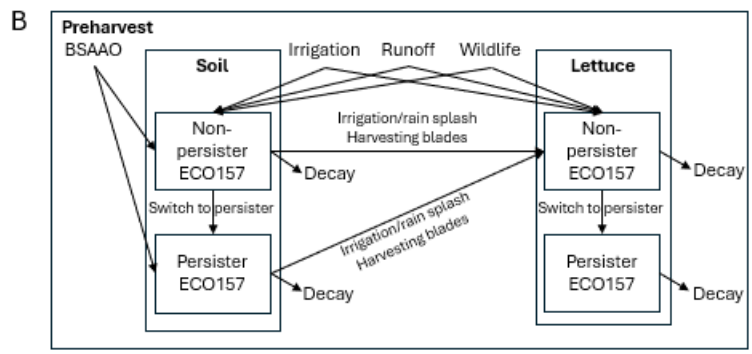
[d] CC: Coastal Region and Central Valley, CA; DR: Desert Region (Imperial), CA; Y:Yuma Region, AZ.

Results

Preharvest model: Under the baseline model conditions, our model predicted a 1.00% prevalence of ECO157 in romaine batches across all three regions. The median ECO157 count in contaminated romaine batches (637.5kg romaine/batch or 2,125 plants/batch) after harvest (C_{Lh_f}) was 3.30 Log_{10}CFU/romaine batch (5th – 95th percentile: 0.59 – 5.41).



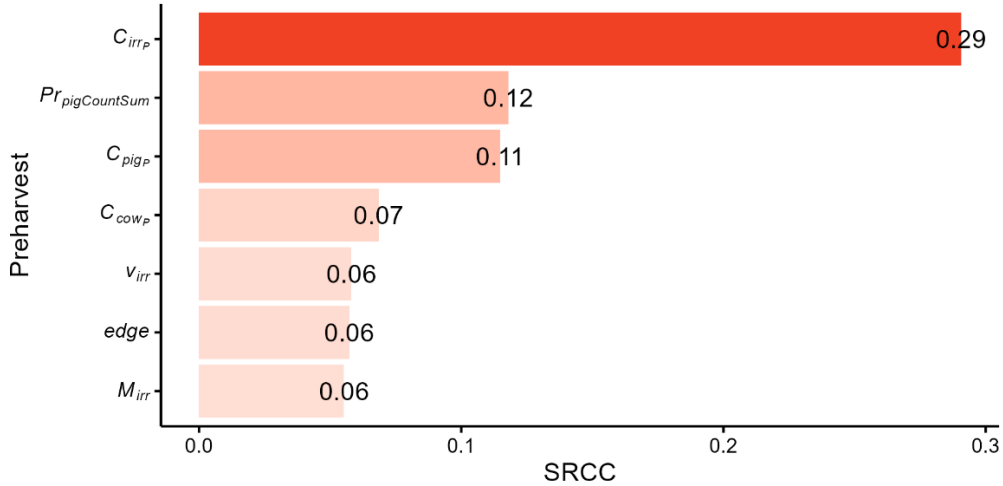
BSAAO: Biological soil amendment of animal origin**; Cattle vac**: Cattle vaccination; **Tr from O to F irr**: Transition from overhead to furrow irrigation; **Tr to D irr**: Transition from overhead and furrow to drip irrigation; **Chlorine, PAA or UV irr water treatments**: Surface irrigation treatments either with chlorine, peracetic acid or ultraviolet radiation; ***CC***: Cross-contamination; ***Inactiv***: Inactivation; **ECO157**: *Escherichia coli* O157:H7.



BSAAO: Biological soil amendment of animal origin; **ECO157**: *Escherichia coli* O157:H7.

Figure 1. (A) Conceptual framework of the quantitative microbial risk assessment (QMRA) model for predicting illnesses due to *Escherichia coli* O157 (ECO157) contamination in the farm-to-fork fresh-cut romaine supply chain. **(B)** Conceptual framework of the difference equation model for population dynamics on romaine and in soil during the 14-day preharvest countdown period.

The sensitivity analysis of the variable parameters in the preharvest model (Figure 2), determined by SRCC, indicated that the ECO157 count in irrigation water (C_{irr_P}) has a fair-strength (0.29) positive correlation with the ECO157 count in romaine batches (outcome of preharvest). The rest of the parameters were either poorly correlated or showed no correlation.



ECO157: *Escherichia coli* O157:H7; **SRCC**: The Spearman rank-order correlation coefficients; $\boldsymbol{C_{irr_P}}$: ECO157 count per ml of contaminated irrigation water applied on day t; $\boldsymbol{Pr_{pigCountSum}}$: Total number of days during the 14-day countdown when wildlife intrusion occurred; $\boldsymbol{C_{pig_P}}$: ECO157 count on any swine feces, which returned 0 CFU/g for non-contaminated feces and C_{pig} for contaminated feces; $\boldsymbol{C_{cow_P}}$: ECO157 count on any cattle feces, which returned 0 CFU/g for non-contaminated feces and C_{cow} for contaminated feces; $\boldsymbol{v_{irr}}$: Volume of irrigation water captured by romaine batches per an irrigation event; $\boldsymbol{edge}$: Batch area was at the edge of the field and thus could be contaminated by runoff: yes=1, no=0; $\boldsymbol{M_{irr}}$: Mass of soil transferred to a romaine batch via splash per irrigation event.

Figure 2. Tornado plot showing the impact of parameter variability on the preharvest output: the predicted *E. coli* O157:H7 (ECO157) count on romaine batches after harvest, ready for processing. The Spearman rank-order correlation coefficients (SRCC) are shown next to each parameter bar. Only variables statistically significant after the Bonferroni correction and with a correlation coefficient > 0.05 are shown. All values are positive, indicating positive correlation (i.e., increase in each parameter value increases ECO157 count in romaine batches).

To assess the impact of preharvest contamination sources (irrigation, wildlife, runoff, and BSAAO) on the ECO157 counts in romaine batches, we ran the baseline preharvest model separately for each source and visualized their effects using violin plots (i.e., we evaluated the effect of each source individually, excluding the effect of the other three sources). Figure 3. (A) shows the distribution of ECO157 counts in contaminated romaine batches (i.e., with ECO157 counts ≥1 CFU/romaine batch), for each preharvest source, including the median, interquartile range, and 5th to 95th percentile across 100,000 iterations. Irrigation was identified as the primary contributor to ECO157 counts in romaine batches (Figure 3 (B)) with a median ECO157 count of 3.91 Log_{10}CFU/romaine batch (5th – 95th percentile: 0.00 – 5.29) and an ECO157 prevalence of 0.53%. This was followed by wildlife, runoff and BSAAO contamination sources (Figure 3 (B)). While wildlife (median ECO157 count: 2.27 Log_{10}CFU/romaine batch, 5th – 95th percentile: 0.47 – 4.47) and runoff (median ECO157 count: 2.28 Log_{10}CFU/romaine batch, 5th – 95th percentile: 0.48 – 5.28) both led to similar ECO157 counts when contamination occurred, contamination by wildlife intrusion events was more frequent (0.26%) than runoff contamination (0.21%). The predicted ECO157 count in romaine batches contributed by BSAAO contamination was low but occurred in 2 out of 100,000 batches (0.002% occurrence), with ECO157 counts of 0.90 and 1.23 Log_{10}CFU/romaine batch. This result on BSAAO contamination was not negligible given the volume of romaine production in the USA: Between 2018 and 2022, an average of 3,578 million kilograms of lettuce were produced annually, with romaine and leaf lettuce together accounting for 53% of the total production *[102]*.

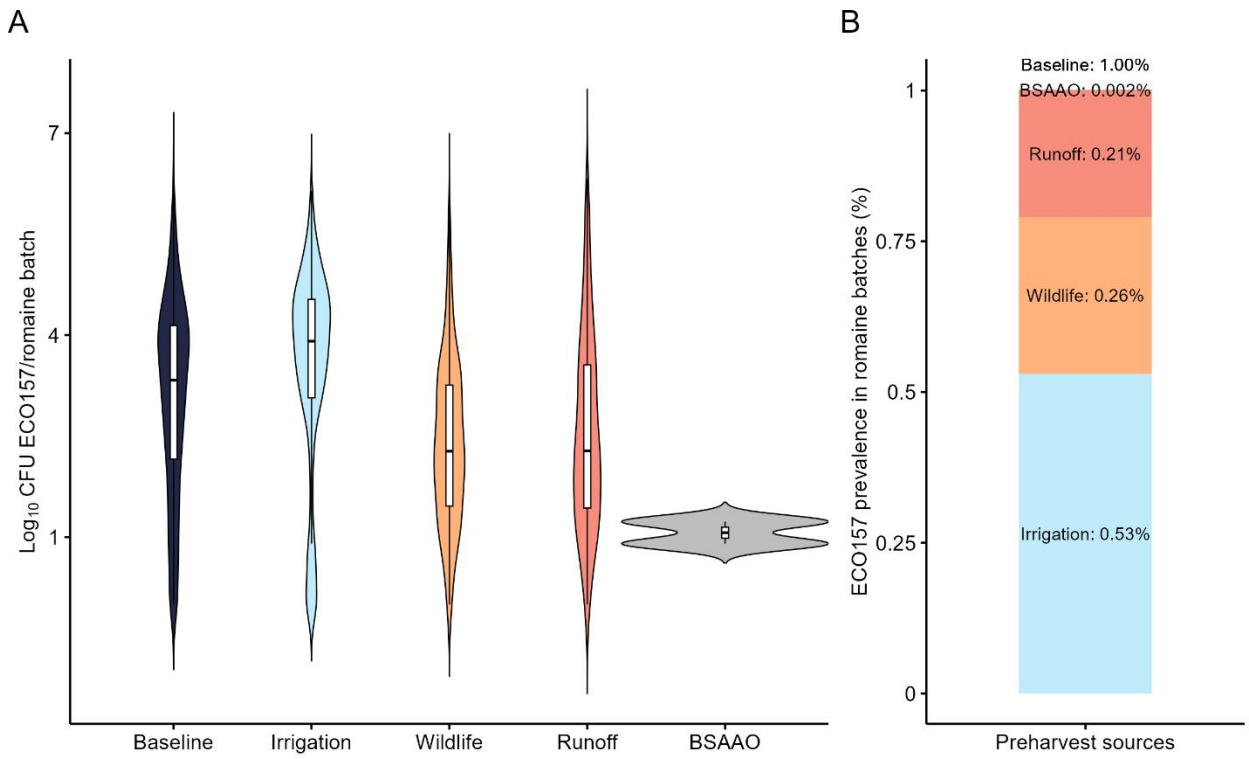


Figure 3. *Escherichia coli* O157:H7 (ECO157) counts in contaminated romaine batches in the baseline model (with all four sources of preharvest contamination accounted for), as well as in models considering each individual source of preharvest contamination separately **(A).** The ECO157 prevalence in romaine batches contributed by each individual preharvest contamination source included in the model, one at a time **(B)**.

We also evaluated the impact of three irrigation systems—overhead spray, furrow, and drip—on EC counts in romaine batches (Figure 4). We found that overhead spray irrigation was the primary contributor to ECO157 counts in romaine batches (median ECO157 count: 4.00 Log_{10}CFU/romaine batch, 5th – 95th percentile: 2.53 – 5.30), leading to ECO157 counts in romaine batches most frequently (0.47%). In contrast, we predicted that furrow and drip irrigation rarely caused ECO157 counts in romaine batches (0.04% and 0.02%, respectively) and resulted in substantially lower ECO157 counts when contamination occurred (Furrow median ECO157 count: 0.00 Log_{10}CFU/romaine batch, 5th – 95th percentile: 0.00 – 1.40; drip median ECO157 count: 0.30 Log_{10}CFU/romaine batch, 5th – 95th percentile: 0.00 – 1.54).

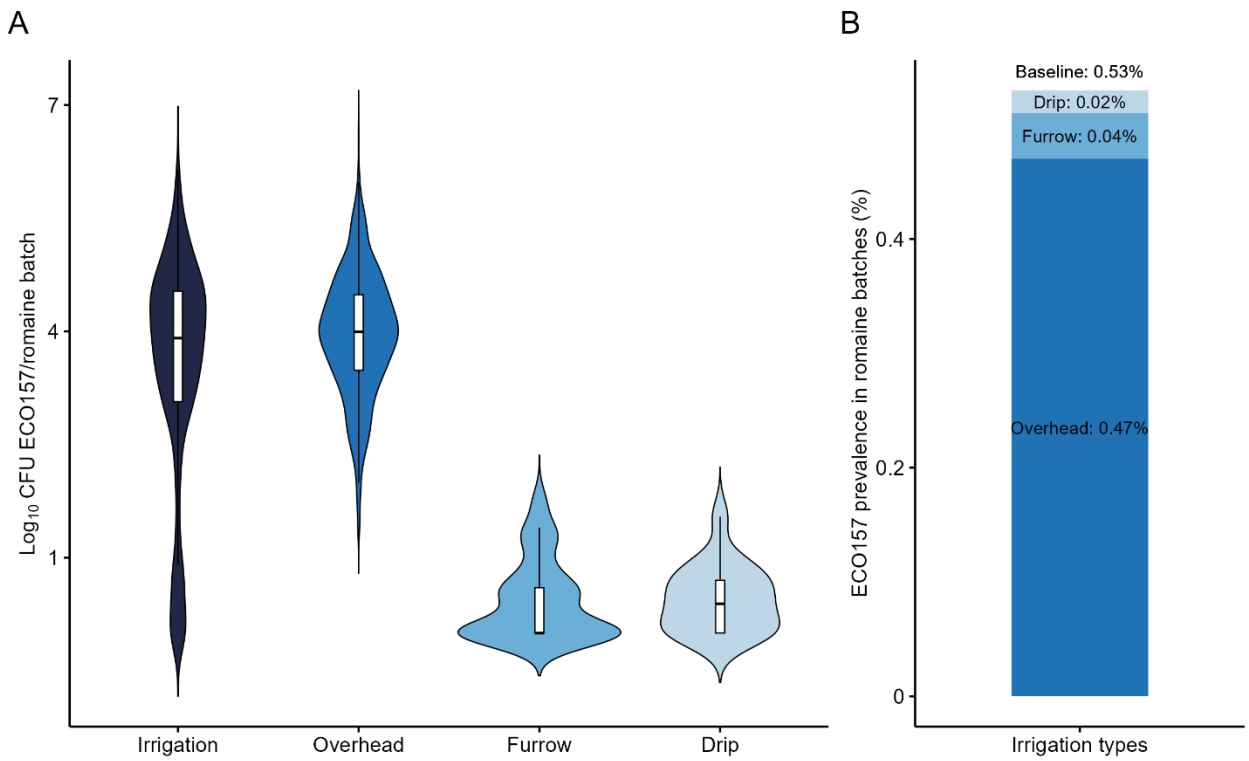
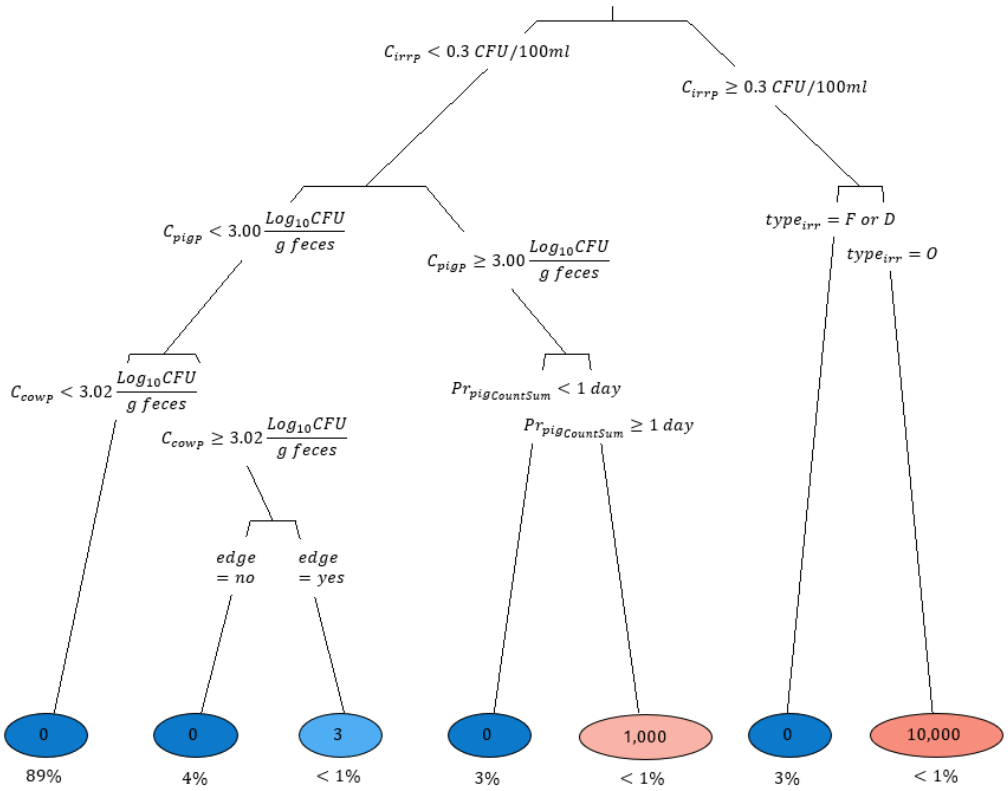


Figure 4. *E. coli* O157:H7 (ECO157) counts in contaminated romaine batches in the baseline irrigation model (considering only the three irrigation systems with no other preharvest sources), as well as in irrigation models considering each individual irrigation source contamination separately **(A).** The ECO157 prevalence in romaine batches contributed by each individual irrigation contamination source included in the irrigation model, one at a time **(B)**.

We ran a sensitivity analysis (using SRCC) for the variable parameters in the baseline irrigation model (considering only the three irrigation systems with no other preharvest sources), as well as for each individual irrigation source contamination separately — i.e., (i) baseline irrigation model with contamination from only overhead spray irrigation, (ii) baseline irrigation model with contamination from only furrow irrigation, and (iii) baseline irrigation model with contamination from only drip irrigation. Similar to what was observed in the baseline model, we found a poor to fair correlation between the ECO157 counts in romaine batches and the ECO157 counts in irrigation water (C_{irr_P}). Other variables were either poorly correlated or had no correlation with ECO157 counts in romaine batches, regardless of the irrigation system (Supplementary Table S7).

The preharvest regression tree captured the importance of irrigation and wildlife-related parameters in predicting ECO157 count in a romaine batch (Figure 5). The highest ECO157

counts in romaine batches were predicted when the ECO157 count in irrigation water was ≥ 0.3 CFU/100ml and applied via an overhead system. On the other hand, irrigation from furrow or drip systems led to substantially lower ECO157 counts in romaine batches. Also demonstrated was the importance of wildlife intrusions. When wildlife intrusion occurred for one or more days during the countdown period, it resulted in a mean of 1,000 CFU ECO157 per romaine batch, given that the fecal material from the feral swine contained 3 $\log_{10}$CFU ECO157 or more per gram of feces.



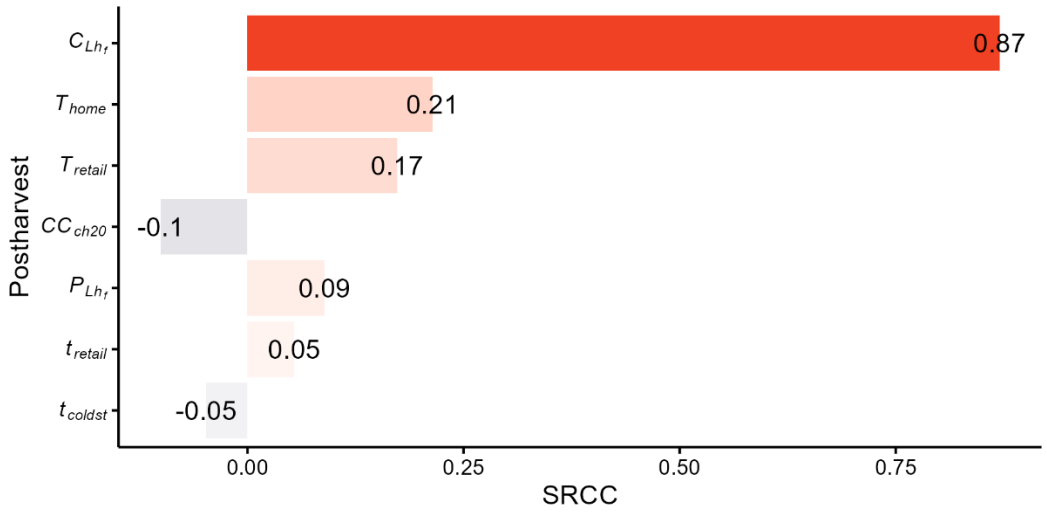
$\boldsymbol{C_{irrp}}$: ECO157 count per 100 ml of contaminated irrigation water applied on day t; $\boldsymbol{C_{pigp}}$:ECO157 count on any swine feces, which returns 0 CFU for non-contaminated feces and C_{pig} for contaminated feces; $\boldsymbol{C_{cowp}}$: ECO157 count on any cattle feces, which returns 0 CFU/g for non-contaminated feces and C_{cow} for contaminated feces; $\boldsymbol{type_{irr}}$: Probability the selected irrigation type is occurring on a given day; **O**: Overhead irrigation; **F**: Furrow irrigation; **D**: Drip irrigation; $\boldsymbol{Pr_{pigCountSum}}$: Total number of days during the 14-day countdown when wildlife intrusion occurred; $\boldsymbol{edge}$: Batch area was at the edge of the field and thus could be contaminated by runoff.

Figure 5. Preharvest model regression tree predicting the mean *Escherichia coli* O157:H7 (ECO157) counts in romaine batches under various preharvest conditions. Branch labels indicate the rules used for splitting each node, and the circled numbers at the end of each branch indicate the predicted mean ECO157 counts, expressed in CFU/romaine batch. The intensity of the circle colors was proportional to the value of predicted ECO157 counts for a given branch (blue: low; red: high). Underneath each circle (e.g., 0), the percentage of model iterations falling in that branch (e.g., 89%) were displayed (e.g., 89% was shown under 0, indicating that a mean

ECO157 count of 0 CFU/romaine batch was observed for this branch, which appeared in 89% of 100,000 iterations).

Postharvest model: The postharvest model predicted that the median number of cases resulting from the consumption of romaine was 19,040 annually (5th – 95th percentile: 42 – 3.6×10^6). Although our predicted median number of illness cases aligned with the reported range for the annual median illness cases (12,189 – 19,893), the predicted distribution of illness cases was wider than the reported validation data (reported 90% CI: 3,394 – 47,134 cases annually). The validation data fell between 32nd and 64th percentiles of the predicted illness cases. Considering the number of romaine servings consumed per year in the USA, the annual median risk of ECO157 due to consumption of ECO157-contaminated fresh-cut romaine was predicted as 0.000132% (or 1 illness in approximately 759,641 consumed servings). In addition, the predicted median prevalence of ECO157 contaminated romaine packages was 0.14% following the retail storage, but in 51% of the iterations the contaminated package (225g) had only 1 CFU per package.

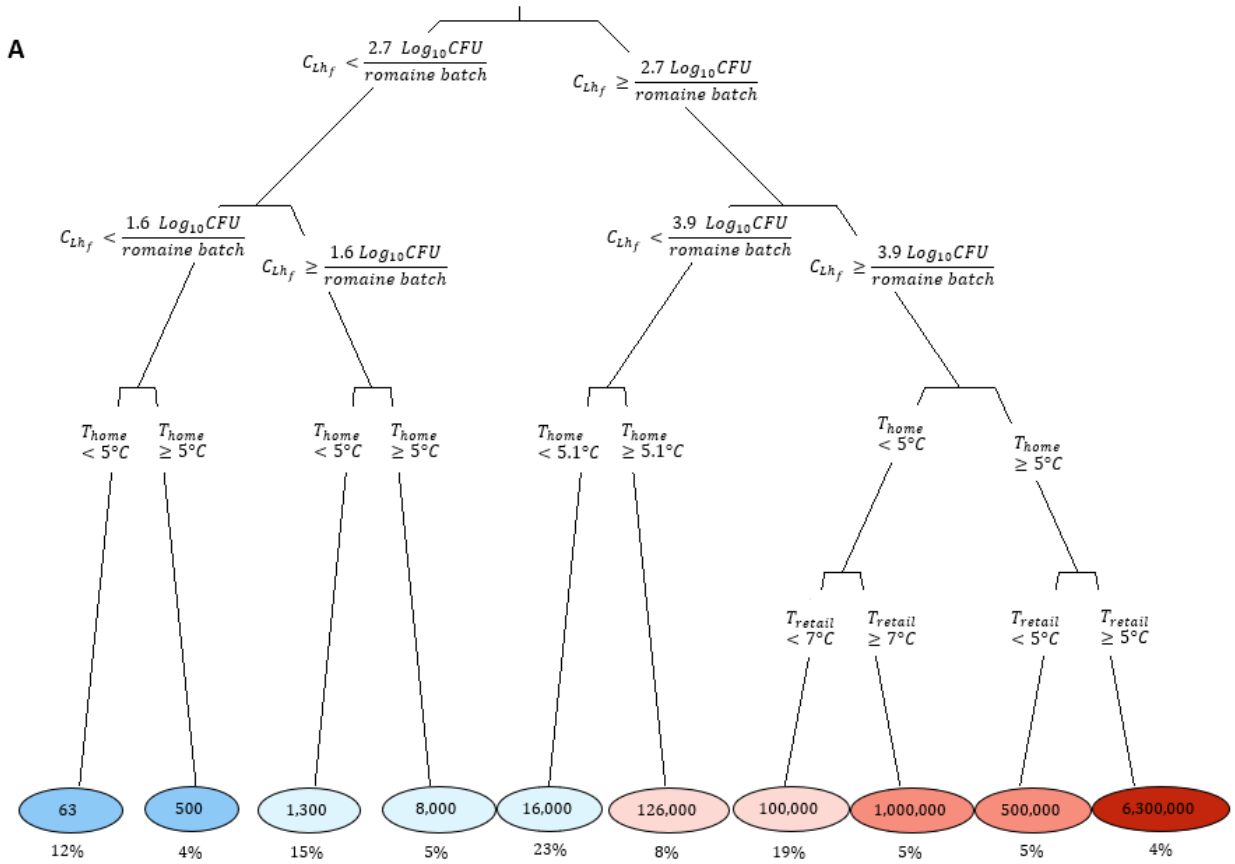
The sensitivity analysis of variable parameters in postharvest model (Figure 6), determined by SRCC, showed that the median number of illness cases was very strongly correlated with the ECO157 count in romaine batches entering processing (C_{Lh_f}). The remaining parameters were either poorly correlated or showed no correlation with the number of illness cases.

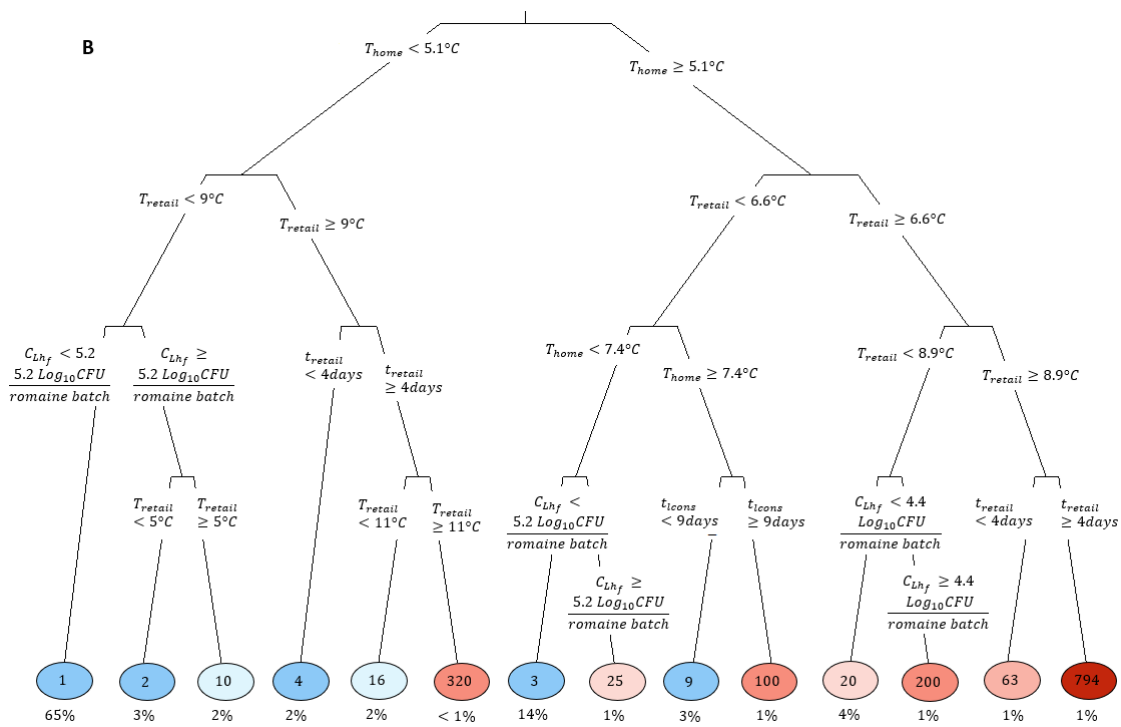


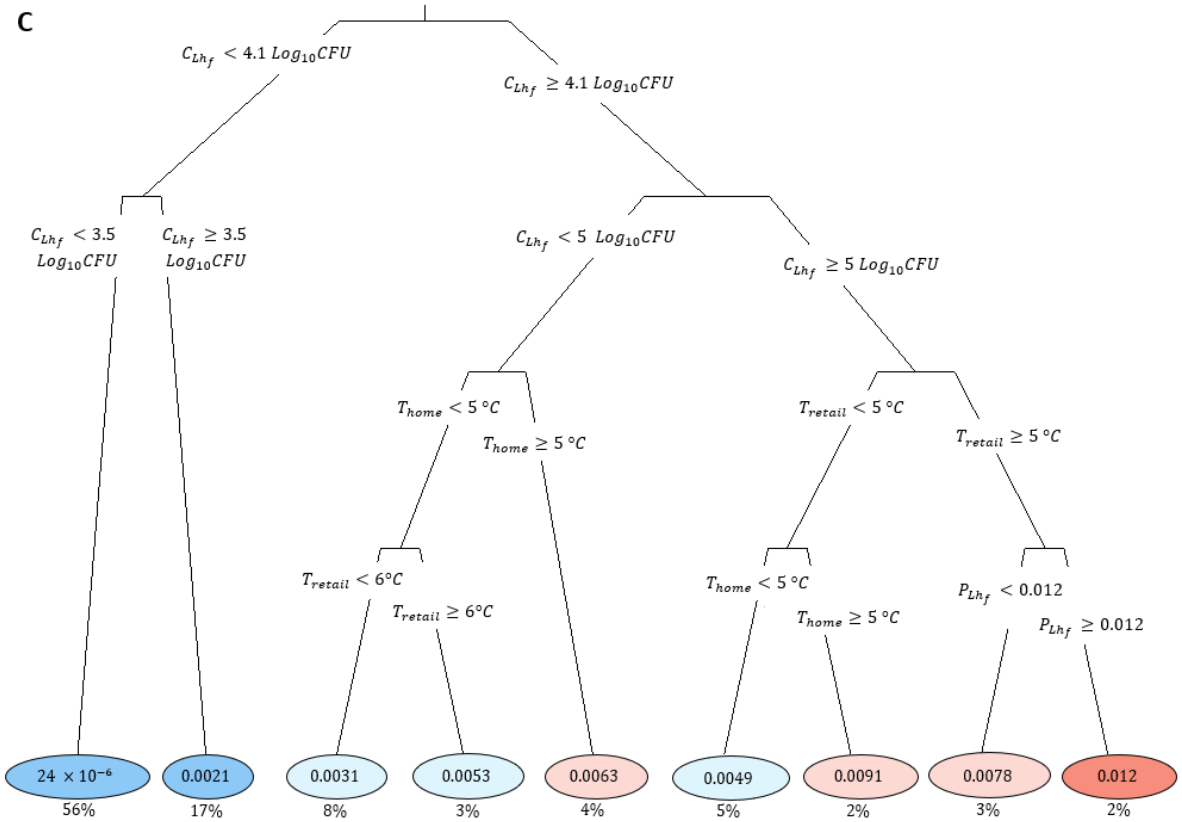
$\boldsymbol{C_{Lh_f}}$: ECO157 count on contaminated romaine batch after harvest; $\boldsymbol{T_{home}}$: Home storage temperature; $\boldsymbol{T_{retail}}$: Retail storage temperature; $\boldsymbol{CC_{ch20}}$: Log_{10} reduction with chlorine wash; $\boldsymbol{P_{Lh_f}}$: Probability of ECO157 contamination in romaine batch after harvest; $\boldsymbol{t_{retail}}$: Retail storage time; $\boldsymbol{t_{coldst}}$: Cold storage time.

Figure 6. Tornado plot showing the impact of variable parameters on the postharvest output, i.e., the median predicted annual number of *E. coli* O157:H7 (ECO157) illness cases. The spearman rank-order correlation coefficients (SRCC) are shown next to each parameter bar. Only variables statistically significant after the Bonferroni correction and with a correlation coefficient > 0.05 are shown. Positive values indicate positive correlation (i.e., an increase in a given parameter increases the number of illness cases), while negative values indicate negative correlation (i.e., an increase in a given parameter decreases the number of illness cases).

The postharvest regression trees demonstrated how ECO157 counts, and consequently the number of illness cases, substantially increases due to both preharvest and postharvest conditions (Figure 7). Specifically, the ECO157 prevalence and counts in preharvest romaine batches, as well as the temperatures and durations of postharvest retail and home storage, play a critical role in this increase. In Figure 7. (A), the regression tree predicted a high number of illness cases, although occurring with a small probability, when $C_{Lh_f} \geq 3.9$ Log_{10}CFU/romaine batch, $T_{retail} \geq 5°C$ and $T_{home} \geq 5°C$. Illness cases were lowest when $C_{Lh_f} < 1.6$ Log_{10}CFU/romaine batch and $T_{home} < 5$ °C. Figure 7. (B) indicated that the ECO157 counts in contaminated servings were increased by a combination of time-temperature abuses during retail and home storage; ≥ 4 day storage at $\geq 8.9°C$ in retail, combined with home storage at $\geq 5.1°C$ resulted in the highest ECO157 counts in romaine servings. On the other hand, the ECO157 counts were lowest when $C_{Lh_f} < 5.2$ Log_{10}CFU/romaine batch, $T_{retail} < 9°C$ and $T_{home} < 5.1°C$. Figure 7. (C) indicated that the highest prevalence of ECO157 contaminated romaine servings occurred when the preharvest ECO157 counts and prevalence in romaine batches was higher, along with retail storage at $\geq 5°C$.


A
$C_{Lh_f} < \frac{2.7\ Log_{10}CFU}{romaine\ batch}$
$C_{Lh_f} \geq \frac{2.7\ Log_{10}CFU}{romaine\ batch}$
$C_{Lh_f} < \frac{1.6\ Log_{10}CFU}{romaine\ batch}$
$C_{Lh_f} \geq \frac{1.6\ Log_{10}CFU}{romaine\ batch}$
$C_{Lh_f} < \frac{3.9\ Log_{10}CFU}{romaine\ batch}$
$C_{Lh_f} \geq \frac{3.9\ Log_{10}CFU}{romaine\ batch}$
$T_{home} < 5°C$
$T_{home} \geq 5°C$
$T_{home} < 5°C$
$T_{home} \geq 5°C$
$T_{home} < 5.1°C$
$T_{home} \geq 5.1°C$
$T_{home} < 5°C$
$T_{home} \geq 5°C$
$T_{retail} < 7°C$
$T_{retail} \geq 7°C$
$T_{retail} < 5°C$
$T_{retail} \geq 5°C$
63
12%
500
4%
1,300
15%
8,000
5%
16,000
23%
126,000
8%
100,000
19%
1,000,000
5%
500,000
5%
6,300,000
4%


B
$T_{home} < 5.1°C$
$T_{home} \geq 5.1°C$
$T_{retail} < 9°C$
$T_{retail} \geq 9°C$
$T_{retail} < 6.6°C$
$T_{retail} \geq 6.6°C$
$C_{Lh_f} < \frac{5.2\ Log_{10}CFU}{romaine\ batch}$
$C_{Lh_f} \geq \frac{5.2\ Log_{10}CFU}{romaine\ batch}$
$t_{retail} < 4days$
$t_{retail} \geq 4days$
$T_{home} < 7.4°C$
$T_{home} \geq 7.4°C$
$T_{retail} < 8.9°C$
$T_{retail} \geq 8.9°C$
$T_{retail} < 5°C$
$T_{retail} \geq 5°C$
$T_{retail} < 11°C$
$T_{retail} \geq 11°C$
$C_{Lh_f} < \frac{5.2\ Log_{10}CFU}{romaine\ batch}$
$C_{Lh_f} \geq \frac{5.2\ Log_{10}CFU}{romaine\ batch}$
$t_{lcons} < 9days$
$t_{lcons} \geq 9days$
$C_{Lh_f} < \frac{4.4\ Log_{10}CFU}{romaine\ batch}$
$C_{Lh_f} \geq \frac{4.4\ Log_{10}CFU}{romaine\ batch}$
$t_{retail} < 4days$
$t_{retail} \geq 4days$
1
65%
2
3%
10
2%
4
2%
16
2%
320
< 1%
3
14%
25
1%
9
3%
100
1%
20
4%
200
1%
63
1%
794
1%



$\boldsymbol{C_{Lh_f}}$: ECO157 count on a contaminated romaine batch after harvest; $\boldsymbol{T_{home}}$: Home storage temperature; $\boldsymbol{T_{retail}}$: Retail storage temperature; $\boldsymbol{t_{lcons}}$: Time to last consumption; $\boldsymbol{P_{Lh_f}}$: ECO157 prevalence in a romaine batch after harvest; $\boldsymbol{t_{retail}}$: Retail storage time.

Figure 7. Regression trees for postharvest model, predicting illness outcome: **(A)** the mean number of illness cases from *Escherichia coli* O157:H7 (ECO157), **(B)** the ECO157 counts in contaminated servings in CFU, and **(C)** the prevalence of ECO157 contaminated romaine servings (unitless). Branch labels indicated the rules used for splitting each node, and the circled numbers at the end of each branch indicated the predicted mean value for the outcomes. The intensity of the circle colors was proportional to the value of predictions at the branch (blue: low; red: high). Underneath each circle (e.g., 63 cases), the percentage of model iterations falling in that branch (e.g., 12%) were displayed (e.g., 12% was shown under 63, indicating that a mean illness case number of 63 was observed for this branch, which appeared in 12% of 100,000 iterations).

In the preharvest model, none of the uncertainty variables were correlated with the predicted ECO157 count in romaine batches. Additionally, only approximately 0.3% of illness cases were attributed to the presence of persister cells in the baseline model, as detailed in Supplementary Table S8.

Scenario analysis

In the 9 scenarios we tested (excluding clustering effect), the median number of illness cases decreased by 17.1% to 99.5% (Table 4). Out of the nine scenarios tested independently, six targeted reducing ECO157 counts at the preharvest stage. However, the most effective method was the postharvest chlorine wash alternative (with a median of 3.22 Log_{10}CFU reduction instead of the baseline chlorine wash with 0.52 Log_{10}CFU median reduction), which achieved a 99.5% reduction in median illness cases. Surface water treatments and transitioning to furrow or drip irrigation resulted in slightly lower but highly effective reductions in ECO157 counts at the preharvest stage, achieving between 90.5% and 96.8% reductions in median illness cases. In contrast, consumer washing (28.0%), cattle vaccination (25.9%) and temperature reduction at retail (17.1%) were considerably less effective at reducing the number of illness cases.

The scenario analyses also showed that the degree of clustering assumed during the partitioning process affects the risk of illness. Increased clustering (in a batch and package) increased the dose of exposure and thus the probability of illness given exposure, but it also reduced the prevalence of ECO157 positive servings, which led to reduced number of cases (Table 4).

Table 4. Description of the scenarios and changes in the illness outcome for potential risk management strategies

Scenario	Definition (Reference)	ID of parameters affected (unit)	Parameter value in the scenario	Predicted illness cases resulted from the tested scenario: 5th; 50th; 95th percentile
Baseline	No changes made in the model.	None	All parameters were at their baseline conditions.	42; 19,040; 3.6×10^6
Cattle vaccination	ECO157 counts were reduced by 1-Log_{10}CFU/g in BSAAO and cattle feces *[87]*.	C_{man} (Log_{10}CFU/g BSAAO) and C_{cow} (Log_{10}CFU/g feces)	C_{man}= $Uniform(3.1,8.4) - 1$ C_{cow}= $Pert(0.89,3.08,8.4) - 1$	43; 14,113; 2.6×10^6 Median number of cases reduced by 25.9%.
Surface water treatment for overhead spray irrigation by chlorine	ECO157 counts in surface water were reduced when chlorine was applied *[90]*.	C_{irr_S} (CFU/ml)	$\frac{C_{irr_S}}{10^{\text{Uniform}(1.00,2.78)}}$	27; 1,803; 4.7×10^5 Median number of cases reduced by 90.5%.
Surface water treatment for overhead spray irrigation by peracetic acid (PAA)	ECO157 counts in surface water were reduced when PAA was applied *[90,91]*.	C_{irr_S} (CFU/ml)	$\frac{C_{irr_S}}{10^{\text{Uniform}(2.48,3.09)}}$	20; 608; 2.7×10^5 Median number of cases reduced by 96.8%.
Surface water treatment for overhead spray irrigation by Ultraviolet radiation (UV)	ECO157 counts in surface water were reduced when UV was applied *[91]*.	C_{irr_S} (CFU/ml)	$\frac{C_{irr_S}}{10^{\text{Uniform}(1.30,2.91)}}$	27; 1,705; 4.4×10^5 Median number of cases reduced by 91.1%.
Transition from overhead to furrow irrigation (i.e., all	All overhead irrigated fields were transitioned to furrow irrigated fields.	Prp_O and Prp_F (unitless)	Prp_O= $CC: 0.0;\ DR: 0.0;\ Y: 0.0$ Prp_F= $CC: 0.4;\ DR: 0.7;\ Y: 1.0$	12; 715; 6.3×10^5 Median number of cases reduced by 96.3%.

irrigation was through either furrow or drip)				
Transition from overhead and furrow to drip irrigation (all irrigation was through drip)	All overhead and furrow irrigated fields were transitioned to drip irrigated fields.	Prp_O, Prp_F and $Prpr_D$ (unitless)	Prp_O and Prp_F= CC: 0.0; DR: 0.0; Y: 0.0 Prp_D= CC: 1.0; DR: 1.0; Y: 1.0	14; 706; 5.3×10^5 Median number of cases reduced by 96.3%.
Chlorine wash alternative with a median of 3.22 Log_{10}CFU reduction instead of the baseline chlorine wash	Chlorine wash with higher reduction was used in postharvest instead of the baseline chlorine wash *[94]*.	CC_{ch20} (Log_{10}CFU)	CC_{ch20}= $Pert(0.01,0.52,1.03)$ was changed to CC_{ch30}= $Pert(3.04,3.22,3.4)$	19; 97; 1.4×10^4 Median number of cases reduced by 99.5%.
Temperature reduction during retail	Temperature distribution was changed *[101]*.	T_{retail} (°C)	T_{retail}= $Normal(4.4441,2.9642,$ $Truncate(0,20.56))$ was changed to $Uniform(0,4.4)$	38; 15,792; 5.5×10^5 Median number of cases reduced by 17.1%.
Consumer wash	62% of consumers washed the product (yielding a median of 0.99 Log_{10} CFU ECO157 reduction before consumption) *[99]*.	CC_{conw} (unitless)	CC_{conw}=0 was changed to $Triang(0.65,0.99,0.99)\times 0.62$	32; 13,717; 1.2×10^6 Median number of cases reduced by 28.0%.
Clustering effect	Clustering was allowed during partitioning of a romaine batch to packages and package to servings. Smaller $Clstr_{parb}$ or $Clstr_{parp}$ values indicated higher clustering of cells.	$Clstr_{parb}$ (unitless), $Clstr_{parp}$ (unitless)	$Clstr_{parb}$=0.1 $Clstr_{parp}$=0.1	38; 9,862; 2.8×10^5 Median number of cases reduced by 48.2%.
			$Clstr_{parb}$=1 $Clstr_{parp}$=1	44; 20,103; 2.9×10^6 Median number of cases increased by 5.6%.

Discussion

This risk assessment model simulated ECO157 counts during the production and consumption of fresh-cut romaine, providing a comprehensive view of how various preharvest and postharvest factors influence food safety risk. Our findings indicated that the model's predictions of ECO157 counts, and associated illness cases aligned with surveillance studies in the USA. Irrigation emerged as the most significant source of ECO157 counts in fresh-cut romaine, particularly when considering the impact of other contamination sources: inadequately treated BSAAO, runoff, and wildlife intrusion. However, the public health impact could be mitigated through effective interventions. Our predictions also indicated a high number of illness cases due to the combined effects of preharvest and postharvest factors. Additionally, we identified key knowledge gaps in assessing human ECO157 exposure through contaminated leafy greens.

The developed risk assessment tracks the dynamics of ECO157 cross-contamination between romaine and soil prior to harvest.

We developed a novel difference equation model, which was incorporated into our preharvest model, quantifying ECO157 counts in both soil and romaine in the field. Similar approaches have been used in other risk assessment models, such as those describing the transmission of *Listeria monocytogenes* contamination in a smoked salmon facility *[103]*, *Salmonella* in pig slaughterhouses *[104]*, and *Salmonella* Typhimurium among farm swine *[105]*. This approach enabled us to track daily dynamics of ECO157 counts on romaine during the days prior to harvest. It allowed us to track the specific impacts of irrigation, runoff, BSAAO and wildlife; as well as to perform calculations related to ECO157 growth, biphasic decay and cross-contamination. Each of these factors has distinct characteristics that would be impossible to accurately capture with a static equation.

Model predictions for preharvest and retail level ECO157 counts are consistent with previous reports.

Our findings generally aligned with existing literature and foodborne illness surveillance data for the USA. The predicted annual median illness cases were in agreement with the average observed foodborne illnesses linked to ECO157-contaminated romaine in the USA. However, the distribution of the predicted illness cases was wider than the reported illness range. One potential reason for this discrepancy was that that we did not account for romaine loss/waste at the consumer stage, assuming instead that all retail romaine would be consumed. This assumption is likely inaccurate, given that it has been estimated that 39.7% lettuce loss occurred at the consumer level in the USA in 2011 *[81]*. Another reason for the right skewness may be explained by the interactions between various parameters in the model, which will be discussed in the following sections. The predicted ECO157 prevalence at preharvest (1.00%) falls within the range of previously reported values for leafy greens from California's central coast region *[106]*, which ranged from <0.01% in 2007 to 2.5% in 2013. While a previous study by Zhang et al. (2018) reported a 0.02% average prevalence of ECO157 on romaine at the retail level in the USA (95% Confidence Interval by Clopper-Pearson method: <0.01 - 0.10, based on 5,548 samples of 25g each) *[11]*, our predicted median ECO157 prevalence of contaminated romaine packages was over 7 times higher (0.14%, predicted for 225g packages). One potential explanation for the discrepancy between the reported and predicted prevalence could be the limitation of microbial tests in detecting contamination when smaller sample sizes are used. For instance, in Zhang et al.'s (2018) study, the detection limit was 1 CFU in a 25g sample (BAM method for diarrheagenic *E. coli*), equivalent to 0.04 CFU/g. For a 225g sample—such as the packages in our model—the detection limit would be 1 CFU per 225g, or 0.0044 CFU/g. Under this scenario, testing the entire 225g package could detect ECO157 levels 9 times lower than a 25g sample. Assuming a 225g sample size, estimating ECO157 counts could yield a prevalence 9 times higher than that observed in retail samples, potentially leading to 0.18% (0.02% × 9) ECO157 prevalence. In our risk assessment model, we considered any package (and serving)

contaminated if it contained even a single ECO157 cell. Nevertheless, the previously observed ECO157 prevalence (0.02%) *[11]* falls within the predicted distribution of ECO157 prevalence simulated in our study (P_{pack}, Table 2).

Irrigation is predicted to be the most important source of ECO157 contamination of fresh-cut romaine with appropriate interventions reducing public health impact.

Our findings suggested that mitigation strategies at preharvest should prioritize reducing contamination of irrigation water, particularly in cases of overhead spray irrigation supplied via surface water sources. Aligned with our findings, contaminated irrigation water has been defined as a major risk factor in the contamination of leafy greens *[11]*. Notably, contaminated surface water irrigation has either been directly linked to or strongly associated with several multistate ECO157 outbreaks of leafy greens in the USA *[107,108,109]*, including spinach *[107]* and iceberg lettuce outbreaks in 2006 *[109]*, and a romaine outbreak in 2018 *[108]*. For the romaine outbreak in 2018, ECO157 introduction into the irrigation water through aerial and land-based spray applications of crop protection chemicals was proposed as a plausible route of introduction *[110)*. On the other hand, the spinach and iceberg lettuce outbreaks appeared to have been facilitated by cross-contamination, where contaminated surface water and wastewater flowed into groundwater or surface water used for irrigating the lettuce plants *[107,109]*.

Our model also suggests that treatment of surface irrigation water—an important intervention strategy already required by the California and Arizona LGMAs since the 2023 update—may achieve a substantial reduction of human illness cases. This outcome was predicted for various surface water treatments, including chlorination, PAA or UV. We consistently observed a strong effectiveness across all treatment methods in reducing illness cases, based on the assumption that the different methods achieved ECO157 reduction in surface irrigation water, ranging from a minimum of 1.00 Log_{10} CFU/ml (chlorination) to a maximum of 3.09 Log_{10} CFU/ml (PAA). Previous research has also reported that these methods reduce *E. coli* counts by up to 6 Log_{10}CFU/ml in agricultural water sources, with PAA typically having the strongest impact *[89,111-116]*. However, the effectiveness of these methods can be highly variable, even when implemented correctly *[115]*. Their site-specific applicability is influenced by a range of factors including water quality (such as total suspended solids, turbidity, pH, oxygen demand and temperature), applied dose, contact time, and contaminant characteristics *[117-119]*. In an open leafy green production field, these factors can be affected by rainfall, runoff or sediment activity upstream of the water sources *[116]*. On the other hand, both Arizona and California LGMAs prohibit the use of surface irrigation waters that fail to meet the performance requirements. To use such waters, treatment and testing are required to demonstrate treatment efficacy and compliance with microbial standards. Therefore, future models should

(i) better account for the variability in treatment effectiveness and (ii) assess the impact of reducing treatment variability: A less variable treatment with lower Log_{10} CFU/ml reduction may be more effective than one with a higher reduction but greater variability. Addressing these challenges can help leafy green growers, who must navigate multiple treatment options influenced by various factors while meeting LGMAs' requirements, make more informed decisions.

Our data also suggested that transitioning from overhead spray irrigation to furrow or, ideally, drip irrigation systems could provide a viable solution for reducing illness cases (Figure 4, Table 4). Despite introducing numerous parameters specific to each irrigation system that reasonably differentiate between the three irrigation systems, our model suggested that the primary driver behind the higher number of illness cases associated with overhead spray irrigation was its capacity for increased direct contact between irrigation water and romaine (Supplementary Table S7). Conversely, drip irrigation, delivering irrigation water to the plant roots, minimized water contact, and consequently resulted in the lowest number of illness cases among the three irrigation systems. Numerous field studies *[120-122]* and risk assessments *[40,41]* have consistently demonstrated that overhead spray irrigation system leads to higher counts of generic *E. coli* and ECO157 on romaine compared to furrow or drip irrigation systems, and argued that the type of irrigation system plays a critical role in mitigating the risk of ECO157 contamination on leafy greens *[111,123,124]*. Aligned with our model, research also supports that furrow irrigation leads to greater produce contamination than drip irrigation, highlighting drip irrigation as the most effective irrigation method for reducing health risks from contaminated irrigation water *[40,66,125]*. In our study, transitioning fully from overhead and furrow irrigation to drip irrigation was predicted to be more effective than chlorination and UV treatments of surface irrigation water, though less effective than PAA treatment.

The combined effect of preharvest and postharvest factors defines the public health risk. The developed model illustrated that interactions of preharvest and postharvest factors affect the number of predicted illness cases (Figure 7). This aligns with findings from observational studies, which have confirmed the multifactorial nature of food safety *[126]*. Our model highlighted how certain co-occurring factors could, though rarely, lead to a very large number of illnesses. Specific factors, such as higher EC counts in irrigation water, the use of overhead irrigation systems instead of furrow or drip irrigation, or increased number of wildlife intrusion, led to elevated ECO157 counts on preharvest romaine (Figure 5), which were further increased by improper handling and storage conditions during postharvest, including time-temperature abuse from harvest until consumption (Figure 7). This highlights the importance of simulation tools like the one developed here that can be used to study impacts and prevention of rare events which would be challenging to study empirically.

Several aspects of farm-to-fork continuum of fresh-cut romaine production and consumption were included in the model but did not emerge as significant risk factors. These included harvesting months, weather conditions (i.e., precipitation intensity, number of sunny days, decay related to sun exposure) and several others. Weather conditions, including temperature, seasonality and precipitation events, have previously shown to influence pathogen counts on leafy greens *[10,35,127]*. In addition, historically, more leafy greens outbreaks have occurred in summer and fall seasons than spring and winter seasons in the USA *[1]*. The likely reason why our model didn't capture the significance of these factors, including weather conditions, was that weather conditions alone was insufficient to exacerbate the risk of contamination unless cooccurring with other factors that introduce ECO157 contamination into the field. For example, heavy rain will introduce contamination via runoff from a neighboring dairy farm only if there was a nearby dairy farm and cattle on the farm shed ECO157. Similarly, heavy rain would contaminate romaine via splash but only if soil was already contaminated. In addition, it was possible that certain factors that may increase ECO157 contamination during certain months or under certain weather conditions were not or not fully included in our model (e.g., increased leaf damage or blistering induced by certain weather conditions) *[128]*.

Our study demonstrated that an effective process washing could reduce the illness cases by over 99%, even with high preharvest contamination levels (Table 4). While washing was important, its impact on illness risk was limited in the baseline model due to the use of a conservative ECO157 reduction - a maximum of 1.03 Log_{10}CFU/g reduction (Figure 6). In practice, chlorination can achieve higher ECO157 reductions > 3 Log_{10}CFU/g, by preventing cross-contamination of ECO157 in the romaine batch during washing *[17,129]*. The reduction is influenced by various factors, including the concentration and state (free or bound) of chlorine, pH of the water, the duration of contact time, and the characteristics of the organism *[77,130,131]*. Addressing these parameters has been the focus of multiple studies, which aim to close knowledge gaps on chlorine wash efficacy *[130,132-134]*. Our study supports this effort, as the model with postharvest washes predicted reduced human illness cases along the farm-to-fork continuum of fresh-cut romaine production and consumption.

Model limitations and key knowledge gaps

Our model here reflected a synthesis of the current knowledge on ECO157 contamination throughout the farm-to-fork supply chain of fresh-cut romaine lettuce in the USA. However, we acknowledged limitations related to the model structure and parameterization. The most consequential assumption in our model was about dilution factors in preharvest, which facilitated modeling the spread of ECO157 from BSAAO, runoff, irrigation and fecal material on the field. For ECO157 spread via BSAAO, we assumed a typical tiling depth of 15cm to 20cm. In

contrast, for irrigation, swine feces introduction caused by feral pigs, and runoff, we assumed a uniform spread of ECO157 to a depth of 1cm. As shown in Supplementary Table S9, the depth significantly impacts the predicted ECO157 counts. Therefore, our analysis indicates that using a 1cm depth was a conservative approach because the risk of romaine contamination from contaminated soil decreases as the depth of bacterial spread increases. Despite this, our assumptions about spread were considered reasonable based on the complexity of acquiring an accurate depth of ECO157 spread in these various media (runoff, swine feces and BSAAO) within 24 hours of their exposure to the field. Another limitation was that, while BSAAOs of poultry origin are the most common in fields across the USA, our study focused on those of dairy origin. Another assumption in the preharvest model was a constant and low-level ECO157 transfer from water and soil to romaine through furrow and drip irrigation. However, since these irrigation methods involve less direct contact with the romaine compared to overhead spray irrigation, we do not expect this assumption to have a major impact on the study conclusions —a finding also supported by the preharvest uncertainty analysis. Additionally, we did not account for the exterior leaves removed in the field, as we do not know the fraction of cells on these outer vs inner leaves. However, based on our assumptions, the impact of this practice was minimal as a 33.3% reduction in the ECO157 count on romaine resulted in 14.4% reduction in the median predicted illness cases (as detailed in Supplementary Table S10). Preharvest contamination was modeled at batch level, ignoring the spatial distribution of contamination among plants and soil in different locations of the field. However, we consider this to be a reasonable simplification since romaine was washed in batches, and the process of washing may further distribute ECO157 within a romaine batch. The preharvest model focused on open-field romaine production, excluding other potential production methods, such as greenhouses or vertical farming. This decision is expected to have minimal impact, as the majority of lettuce in the USA is grown in open fields *[135]*.

During postharvest, presence of cell clustering during partitioning — a poorly understood phenomenon in romaine packaging and servings — was predicted to moderately reduce illness cases (1% - 51% reduction). This reduction was due to the prediction of a lower ECO157 prevalence in batches, as clustering decreased the probability of consumers being exposed to servings with high ECO157 counts, as also observed by Zoellner et al. (2019). Clustering acted as a form of containment, minimizing the spread of contamination. For the industry, this insight may encourage mechanical processing that minimizes disturbance of clustered cells, preventing cells from spreading across the produce. For example, preventing cross-contamination during postharvest wash process (e.g., by adding chlorine), helps contain contamination within a single romaine batch, thereby reducing the risk of human illness (Table 4).

Conclusion

The risk assessment model presented here provides a comprehensive analysis of the impact of various preharvest and postharvest factors and their interactions. Our results suggest that illnesses occur mainly (52%) due to ECO157 contamination originating from untreated irrigation water applied through overhead spray irrigation. Our data further suggests that preharvest ECO157 risk, including that from irrigation, can be reduced either through water treatments or by switching to furrow or drip irrigation. In postharvest, maintaining the cold chain is crucial to prevent high numbers of illness cases. While our data indicate that an effective postharvest wash can help mitigate the public heath impact of both preharvest and postharvest contamination, future work on the impact of postharvest wash variability on public health risk is important. We identify knowledge gaps regarding the microbial quality of irrigation waters, as well as the effectiveness of preharvest surface water treatments and postharvest wash processes. In conclusion, the developed QMRA is expected to inform development of sustainable food safety strategies along the supply chain and guide further research to address key knowledge gaps.

References

1. Turner, K., Moua, C. N., Hajmeer, M., Barnes, A. & Needham, M. Overview of leafy greens–related food safety incidents with a California link: 1996 to 2016. *J. Food Prot.* **82**, 405–414 (2019).
2. CDC. NORS Dashboard. https://wwwn.cdc.gov/norsdashboard/ (2022).
3. CDC. Active investigations in foodborne outbreaks. https://www.cdc.gov/foodborne-outbreaks/active-investigations/all-foodborne-outbreak-notices.html (2024).
4. Coulombe, G., Catford, A., Martinez-Perez, A. & Buenaventura, E. Outbreaks of *Escherichia coli* O157:H7 infections linked to romaine lettuce in Canada from 2008 to 2018: An analysis of food safety context. *J. Food Prot.* **83**, 1444–1462 (2020). https://doi.org/10.4315/JFP-20-029
5. Benjamin, L. A. et al. Risk factors for *Escherichia coli* O157 on beef cattle ranches located near a major produce production region. *Epidemiol. Infect.* **143**, 81–93 (2015).
6. Jay, M. T. et al. *Escherichia coli* O157:H7 in feral swine near spinach fields and cattle, central California coast. *Emerg. Infect. Dis.* **13**, 1908 (2007).
7. California Food Emergency Response Team (CalFERT). Investigation of an *Escherichia coli* O157:H7 outbreak associated with Dole pre-packaged spinach. California Department of Health Services, Sacramento, CA (2007).
8. Cooley, M. B. et al. Development of a robust method for isolation of Shiga toxin-positive *Escherichia coli* (STEC) from fecal, plant, soil, and water samples from a leafy greens production region in California. *PLOS ONE* **8**, e65716 (2013). https://doi.org/10.1371/journal.pone.0065716

9. Atwill, E. R. et al. Transfer of *Escherichia coli* O157:H7 from simulated wildlife scat onto romaine lettuce during foliar irrigation. *J. Food Prot.* **78**, 240–247 (2015). https://doi.org/10.4315/0362-028X.JFP-14-277

10. Oliveira, M., Viñas, I., Usall, J., Anguera, M. & Abadias, M. Presence and survival of *Escherichia coli* O157:H7 on lettuce leaves and in soil treated with contaminated compost and irrigation water. *Int. J. Food Microbiol.* **156**, 133–140 (2012). https://doi.org/10.1016/j.ijfoodmicro.2012.03.014

11. Zhang, G. et al. Survey of foodborne pathogens, aerobic plate counts, total coliform counts, and *Escherichia coli* counts in leafy greens, sprouts, and melons marketed in the United States. *J. Food Prot.* **81**, 400–411 (2018).

12. Ackers, M. L. et al. An outbreak of *Escherichia coli* O157:H7 infections associated with leaf lettuce consumption. *J. Infect. Dis.* **177**, 1588–1593 (1998).

13. Hillborn, E. D. et al. A multistate outbreak of *Escherichia coli* O157:H7 infections associated with consumption of mesclun lettuce. *Arch. Intern. Med.* **159**, 1758–1764 (1999).

14. FAO. Prevention and control of microbiological hazards in fresh fruits and vegetables, Part 4: Specific commodities. https://openknowledge.fao.org/server/api/core/bitstreams/151c9dda-508b-4b70-972a-85725eb12145/content (2023).

15. Yang, Y., Luo, Y., Millner, P., Turner, E. & Feng, H. Assessment of *Escherichia coli* O157:H7 transference from soil to iceberg lettuce via a contaminated field coring harvesting knife. *Int. J. Food Microbiol.* **153**, 345–350 (2012).

16. Buchholz, A. L., Davidson, G. R., Marks, B. P., Todd, E. C., Ryser, E. T. Transfer of *Escherichia coli* O157:H7 from equipment surfaces to fresh-cut leafy greens during processing in a model pilot-plant production line with sanitizer-free water. *J. Food Prot.* **75**, 1920–1929 (2012). https://doi.org/10.4315/0362-028X.JFP-11-558

17. López-Gálvez, F., Allende, A., Selma, M. V. & Gil, M. I. Prevention of *Escherichia coli* cross-contamination by different commercial sanitizers during washing of fresh-cut lettuce. *Int. J. Food Microbiol.* **133**, 167–171 (2009). https://doi.org/10.1016/j.ijfoodmicro.2009.05.017

18. USDA ERS. Food availability per capita data system. https://www.ers.usda.gov/data-products/food-availability-per-capita-data-system/ (2020).

19. California LGMA. About us. https://lgma.ca.gov/about-us (2024).

20. Arizona LGMA. About us. https://www.arizonaleafygreens.org/about-us (2023).

21. FDA. FSMA Final Rule on Produce Safety: Standards for the Growing, Harvesting, Packing, and Holding of Produce for Human Consumption. https://www.fda.gov/food/food-safety-modernization-act-fsma/fsma-final-rule-produce-safety (2024).

22. FSMA PSR § 112.41, § 112.42. https://www.ecfr.gov/current/title-21/chapter-I/subchapter-B/part-112 (2024).

23. CFR § 205.203. https://www.ecfr.gov/current/title-7/subtitle-B/chapter-I/subchapter-M/part-205/subpart-C/section-205.203 (2024).

24. CFR § 112.55. https://www.ecfr.gov/current/title-21/chapter-I/subchapter-B/part-112/subpart-F/section-112.55 (2024).

25. CFR § 112.83 and §112.112. https://www.regulations.gov/document/FDA-2011-N-0921-18558 (2024).

26. FSMA PSR. https://www.regulations.gov/document/FDA-2011-N-0921-18558 (2015).

27. Allende, A., Truchado, P., Lindqvist, R. & Jacxsens, L. 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 Microbiol.* **75**, 82–89 (2018).

28. Koseki, S. & Isobe, S. Prediction of pathogen growth on iceberg lettuce under real temperature history during distribution from farm to table. *Int. J. Food Microbiol.* **104**, 239–248 (2005).

29. Franz, E., Semenov, A. V. & van Bruggen, A. H. Modelling the contamination of lettuce with *Escherichia coli* O157:H7 from manure-amended soil and the effect of intervention strategies. *J. Appl. Microbiol.* **105**, 1569–1584 (2008).

30. Franz, E., Tromp, S. O., Rijgersberg, H. & Van Der Fels-Klerx, H. J. Quantitative microbial risk assessment for *Escherichia coli* O157:H7, *Salmonella*, and *Listeria monocytogenes* in leafy green vegetables consumed at salad bars. *J. Food Prot.* **73**, 274–285 (2010).

31. Tromp, S. O., Rijgersberg, H. & Franz, E. Quantitative microbial risk assessment for *Escherichia coli* O157:H7, *Salmonella enterica*, and *Listeria monocytogenes* in leafy green vegetables consumed at salad bars, based on modeling supply chain logistics. *J. Food Prot.* **73**, 1830–1840 (2010).

32. Danyluk, M. D. & Schaffner, D. W. Quantitative assessment of the microbial risk of leafy greens from farm to consumption: preliminary framework, data, and risk estimates. *J. Food Prot.* **74**, 700–708 (2011).

33. Rodríguez, F. P. et al. A mathematical risk model for *Escherichia coli* O157:H7 cross-contamination of lettuce during processing. *Food Microbiol.* **28**, 694–701 (2011).

34. Ottoson, J. R., Nyberg, K., Lindqvist, R. & Albihn, A. Quantitative microbial risk assessment for *Escherichia coli* O157 on lettuce, based on survival data from controlled studies in a climate chamber. *J. Food Prot.* **74**, 2000–2007 (2011).

35. Castro-Ibáñez, I., Gil, M. I., Tudela, J. A., Ivanek, R. & Allende, A. Assessment of microbial risk factors and impact of meteorological conditions during production of baby spinach in the Southeast of Spain. *Food Microbiol.* **49**, 173–181 (2015).

36. Mishra, A., Pang, H., Buchanan, R. L., Schaffner, D. W. & Pradhan, A. K. A system model for understanding the role of animal feces as a route of contamination of leafy greens before harvest. *Appl. Environ. Microbiol.* **83**, e02775-16 (2017).

37. Kundu, A., Wuertz, S. & Smith, W. A. Quantitative microbial risk assessment to estimate the risk of diarrheal diseases from fresh produce consumption in India. *Food Microbiol.* **75**, 95–102 (2018).

38. Pang, H, Lambertini, E., Buchanan, R. L., Schaffner, D. W. & Pradhan, A. K. Quantitative microbial risk assessment for *Escherichia coli* O157:H7 in fresh-cut lettuce. *J. Food Prot.* **80**, 302–311 (2017).

39. O'Flaherty, E., Solimini, A. G., Pantanella, F., De Giusti, M. & Cummins, E. Human exposure to antibiotic-resistant *Escherichia coli* through irrigated lettuce. *Environ. Int.* **122**, 270–280 (2019).

40. Rock, C. M. et al. Review of water quality criteria for water reuse and risk-based implications for irrigated produce under the FDA Food Safety Modernization Act, produce safety rule. *Environ. Res.* **172**, 616–629 (2019).

41. Bozkurt, H., Bell, T., van Ogtrop, F., Phan-Thien, K. Y. & McConchie, R. Assessment of microbial risk during Australian industrial practices for *Escherichia coli* O157:H7 in fresh-cut cos lettuce: A stochastic quantitative approach. *Food Microbiol.* **95**, 103691 (2021).

42. Nauta, M. J. Microbiological risk assessment models for partitioning and mixing during food handling. *Int. J. Food Microbiol.* **100**, 311–322 (2005).

43. Munther, D. S., Carter, M. Q., Aldric, C. V., Ivanek, R. & Brandl, M. T. Formation of *Escherichia coli* O157:H7 persister cells in the lettuce phyllosphere and application of differential equation models to predict their prevalence on lettuce plants in the field. *Appl. Environ. Microbiol.* **86**, e01602-19 (2020).

44. Turini, T. et al. Iceberg lettuce production in California. https://escholarship.org/uc/item/7w47j6zv (2011).

45. Scallan, E. et al. Foodborne illness acquired in the United States—major pathogens. *Emerg. Infect. Dis.* **17**, 7–15 (2011).

46. Painter, J. A. et al. Attribution of foodborne illnesses, hospitalizations, and deaths to food commodities by using outbreak data, United States, 1998–2008. *Emerg. Infect. Dis.* **19**, 407–415 (2013).

47. Rural Migration News (RMN). https://migration.ucdavis.edu/rmn/blog/post/?id=2699 (2022).

48. Agriculture Marketing Resource Center. Lettuce. https://www.agmrc.org/commodities-products/vegetables/lettuce (2021).

49. Geisseler, D. & Horwath, W. R. Lettuce production in California. https://www.cdfa.ca.gov/is/ffldrs/frep/FertilizationGuidelines/pdf/Lettuce_Production_CA.pdf (2016).

50. Brouwer, A. F. et al. Modeling biphasic environmental decay of pathogens and implications for risk analysis. *Environ. Sci. Technol.* **51**, 2186–2196 (2017).

51. Bezanson, G. et al. Comparative examination of *Escherichia coli* O157:H7 survival on romaine lettuce and in soil at two independent experimental sites. *J. Food Prot.* **75**, 480–487 (2012).

52. Erickson, M. C. et al. Surface and internalized *Escherichia coli* O157:H7 on field-grown spinach and lettuce treated with spray-contaminated irrigation water. *J. Food Prot.* **73**, 1023–1029 (2010).

53. Moyne, A. L. et al. Fate of *Escherichia coli* O157:H7 in field-inoculated lettuce. *Food Microbiol.* **28**, 1417–1425 (2011).

54. Williams, T. R., Moyne, A. L., Harris, L. J. & Marco, M. L. Season, irrigation, leaf age, and *Escherichia coli* inoculation influence the bacterial diversity in the lettuce phyllosphere. *PLoS One* **8**, e68642 (2013).

55. Chen, Z. et al. Prevalence of *Escherichia coli* O157 and *Salmonella* spp. in solid bovine manure in California using real-time quantitative PCR. *Lett. Appl. Microbiol.* **69**, 23–29 (2019).

56. Rogers, S. W. et al. Decay of bacterial pathogens, fecal indicators, and real-time quantitative PCR genetic markers in manure-amended soils. *Appl. Environ. Microbiol.* **77**, 4839–4848 (2011).

57. Chen, H. Contribution of the persister state to the tolerance of *Escherichia coli* O157:H7 to stresses encountered within the fresh produce chain. Master's Thesis, University of Guelph (2020).

58. Aqua-Calc. https://www.aqua-calc.com/ (2023).

59. Valenzuela, H. R., Kratky, B. & Cho, J. Lettuce production guidelines for Hawaii. https://www.ctahr.hawaii.edu/oc/freepubs/pdf/res-164.pdf (1995).

60. California Department of Fish and Wildlife. Wild pig. https://wildlife.ca.gov/Hunting//Wild-Pig (2001).

61. Hamilton, D. W., Luce, W. G. & Heald, A. D. Production and characteristics of swine manure. Oklahoma State University Extension Facts Publication F-1735 (1997).

62. Ferretti, F., Storer, K., Coats, J. & Massei, G. Temporal and spatial patterns of defecation in wild boar. *Wildl. Soc. Bull.* **39**, 65–69 (2015).

63. USGS. Freshwater withdrawals in the United States. https://www.usgs.gov/special-topics/water-science-school/science/freshwater-withdrawals-united-states (2018).

64. Smith, B. A., Fazil, A. & Lammerding, A. M. A risk assessment model for *Escherichia coli* O157:H7 in ground beef and beef cuts in Canada: Evaluating the effects of interventions. *Food Control* **29**, 2 (2013).

65. Benjamin, L. et al. Occurrence of generic *Escherichia coli*, *E. coli* O157, and *Salmonella* spp. in water and sediment from leafy green produce farms and streams on the Central California coast. *Int. J. Food Microbiol.* **165**, 65–76 (2013).

66. Stine, S. W. et al. Application of microbial risk assessment to the development of standards for enteric pathogens in water used to irrigate fresh produce. *J. Food Prot.* **68**, 913–918 (2005).

67. Allende, A. et al. Quantitative contamination assessment of *Escherichia coli* in baby spinach primary production in Spain: Effects of weather conditions and agricultural practices. *Int. J. Food Microbiol.* **257**, 238–246 (2017).

68. FDA. Standards for the Growing, Harvesting, Packing, and Holding of Produce for Human Consumption. Federal Register. https://www.federalregister.gov/documents/2015/11/27/2015-28159/standards-for-the-growing-harvesting-packing-and-holding-of-produce-for-human-consumption (2015).

69. Alzahrani, A. Persister formation and revival in Shiga Toxin-Producing *Escherichia coli* subjected to environmental stress conditions. *PhD Thesis, University of Guelph.* *https://atrium.lib.uoguelph.ca/server/api/core/bitstreams/e5a584c9-3e95-41a1-aa4a-7e0abf5e0c42/content* (2020).

70. Hofsteenge, N., van Nimwegen, E. & Silander, O.K. Quantitative analysis of persister fractions suggests different mechanisms of formation among environmental isolates of *E. coli*. *BMC Microbiol* **13**, 25 (2013).

71. Niu, H., Gu, J. & Zhang, Y. Bacterial persisters: molecular mechanisms and therapeutic development. *Sig. Transduct. Target. Ther.* **9**, 174 (2024).

72. Wang, D. et al. Comparative persistence of *Salmonella* and *Escherichia coli* O157:H7 in loam or sandy loam soil amended with bovine or swine manure. *Can. J. Microbiol.* **64**, 979–991 (2018).

73. Tran, D. T. Q. et al. Environmental drivers for persistence of *Escherichia coli* and *Salmonella* in manure-amended soils: A meta-analysis. *J. Food Prot.* **83**, 1268–1277 (2020).

74. McKellar, R. C., LeBlanc, D. I., Rodríguez, F. P. & Delaquis, P. Comparative simulation of *Escherichia coli* O157:H7 behaviour in packaged fresh-cut lettuce distributed in a typical Canadian supply chain in the summer and winter. *Food Control* **35**, 192–199 (2014).

75. McKellar, R. C. LeBlanc, D. I., Lu, J. & Delaquis, P. Simulation of *Escherichia coli* O157:H7 behavior in fresh-cut lettuce under dynamic temperature conditions during distribution from processing to retail. *Foodborne Pathog. Dis.* **9**, 239–244 (2012).

76. When to harvest vegetables. University of Georgia Extension. https://extension.uga.edu/publications/detail.html?number=C935&title=when-to-harvest-vegetables (2022).

77. Keskinen, L. A., Burke, A. & Annous, B. A. Efficacy of chlorine, acidic electrolyzed water and aqueous chlorine dioxide solutions to decontaminate *Escherichia coli* O157:H7 from lettuce leaves. *Int. J. Food Microbiol.* **132**, 134–140 (2009).

78. Zoellner, C., Wiedmann, M. & Ivanek, R. An assessment of Listeriosis risk associated with a contaminated production lot of frozen vegetables consumed under alternative consumer handling scenarios. *J. Food Prot.* **82**, 2174–2193 (2019).

79. Cassin, M. H., Lammerding, A. M., Todd, E. C., Ross, W., & McColl, R. S. Quantitative risk assessment for *Escherichia coli* O157:H7 in ground beef hamburgers. *Int. J. Food Microbiol.* **41**, 21–44 (1998).

80. U.S. Census Bureau. Census Bureau projects U.S. population of 312.8 million on New Year's Day. https://www.census.gov/newsroom/releases/archives/population/cb11-219.html (2011).

81. USDA ERS. Food Availability (Per Capita) Data System. https://www.ers.usda.gov/data-products/food-availability-per-capita-data-system/ (2011).

82. Wickham, H., François, R., Henry, L., Müller, K. & Vaughan, D. dplyr: A Grammar of Data Manipulation. R package version 1.1.4. https://dplyr.tidyverse.org. (2023).

83. Akoglu, H. User's guide to correlation coefficients. *Turk. J. Emerg. Med.* **18**, 91–93 (2018).

84. Mukaka, M. M. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Med. J.* **24**, 69–71 (2012).

85. Therneau, T., Atkinson, B. & Ripley, B. rpart: Recursive Partitioning and Regression Trees. https://cran.r-project.org/package=rpart (2018).

86. Longfellow's Greenhouses. http://plants.longfellowsgreenhouses.com/12100007/Plant/Print?id=29076.

87. Mir, R. et al. Cattle intestinal microbiota shifts following *Escherichia coli* O157:H7 vaccination and colonization. *PLOS ONE* **14**, e0226099 (2019).

88. Gravuer, K. Compost application rates for California croplands and rangelands for a CDFA Healthy Soils Incentives Program. https://www.cdfa.ca.gov/oefi/healthysoils/docs/CompostApplicationRate_WhitePaper.pdf (2016).

89. Beauvais, W. et al. The effectiveness of treating irrigation water using ultraviolet radiation or sulphuric acid fertilizer for reducing generic *Escherichia coli* on fresh produce – a controlled intervention trial. *J. Appl. Microbiol.* **131**, 1360–1377 (2021).

90. Chang, T. Evaluation of multiple disinfection methods to mitigate contaminated irrigation water. Master's Thesis, University of Tennessee https://trace.tennessee.edu/utk_gradthes/3352 (2015).

91. Dery, J. L., Gerrity, D. & Rock, C. Minimizing Risks: Use of surface water in pre-harvest agricultural irrigation; part ii: sodium and calcium hypochlorite (chlorine) treatment

methods. https://repository.arizona.edu/bitstream/handle/10150/669941/AZ1831-2020.pdf?sequence=1&isAllowed=y (2011).

92. Girardin, L. et al. Behaviour of the pathogen surrogates Listeria innocua and Clostridium sporogenes during production of parsley in fields fertilized with contaminated amendments. *FEMS Microbiol. Ecol.* ***54***, 287-295 (2005).

93. Rodríguez, F. P. et al. A mathematical risk model for *Escherichia coli* O157:H7 cross-contamination of lettuce during processing. *Food Microbiol.* ***28***, 694–701 (2011).

94. Zhang, G., Ma, L., Phelan, V. H., Doyle & M. P. Efficacy of antimicrobial agents in lettuce leaf processing water for control of *Escherichia coli* O157:H7. *J. Food Prot.* **72**, 1392–1397 (2009).

95. EcoSure 2007 Cold Temperature Database. https://www.foodrisk.org/resources/display/21

96. Latorre, A. et al. Quantitative risk assessment of listeriosis due to consumption of raw milk. *J. Food Prot.* **74**, 1268–1281 (2011).

97. Pouillot, R., Lubran, M. B., Cates, S. C. & Dennis, S. Estimating parametric distributions of storage time and temperature of ready-to-eat foods for U.S. households. *J. Food Prot.* **73**, 312–321 (2010).

98. USDA. *Lettuce, Cos or Romaine, Raw; Portion*. https://fdc.nal.usda.gov/food-details/746769/nutrients/ (2015).

99. Verrill, L. Consumer vegetable and fruit washing practices in the United States, 2006 and 2010. *Food Prot. Trends* **32**, 164-172 (2012).

100. Climate Data Online. https://www.ncei.noaa.gov/cdo-web/

101. FDA, University of California Davis. *Head Lettuce*. https://www.wifss.ucdavis.edu/wp-content/uploads/2016/10/HeadLettuce_PDF.pdf

102. USDA ERS. *Vegetables and Pulses Outlook* (2023). Available at: https://downloads.usda.library.cornell.edu/usda-esmis/files/cj82k729x/pn89fk86d/3f463k190/VGS-370.pdf

103. Ivanek, R., Gröhn, Y. T., Wiedmann, M. & Wells, M. T. Mathematical model of *Listeria monocytogenes* cross-contamination in a fish processing plant. *J. Food Prot.* **67**, 2688–2697 (2004).

104. de Freitas Costa, E., Corbellini, L. G., da Silva, A. P. S. P. & Nauta, M. A stochastic model to assess the effect of meat inspection practices on the contamination of pig carcasses. *Risk Anal.* **37**, 1849–1864 (2017).

105. Ivanek, R., Snary, E. L., Cook, A. J., & Gröhn, Y. T. A mathematical model for the transmission of *Salmonella Typhimurium* within a grower-finisher pig herd in Great Britain. *J. Food Prot.* **67**, 2403–2409 (2004).

106. Karp, D. S. et al. Co-managing fresh produce for nature conservation and food safety. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 11126–11131 (2015).

107. Gelting, R. J. et al. Irrigation water issues potentially related to the 2006 multistate *E. coli* O157:H7 outbreak associated with spinach. *Agric. Water Manag.* **98**, 1395–1402 (2011).

108. Bottichio, L. et al. Shiga toxin-producing *Escherichia coli* infections associated with romaine lettuce—United States, 2018. *Clin. Infect. Dis.* **71**, e323–e330 (2020).

109. Gelting, R. J. et al. A systems analysis of irrigation water quality in an environmental assessment of an *E. coli* O157:H7 outbreak in the United States linked to iceberg lettuce. *Agric. Water Manag.* **150**, 11–118 (2015).

110. FDA. *Environmental Assessment Factors Potentially Contributing to Contamination of Romaine Lettuce Implicated in Outbreaks*. (2019). Available at: https://www.fda.gov/food/outbreaks-foodborne-illness/environmental-assessment-factors-potentially-contributing-contamination-romaine-lettuce-implicated#summary

111. Allende, A. & Monaghan, J. Irrigation water quality for leafy crops: a perspective of risks and potential solutions. *Int. J. Environ. Res. Public Health* **12**, 7457–7477 (2015).

112. Dery, J. L., Choppakatla, V., Sughroue, J. & Rock, C. Minimizing Risks: Use of Surface Water in Pre-Harvest Agricultural Irrigation; Part III: Peroxyacetic Acid (PAA) Treatment Methods. Available at: https://extension.arizona.edu/sites/extension.arizona.edu/files/pubs/az1884-2021.pdf

113. Banach, J. L. et al. Application of water disinfection technologies for agricultural waters. *Agric. Water Manag.* **244**, 106527 (2021).

114. Collvignarelli, M. et al. Overview of the main disinfection processes for wastewater and drinking water treatment plants. *Sustainability* **10**, 86 (2018).

115. Rock, C. CPS Rapid Response – Yuma Valley, Final Report. *Center for Produce Safety* (2019).

116. Rock, C., Cooper, K. & Yemmireddy, V. K. Agriculture Water Treatment – Southwest Region Final Project Report. Center for Produce Safety (2021).

117. USEPA. Peracetic Acid Fact Sheet https://www.epa.gov/sites/default/files/2019-08/documents/disinfection_-_paa_fact_sheet_-_2012.pdf (2012).

118. *USEPA.* Chlorination Fact Sheet. https://www3.epa.gov/npdes/pubs/chlo.pdf (1999).

119. *USEPA.* UV Disinfection Fact Sheet. https://www3.epa.gov/npdes/pubs/uv.pdf (1999).

120. Fonseca, J. M. et al. *Escherichia coli* survival in lettuce fields following its introduction through different irrigation systems. *J. Appl. Microbiol.* **110**, 893–902 (2011).

121. Moyne, A. L. et al. Fate of *Escherichia coli* O157:H7 in field-inoculated lettuce. *Food Microbiol.* **28**, 1417–1425 (2011).

122. Solomon, E. B., Potenski, C. J. & Matthews, K. R. Effect of irrigation method on transmission to and persistence of *Escherichia coli* O157:H7 on lettuce. *J. Food Prot.* **65**, 673–676 (2002).

123. Gil, M. I., Tudela, J. A., Luna, M. C. & Allende, A. Water management and its effect on the postharvest quality of fresh-cut vegetables. *Stewart Postharvest Rev.* **9**, 1–8 (2013).

124. Gurtler, J. B. & Gibson, K. E. Irrigation water and contamination of fresh produce with bacterial foodborne pathogens. *Curr. Opin. Food Sci.* **47**, 100889 (2022).

125. Song, I., Stine, S., Choi, C. & Gerba, C. Comparison of crop contamination by microorganisms during subsurface drip and furrow irrigation. *J. Environ. Eng.* **132**, 1243–1248 (2006).

126. Park, S. et al. Multifactorial effects of ambient temperature, precipitation, farm management, and environmental factors determine the level of generic *Escherichia coli* contamination on preharvested spinach. *Appl. Environ. Microbiol.* **81**, 2635–2650 (2015).

127. Holvoet, K., Sampers, I., Seynnaeve, M., & Uyttendaele, M. Relationships among hygiene indicators and enteric pathogens in irrigation water, soil and lettuce and the impact of climatic conditions on contamination in the lettuce primary production. *Int. J. Food Microbiol.* **171**, 21–31 (2014).

128. Erickson, M. C. et al. Preharvest internalization of *Escherichia coli* O157:H7 into lettuce leaves, as affected by insect and physical damage. *J. Food Prot.* **73**, 1809–1816 (2010).

129. Petri, E., Rodríguez, M. & García, S. Evaluation of combined disinfection methods for reducing *Escherichia coli* O157:H7 population on fresh-cut vegetables. *Int. J. Environ. Res. Public Health* **12**, 8678–8690 (2015).

130. Mokhtari, A. et al. Evaluation of chlorine dioxide and peracetic acid on *Escherichia coli* O157:H7 reduction on lettuce. *Food Control* **62**, 36–40 (2016).

131. Fu, T.-J., Li, Y., Awad, D., Zhou, T.-Y. & Liu, L. Factors affecting the performance and monitoring of a chlorine wash in preventing *Escherichia coli* O157:H7 cross-contamination during postharvest washing of cut lettuce. *Food Control* **94**, 212–221 (2018).

132. Gombas, D. et al. Guidelines to validate control of cross-contamination during washing of fresh-cut leafy vegetables. *J. Food Prot.* **80**, 312–330 (2017).

133. Luo, Y. et al. Determination of free chlorine concentrations needed to prevent *Escherichia coli* O157:H7 cross-contamination during fresh-cut produce wash. *J. Food Prot.* **74**, 352–358 (2011).

134. Gómez-López, V. M., Lannoo, A. S., Gil, M. I. & Allende, A. Minimum free chlorine residual level required for the inactivation of *Escherichia coli* O157:H7 and trihalomethane generation during dynamic washing of fresh-cut spinach. *Food Control* **42**, 132–138 (2014).

135. USDA ERS, Vegetables and Pulses Outlook: April 2024 https://ers.usda.gov/sites/default/files/_laserfiche/outlooks/109067/VGS-372.pdf?v=19742 (2024).

Acknowledgements

This work was supported through a grant (award 2019-51181-30016) from the U.S. Department of Agriculture (USDA), National Institute of Food and Agriculture (NIFA). Partial support received from a grant (award 2020-67021-32855) from the USDA NIFA.

Author contributions

E.B.: conceptualization, data curation, formal analysis, investigation, methodology, software, validation, visualization, writing- original draft, writing – review and editing; S.I.M.: conceptualization, data curation, methodology, software, writing – review and editing; L.K.S.: conceptualization, funding acquisition, writing – review and editing; M.D.D.: conceptualization, funding acquisition, writing – review and editing; M.W.: conceptualization, funding acquisition, writing – review and editing; R.I.: conceptualization, funding acquisition, investigation, methodology, project administration, software, supervision, validation, visualization, writing – review and editing. All authors have seen and approved the manuscript.

Data availability statement

The code associated with this project is stored in the GitHub repository linked here: https://github.com/IvanekLab/Romaine_ECO157_QMRA and is publicly accessible.

Competing interests

The authors declare no competing interests.