FDA-iRISK® 5.0
Food Safety Modeling Tool

Technical Document

March 2026

Disclaimer

The U.S. Food and Drug Administration (FDA), the Joint Institute for Food Safety and Applied Nutrition (JIFSAN) and Risk Sciences International (RSI) have taken all reasonable precautions in creating the FDA-iRISK® quantitative risk assessment system (version 5.0) and the documentation accompanying it. FDA, JIFSAN and RSI are not responsible for errors, omissions or deficiencies regarding the system and the accompanying documentation. The FDA-iRISK system and the accompanying documentation are being made available “as is” and without warranties of any kind, either expressed or implied, including, but not limited to, warranties of performance, merchantability, and fitness for a particular purpose. FDA, JIFSAN and RSI are not making a commitment in any way to regularly update the system and the accompanying documentation.

Responsibility for the interpretation and use of the system and of the accompanying documentation lies solely with the user. Risk scenarios and data provided in the FDA-iRISK system are for illustration purposes only; they do not represent endorsement by FDA, JIFSAN, or RSI. In no event shall FDA, JIFSAN or RSI be liable for direct, indirect, special, incidental, or consequential damages resulting from the use, misuse, or inability to use the system and the accompanying documentation.

Third parties' use of or acknowledgment of the system and its accompanying documentation, including through the suggested citation, does not in any way represent that FDA, JIFSAN or RSI endorses such third parties or expresses any opinion with respect to their statements.

March 2026

Contents

1 Overview

Stakeholders in the system of food safety, in particular government agencies, need evidence-based, transparent, and rigorous approaches to estimate and compare the risk of foodborne illness from microbial and chemical hazards. FDA-iRISK® is a web-based software tool intended for relatively rapid assessment of the risks associated with microbiological and chemical hazards in food. The tool provides a step-wise data entry, documentation, computing, and reporting environment. In this environment, a user can develop risk scenarios that describe various key aspects of the hazard, the food, and the processing of the food as it relates to the fate of the hazard within the food. An FDA-iRISK scenario includes seven elements: the food, the hazard, the population of consumers, a process pathway (i.e., food production, processing and handling practices), consumption patterns in the population, dose response relationships, and burden of disease measures associated with health effects (e.g., losses in disability-adjusted life years, or DALYs). Once the user has described these key elements, the tool can combine the user's input into a quantitative risk assessment model (i.e., a risk scenario) that estimates the risk of illness or health burden to the consumer. The results of this model are presented to the user in the form of reports (e.g., in PDF format), allowing the user to study the implications of the food and hazard properties that they have entered. For risk assessment model development, FDA-iRISK is intended for users, as well as a team of users with different expertise who collaborate, who are *knowledgeable about the hazards, foods, and processes that they are describing*, but who may not be familiar with risk assessment methodology, particularly as it pertains to developing quantitative estimates of risk.

The tool provides for rapid, quantitative risk assessment. The assessment is rapid in the sense that the user can control, to an extent, the level of complexity of the model. In addition, by maintaining a structured database of the user's previous work, and by allowing copying and sharing of risk assessment model elements, significant efficiencies in the overall conduct of risk assessment activities can be gained by an individual user, or groups of users who choose to share model elements amongst each other.

It is important to understand that FDA-iRISK itself does not contain or provide any scientific data other than what has been entered explicitly by the user. Users of FDA-iRISK provide all of the data, assumptions, and knowledge about hazards and foods. The purpose of FDA-iRISK is to provide an appropriate database and computational infrastructure to support a majority of the types of calculations typically required in risk assessments applied to food safety. A key design principle behind FDA-iRISK is that the combination of the user's technical knowledge and the reliability associated with the computational infrastructure should ensure higher quality and more productive risk assessment activity. Key among the benefits is the avoidance of common conceptual and mathematical challenges that can make quantitative risk assessment either too difficult or too error-prone for some potential users.

As part of the computational infrastructure, FDA-iRISK allows for most quantitative parameters in a model to be characterized in the form of probability distributions, intended to describe the variability in various aspects of the system being described. When the user includes variability distributions, FDA-iRISK performs Monte Carlo simulation to combine the impact of variability from all user inputs into the final estimates of risk. FDA-iRISK also provides the option of specifying quantitative descriptions of uncertainty, or so-called Second Order Monte Carlo simulation, to the majority of model parameters.

A key design driver for FDA-iRISK is to support comparative risk assessment. A key challenge in comparative risk assessment is the generation of estimates that stem from very diverse hazards, which may have different exposure patterns and very diverse health consequences ranging from very mild to fatal. This requires a common measure of risk that can be compared across both acute and chronic hazards and both microbial and chemical hazard types.

A risk estimate can be generally described as having a numerator and a denominator. The numerator generally describes the extent of harm. The denominator describes the context (e.g., the timeframe, the number of people, the amount of food consumed, etc.) in which the harm occurs. There are a great variety of combinations of numerator and denominators, for example:

- Cases of illness per year
- Cases of illness per million servings of food
- Fatalities per million persons per year
- Lifetime probability of cancer per consumer

In different contexts, each of these combinations has potential value to support different types of decisions and comparisons. The current choices of numerator in FDA-iRISK are Disability-Adjusted Life Years (DALYs), Quality-Adjusted Life Year (QALY) Loss, and Cost of Illness (COI). These choices are justified and described below.

Note: Users may override the default risk estimates for ranking purposes when generating reports (see *Section 1.4 Ranking Options*).

1.1 Choices for Numerator (DALYs, COI, or QALY Loss)

A wide variety of hazards and associated health outcomes are associated with foodborne hazards. To accommodate this variety, two composite measures of harm are included in FDA-iRISK:

- Disability-Adjusted Life Years (DALYs): This measure has been used internationally to compare the burden of disease for a variety of health outcomes. As a health metric, the DALY integrates the severity and duration of health outcomes, and the relative frequency of each outcome, and provides a measure that accommodates both non-fatal and fatal

outcomes. The DALY measure is very similar in concept to the measure of Quality-Adjusted Life Years (QALYs), but is more common in the current food safety literature (Havelaar et al., 2012; GBD 2019 Diseases and Injuries Collaborators. 2020).

- Quality-Adjusted Life Years (QALY) Loss: QALY is a measure of disease burden (Batz et al., 2014) similar in many respects to DALYs. However, where DALYs quantify health burden in terms of increasing disability, QALYs quantify health burden in terms of decreasing utility. With DALYs, a severity of 1 indicates death where a severity of 0 indicates perfect health. Conversely, with QALYS a utility of 1 indicates perfect health where a utility of 0 indicates death. For compatibility with how FDA-iRISK computes risk, QALY Loss is used instead of QALY.
- Cost-of-Illness (COI): This measure allows for the accumulation of the economic cost of illness as a composite measure that integrates the health burden and other societal burdens of illness such as lost productivity, medical costs, and other economic indicators of societal burden (Minor et al., 2015.

In all cases, the user specifies the array of individual outcomes or cost category to be considered, and assigns appropriate parameters to quantify the severity and duration (for DALYs), the utility loss and duration (for QALY Loss), and the cost per case (for COI) for each outcome or cost category considered. In the case of DALYs and QALY Loss, the frequency of the health outcome is also specified by the user. The frequency of each health outcome is used to provide a frequency-weighted burden, or average burden per case.

The average DALY per case is given by:

$$DALY = \sum_{j} S_j \times D_j \times w_j$$

Equation 1

where:

- S_j is the severity of health effect j for a given hazard, expressed on a scale from 0 (no disability) to 1 (death).
- D_j is the duration of health effect j, expressed in years. In the case of death, duration is expressed as years of life lost based on the age of the person affected, and severity is set to the maximum value of 1.0.
- w_j is the fraction of cases in which health endpoint j occurs.

For example, a DALY for liver cancer might be based on a combination of morbidity and mortality endpoints:

Table 1_1. DALY calculation based on health endpoints

Health Endpoint	Severity	Duration (years)	Fraction of Cases	DALY for Endpoint
Morbidity: non-fatal liver cancer	0.2	15.1	0.05	0.1510
Morbidity: fatal liver cancer	0.56	0.4	0.95	0.2128
Mortality: fatal liver cancer	1	20	0.95	19.000
			Total DALY per case:	19.3638

For QALY Loss calculations, the average QALY Loss per case is given by:

$$QALYLoss = \sum_j UL_j \times D_j \times w_j$$

Equation 2

where:

- UL_j is the loss of utility of health effect j for a given hazard, expressed on a scale from 0 (no loss) to 1 (death).
- D_j is the duration of health effect j, expressed in years. In the case of death, duration is expressed as years of life lost based on the age of the person affected, and utility loss is set to the maximum value of 1.0.
- w_j is the fraction of cases in which health endpoint j occurs.

For COI calculations, the average cost per case is given by:

$$COI = \sum_j C_j \times w_j$$

Equation 3

where:

- C_j is the cost per cost category (currency is unspecified).
- w_j is the fraction of cases for which this cost would be expected to be incurred.

The average DALY per case, QALY Loss per case, or the average COI per case, is then multiplied by the number of cases of illness predicted by the FDA-iRISK simulation model to yield the overall burden of disease.

1.1.1 Applying External Sources of DALY, QALY Loss, and COI Estimates

If an external source is available that provides an estimate for the average DALY loss per case of illness, QALY Loss per case, or the average cost per case of illness, the user has the option of entering this number directly (i.e., without specifying the individual health outcomes or cost categories). This is treated the same way as the average DALY per case, average QALY Loss per case, or average COI per case computed as described above.

1.1.2 Number of Illness as a Numerator in FDA-iRISK

In addition to the built-in health metrics, users can choose illnesses for risk ranking. Users can also use a value of 1 for the burden of disease (e.g., DALY) so that the numerator in the risk estimate is equivalent to the number of cases of illness. This is useful for comparing to other estimates of the number of cases. However, while the number of illnesses can be used as a metric for risk ranking for one hazard across different foods, due to the great diversity in the harm associated with different hazards and their associated health outcomes, the DALY, QALY Loss, or COI is recommended for use in ranking exercises across different hazards.

1.2 Choice for Denominator (per Year)

In addition to providing a common measuring stick for the harm associated with foodborne disease, it is necessary to provide a common context. For example, it would be inappropriate to compare two numerical results where one is expressed as cases per million persons and the other is expressed as cases per year.

FDA-iRISK employs the common denominator in units of time, specifically one year in a user-specified population. For chronic hazards, where risk accumulates over a lifetime of exposure, the lifetime risk is divided by the total duration of the user-specified life-stages in the population to yield an annualized risk. This measure indicates the amount of the overall risk that can be attributed to each year of consumption, on average over the lifetime.

FDA-iRISK provides users with an option to override the default method for chronic hazards. That is, if only risk scenarios for chronic hazards are selected, the user can instruct FDA-iRISK to report the lifetime risk instead of the annualized risk. This option is not available when risk scenarios for acute and chronic hazards are combined in the same ranking report.

1.3 The Overall Calculation of Risk

For each food-hazard combination consisting of food f and hazard h, the burden of disease is given by:

$$Burden_{f,h} = \frac{Burden_h \times P_{f,h} \times S_f}{T}$$

Equation 4

where:

- $Burden_h$ is the average burden (in DALYs, QALYs or units of currency) per case of illness for hazard h.
- $P_{f,h}$ is the probability of a case of illness for the food–hazard combination f, h (defined below).
- S_f scales the result according to the number of consumers (for chronic exposures) or the number of eating occasions (for acute exposures), and is equal to the user-defined number of consumers for chronic exposures or the number of annual eating occasions for acute exposures, for food f. (The amount of consumption, e.g., serving sizes, average intakes, is included in the estimate of dose).
- T is used to provide a comparable time scale. For chronic hazards resulting from cumulative exposure, the value is generally annualized by dividing by the total duration of exposure (i.e., T = total lifespan). However, for risk scenarios for chronic hazards, T=1 when the user instructs FDA-iRISK to not annualize the results. T=1 for acute hazards in all cases.

This measure incorporates both the numerator (the burden or COI) and the denominator (per year) since the extent of exposure is expressed per year (number of eating occasions per year) for acute hazards and the burden of disease is generally annualized (by dividing the lifetime risk by the duration of exposure) for chronic hazards. As noted above, however, a user may specify not to compute annualized results for chronic exposures.

1.3.1 Probability of Illness: Acute Hazards

For acute hazards the probability of illness, i.e., the mean probability of illness for all contaminated food units, is:

$$P_{f,h} = E[P(\varepsilon_h \mid AD_{f,h}) \times P(\gamma_h \mid \varepsilon_h) \times P_s]$$

Equation 5

where:

- $P(\varepsilon_h \mid AD_{f,h})$ is the probability of response provided by the dose response model specified for hazard *h*, given ingestion of dose $AD_{f,h}$.
- $P(\gamma_h \mid \varepsilon_h)$ is the probability of illness given response ε_h occurs if ε_h is an endpoint other than frank illness. For example, if the dose response relationship predicts infection only,

this value takes into account that illness may only occur for a fraction of the cases of infection.

- P_s is the prevalence of contaminated units of food at the point of consumption, provided by the process model.
- *E* denotes the expectation (e.g., the mean) of the value in the brackets, as computed from the mean of the iterations within a Monte Carlo Simulation.

The calculation of dose, i.e., the distribution of doses from contaminated units, is the result of the Process Model element of FDA-iRISK. (The process model is described in *Section 2: Estimation of the Extent of Contamination: Process Models*.)

1.3.2 Probability of Illness: Chronic Hazards

For chronic hazards the probability of illness is:

$$P_{f,h} = P(\varepsilon_h \mid LADD_{f,h}) \times P(\gamma_h \mid \varepsilon_h)$$

Equation 6

where:

- $P(\varepsilon_h \mid LADD_{f,h})$ is the probability of response provided by the dose response model specified for hazard *h*, given ingestion of lifetime (or long-term) average daily dose $LADD_{f,h}$.
- $P(\gamma_h \mid \varepsilon_h)$ is the probability of illness given response ε_h occurs if ε_h is an endpoint other than frank illness.

The calculation of dose is the result of the Process Model element of FDA-iRISK.

1.4 Ranking Options

When users request ranking reports from FDA-iRISK, they have the option to specify one of four ranking options based on the risk estimates. The user can choose to rank the scenarios by:

- Health Metric (e.g., Total DALYs)
- Health Metric per eating occasion or consumer
- Total Illnesses
- Illnesses per eating occasion or consumer
- Exposure (Dose)

Note: While this provides different options for ranking, it does not change the risk estimates computed for each risk scenario or exposure only scenario.

2 Estimation of the Extent of Contamination: Process Models

The FDA-iRISK process model is responsible for generating the values for unit mass, prevalence, and concentration of hazards in distinct units at the point of consumption.

The user provides the initial values of the three variables and the values of parameters that define various process stages that may affect the mass, prevalence, and/or concentration (see Figure 1).

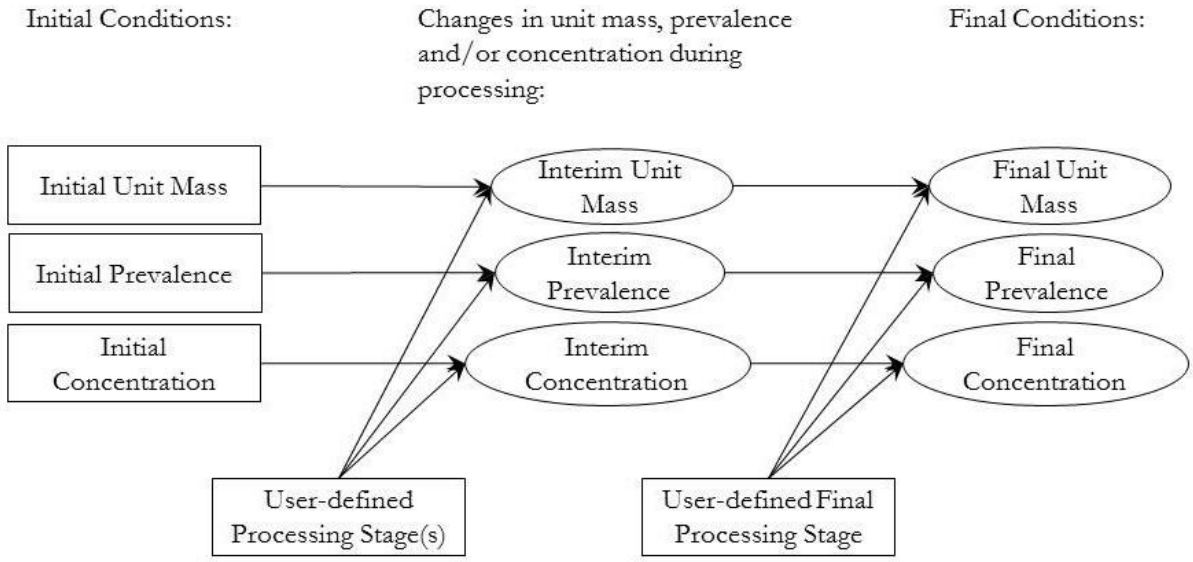


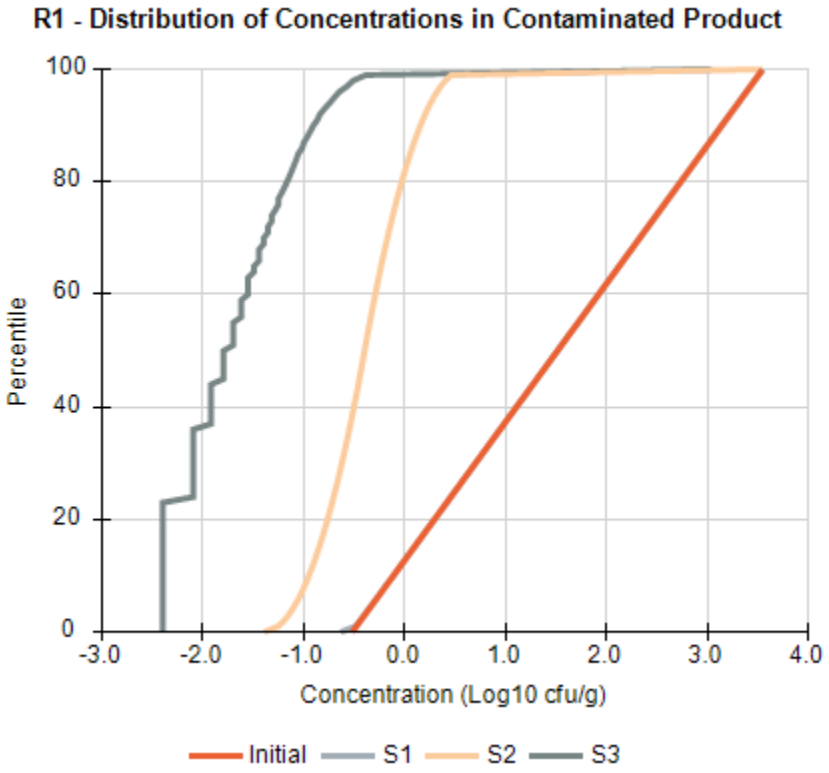
Figure 1. Mathematical structure of a process model. The user inputs initial conditions and defines sequential process stages that affect the mass, prevalence, and/or concentration of the hazard in the food. FDA-iRISK recalculates these values after every stage until the final values are obtained.

Whenever the user specifies a probability distribution for a quantity required in FDA-iRISK, the calculation from that point forward (e.g., any subsequent quantities that depend on this distributed quantity) is computed using Monte Carlo simulation (strictly speaking, FDA-iRISK uses Random Latin Hypercube Sampling). The distribution is intended to describe variability in various physical properties and processes. It is not intended to describe uncertainty which the user can assign to individual parameters separately.

Each iteration within the simulation corresponds to a distinct variant of food exposure that is the result of the calculations that include the values drawn randomly from the user-specified distributions. The FDA-iRISK process model tracks two probability distributions in parallel during the assessment of a single food-hazard combination. The first considers the variability in the level of contamination in contaminated units and is conditional on a unit being contaminated at a level >0. The second distribution is an array of corresponding weight factors (termed “prevalence” within the model, and in this description) which, on an iteration-by-iteration basis, takes account

of the likelihood that a unit is contaminated at a level >0. The weights are a combination of the initial prevalence of contamination and the impact of various process changes that require adjustment of the probability of contamination for the variant of exposure that each iteration describes. Each iteration models a distinct and separate pass of a batch of food units through the process model. All of the process types, as described in subsequent sections, act on the concentration and prevalence value of each iteration separately. In no case would the concentration and prevalence values of units from one iteration be merged with those of another iteration. If units are pooled, for example, they are pooled within each iteration and all contaminated units in the pooled batch are assumed to share the same contamination level and prevalence associated with that iteration.

In the risk estimate reports, FDA-iRISK reports the mean prevalence and the mean concentration values for final conditions after each process stage to provide an overview of the impact of each processing stage. For microbial process models, the concentration value is the arithmetic mean but is reported on the log scale. FDA-iRISK also provides charts of intermediate results for concentration distributions for single food, acute microbial scenarios where users can review how the entire distribution of concentration values in contaminated product varies from stage to stage. The following chart provides an illustrative example, where "Initial" shows the initial concentration and "S1", "S3", and "S3" represent three sequential process stages. Note: the process stage "S1" has limited impact on the concentration distribution and is largely hidden by the "Initial" concentration curve in this example.



Note: these distributions represent the concentration in *contaminated* product and should be considered in the context of mean prevalence which can vary significantly between initial conditions and each stage.

2.1 Description of Initial Contamination

The user identifies a point in the production and process change from which modeling of contamination will start and specifies the initial conditions representing that point in the production and processing chain.

The initial unit quantity of a unit of food is specified in terms of mass or volume, according to the choice made in the definition of the food. The user selects or clears a check box to register the presence or absence of contamination at this stage. If contaminated, a fixed value for the initial prevalence must be provided. Finally, for prevalence $\neq 0$, a concentration of hazard in the food must be defined either as a fixed value or as a distribution (e.g., Beta PERT, Cumulative Empirical, Normal, Triangular, or Uniform). At the initial stage, based on the user-defined prevalence, each unit of food has the same probability of being contaminated.

The numerator units for concentration are $\log_{10}$ count for microbial hazards, where the count refers to colony-forming units (cfu), plaque-forming units (pfu), or other units as specified in the hazard definition. The numerator units for chemical concentration are expressed in units of the mass of hazard.

The choice of units for the denominator (mass or volume) is specified when the food is defined.

In both cases, the denominator unit depends on the unit choice made for the food (either mass or volume). For example, a microbial concentration might be expressed as 3 $\log_{10}$ cfu/g or 5 $\log_{10}$ pfu/ml whereas a chemical concentration might be expressed as 2 ug/kg or 2 ng/l.

Additionally, in the case of microbial hazards, the concentration specified must result in at least one count (cfu, pfu, or other) in the initial unit mass specified. This is due to the definition of concentration as including only contaminated units of food (i.e., not allowing zero concentration, since this is captured by the estimate of prevalence). For example, a concentration of 1 $\log_{10}$ cfu/kg would not be permissible if the initial unit mass was specified as 1 gram, as this would result in an initial microbial load of 0.01 cfu (i.e., less than one organism). Fractions less than 1 are not allowed for estimates of concentration due to the definition of concentration as applying only to contaminated units. A concentration of 1 $\log_{10}$ cfu/g (i.e., 10 cfu/g on the non-log scale) would be acceptable if the unit mass was 0.1 grams, as it would result in an initial microbial load of 1 cfu in the initial unit mass of 0.1 gram. Note that, while users can specify initial concentration on the log scale, FDA-iRISK converts all $\log_{10}$ units to arithmetic units when performing calculations. FDA-iRISK then computes the arithmetic mean of concentration results but converts these means back to the $\log_{10}$ scale for the report. This can result in the reported mean

concentration (also referred to as computed mean) having a value different than the mean specified by the user. For example, if the user specifies an initial concentration in $\log_{10}$ cfu/g as a Normal distribution with a mean of 3 and a standard deviation of 1, the computed arithmetic mean concentration is 13,035 cfu/g. When converted back to the $\log_{10}$ scale, this results in a value of 4.12 $\log_{10}$ cfu/g.

For microbial hazards, the user may also specify a maximum population density (MPD). If the MPD is specified, the concentration of the hazard in the food is compared with the MPD at each stage of the process model and prior to consumption. If the concentration exceeds the MPD, the concentration value is set to the MPD.

These restrictions do not apply to chemical hazards.

2.1.1 Initial Prevalence and Unit Size

The prevalence value specified must be the prevalence of contaminated units for the unit size specified. For example, if the unit size is for a head of lettuce, the prevalence must be the proportion of heads of lettuce that are contaminated and not the proportion of fields or shipping crates that are contaminated. Initial prevalence must be a fixed value. However, unit size may be defined as a fixed value or a distribution.

2.1.2 Distributions for Initial Concentration and Unit Size

The following distributions are available in FDA-iRISK. Note that not all distributions are available in in all context. For example, Normal is not available for unit size as the unit size must have fixed bounds. Users need to define the parameter values for a distribution they select.

Table 2_1. Distributions for initial concentration and unit size

Distribution	Parameters
Fixed Value	Value
Beta General	Alpha, Beta, Lower bound, Upper bound
Beta PERT[a]	Minimum, Mode, Maximum
Chance	Probabilities, Values (corresponding to probabilities).
Empirical (cubic)	Probabilities [b], Values (corresponding to probabilities)
Empirical (linear)	Probabilities [b], Values (corresponding to probabilities)
Gamma	Shape, Rate

Distribution	Parameters
Laplace [c]	Location, Scale
Lognormal	Arithmetic Mean of X, Standard Deviation
Lognormal (median)	Median, Geometric Standard Deviation
Normal	Mean, Standard Deviation
Normal (Truncated)	Mean, Standard Deviation, Lower bound, Upper bound
Triangular	Minimum, Mode, Maximum
Triangular (Percentiles)	5th Percentile, Mode, 95th Percentile
Triangular (Truncated)	Minimum, Mode, Maximum, Lower bound, Upper bound
Uniform	Minimum, Maximum
Uniform (Percentiles)	5th Percentile, 95th Percentile

[a] Also known as PERT

[b] Must have values for probabilities of 0 and 1.

[c] Example of use reported in Pouillot et al., 2010

2.1.3 Linking to other Process Models

When defining the initial conditions an option is available to link to another process model. Linked Process Models are useful, for example, to define foods that are contaminated because they contain a particularly problematic ingredient (that may be found in a variety of foods), or to describe the type of branched process model that results from variations in downstream processing (e.g., frozen vs. refrigerated finished product, different styles of preparation by consumers).

An “upstream” model is built, composed of initial conditions and zero or more process stages, to represent the common conditions and process steps prior to consumer preparation, for example. Any model built subsequently can be “linked” to this, becoming one of any number of “downstream” models. The upstream model needs to be defined first, using the option “Single Set of Parameters”.

In contrast, when defining the “Downstream” model(s), the user selects “Upstream Process Model” to view the menu of models to which to link the new model. After having made the

selection and selected “Change Method”, a dropdown menu will appear that presents the selection of potential upstream models. The downstream model now takes the upstream model endpoints from the process model as the initial conditions.

Note that in order for linking to be possible, the hazard must be the same in both upstream and downstream models, and the units of measurement of the food must be the same. Linking does not allow branching within a single model.

A special case of linked process models converts a microbial pathogen to a chemical contaminant. In this case, the microbial process model constitutes the upstream model, and the chemical process model is the downstream model. The two are linked by means of a conversion rate that relates the amount of toxin or another chemical produced per unit of the microbe. Two options are available for the conversion:

1) A linear conversion (in the standard form y=mx+c) where CT= C_M * R + I where C_M = microbial concentration (e.g., cfu/g), R = Conversion Rate (e.g., mg/cfu), I is the intercept (which can be set to 0) and CT = Toxin concentration (e.g., mg/g). The user will specify the units used for the conversion rate, and unit conversions are applied in the model as appropriate.

2) A log-linear conversion (in the standard form Log(y)=m*Log(x)+c) where LogCT= $LogC_M$ * LR + LI where $LogC_M$ = microbial concentration (e.g., Log cfu/g), LR = Conversion Rate (e.g., Log mg/Log cfu), LI is the intercept (which can be set to 0) and LogCT = Toxin concentration (e.g., Log mg/g). The user specifies the units used for the conversion rate, and unit conversions are applied in the model as appropriate.

In both cases, the user can specify an optional Threshold (in log cfu/g) below which no toxin is produced.

The end result is a chemical concentration in the food. From this point forward, FDA-iRISK treats the associated scenario as a chemical hazard type. As with standard linking models, in order for linking to be possible, the units of measurement of the food must be the same.

2.2 Mathematical Description of the Process Stage Types

In this section, the following notation is used:

i *Current stage of the process model being described.*

C_i *Concentration of the hazard in contaminated food units at the end of stage i, expressed in non-log units for microbial pathogen hazards and chemical hazards.*

P_i *Prevalence (probability of contamination) of units of food contaminated with the hazard at the end of stage i.*

M_i *Mass of a unit of food at the end of stage* i.

X_{i-1} *Value of* X *at the previous stage of the process model,* $i-1$*, where* $X \in \{C, P, M\}$, for example C_{i-1}.

Notes:

- The subscripts h and f (used previously to denote hazard *h* and food *f*) are omitted in this section for clarity; however, any given process model is specific to a single food-hazard combination.
- While some parameters for microbial pathogens are specified on the $\log_{10}$ scale, values are converted as needed to compute concentration on the non-log scale.

As discussed above, FDA-iRISK conducts a Monte Carlo simulation to describe the prevalence, concentration, and other intermediate calculations from any point in the model downstream of (i.e., dependent upon) a quantity that the user has specified as a probability distribution. As such, each of the quantities expressed in this description could be computed separately in each of *N* iterations in a Monte Carlo simulation. To simplify the notation we have suppressed the subscript denoting a specific iteration.

2.2.1 Variability Distributions for Process Types

The table below lists the variability distributions that are available in FDA-iRISK for the following process types:

- Increase by Growth
- Increase by Growth Model
- Increase by Addition
- Increase by Cross Contamination (Amount)
- Increase by Cross Contamination (Concentration)
- Decrease
- Decrease by Inactivation Model
- Evaporation/Dilution
- Partitioning
- Pooling
- Redistribution (Partial)
- Set Maximum Population Density

Table 2_2. Distributions for process types

Distribution	Parameters
Fixed Value	Value
Beta	Alpha, Beta
Beta General	Alpha, Beta, Lower bound, Upper bound
Beta PERT*	Minimum, Mode, Maximum
Chance	Probabilities, Values (corresponding to probabilities).
Empirical (cubic)	Probabilities**, Values (corresponding to probabilities)
Empirical (linear)	Probabilities**, Values (corresponding to probabilities)
Gamma	Shape, Rate
Laplace	Location, Scale
LogUniform	Mean, Standard Deviation
Normal	Mean, Standard Deviation
Normal (Truncated)	Mean, Standard Deviation, Lower bound, Upper bound
Triangular	Minimum, Mode, Maximum
Triangular (Percentiles)	5th Percentile, Mode, 95th Percentile
Triangular (Truncated)	Minimum, Mode, Maximum, Lower bound, Upper bound
Uniform	Minimum, Maximum
Uniform (Percentiles)	5th Percentile, 95th Percentile

* Also known as PERT

**Must have values for probabilities of 0 and 1.

The process types: Placeholder, No Change, and Redistribution (Total), do not have variability distributions associated with them.

2.3 Process Types for Microbial Hazards

Note: If the user specifies a maximum population density (MPD) for the hazard as part of the initial contamination, the concentration of the hazard in the food is compared with the MPD at the end of each stage of the process model. If the concentration exceeds the MPD, the concentration value is set to the MPD. For more information about MPD, see *Section 2.3.16 Set Maximum Population Density (MPD)*.

The process types implemented in FDA-iRISK are similar in nature and purpose to those previously published in the literature (e.g., Nauta, 2002, 2005 & 2008; ILSI, 2010).

The process types include:

- Increase by Growth
- Increase by Growth Model
- Increase by Addition
- Increase by Cross Contamination (Amount)
- Increase by Cross Contamination (Concentration)
- Decrease
- Decrease by Inactivation Model
- Pooling
- Partitioning
- Evaporation/Dilution
- Redistribution (Partial)
- Redistribution (Total)
- Set Maximum Population Density
- Sampling (OC Curve)
- Sampling (Simple Poisson)
- Threshold Exceedance Test
- No Change
- Placeholder

2.3.1 Increase by Growth-Microbial

To describe a growth process, the user specifies:

- *G*, the multiplicative increase in the number of microorganisms, expressed in $\log_{10}$ units (e.g., 1 $\log_{10}$ denotes a 10-fold increase in the concentration of organisms). This can be specified either as a fixed value or as a variability distribution.

Case 1: When $C_{i-1} = 0$ or when $P_{i-1} = 0$, then the new concentration and prevalence following the stage are also 0.

Case 2: The concentration after growth occurs at stage *i* is:

$$C_i = C_{i-1} \times 10^G$$

Equation 7

The concentration is evaluated taking into account the MPD as described above. Prevalence is unaffected by growth and therefore, $P_i = P_{i-1}$. Mass is similarly unaffected, $M_i = M_{i-1}$.

The following example illustrates how many of the process types are applied during the Monte Carlo simulation. Assume the user defines *G* as a Uniform distribution (0,4) and the concentration at the end of the previous stage is as listed for each iteration in the table below. FDA-iRISK will draw samples from the distribution for *G* using the Random Latin Hypercube method. The following may result for the first five iterations of the Monte Carlo simulation of the growth stage:

Table 2_3. Example: application of process types during simulation

Iteration	**1**	**2**	**3**	**4**	**5**
C_{i-1} (cfu/g)	0	2	4	10	5
Sample of G	2.4	3.1	0.5	1.3	1.9
10^G	251.2	1,258.9	3.15	20.0	79.4
C_i (cfu/g)	0	2,518	12.6	200	397

2.3.2 Increase by Growth Model-Microbial

With the increase by growth process type, the user specifies the amount of growth directly. With the increase by growth *model* process type, the user selects pre-defined growth and lag predictive models (see Section 3: Predictive Models), and provides the time, temperature and other parameters required by the models. The process type then computes a growth rate and lag phase duration based by applying the time, temperature and other parameters (e.g., pH) to the selected growth and lag models.

The increase by growth is calculated from $LogIncrease = GR_T t_G$ where GR_T is the growth rate at temperature T, and t_G is the time that growth can occur, given by $t - lag_T$, where lag_T is the lag time at temperature T.

Once the increase by growth amount is calculated, growth is computed using the same method specified above for the Increase by Growth process type. If the user wishes to set the lag to 0, they must first select the no lag model for process stage.

Note that predictive models imported from ComBase are assumed to account for lag and to incorporate assumptions about the temperature experienced by the food. As such, the user will generally select "No Lag" as the option for the lag model. The user will not be prompted for a temperature distribution when using predictive models imported from ComBase unless a lag model is selected, and that lag model requires temperature as a parameter. The temperature will not be used to compare to the growth range for the hazard when using a model imported from ComBase even if specified for the lag model.

2.3.3 Increase by Addition-Microbial

The Increase by Addition process type is specified using two parameters: the amount (not concentration) of contamination added (on the $\log_{10}$ scale) and the likelihood of that addition occurring. FDA-iRISK models increase by addition at the batch level. That is, likelihood is defined as the likelihood that the amount of contamination specified will be added to each unit in a batch.

Four states may result from the application of an increase by addition process:

1) A previously contaminated unit did not experience addition.
2) A previously contaminated unit experienced addition.
3) A previously uncontaminated unit experienced addition.
4) A previously uncontaminated unit did not experience addition.

As the fourth state does not pose any health risk (no contamination), it is not considered separately. Instead, it is incorporated with the first state using prevalence (the proportion of contaminated units in a batch).

To allow for low likelihood values but still maintain an efficient simulation model, FDA-iRISK implements separate pathways to model each state and applies a weight to each pathway that is used to re-integrate the pathways when computing risk downstream in the model. Each state will have a different net concentration and prevalence result. The following table summarizes how concentration and prevalence change for each state, and the weight associated with that state.

Table 2_4. Increase-by-Addition (microbial): changes in concentration and prevalence (definition)

State (Pathway)	Concentration After Addition Process	Prevalence After Addition Process	Probability of Pathway for Any Given Eating Occasion
No addition	$C_i = C_{i-1}$	$P_i = P_{i-1}$	$(1-P_a)$
Addition, previously contaminated	$C_i = C_{i-1} + (10^A / M_{i-1})$	1	$P_{i-1} * P_a$

State (Pathway)	Concentration After Addition Process	Prevalence After Addition Process	Probability of Pathway for Any Given Eating Occasion
Addition, previously uncontaminated	10^{A} / M_{i-1}	1	$(1-P_{i-1}) * P_a$

where:

- A is the amount added per unit on the $\log_{10}$ scale. It is not a concentration value.
- P_a is the probability of addition to any unit.

For example, assuming:

- Initial Contamination: 2 $\log_{10}$ cfu/g (100 cfu/g)
- Initial Prevalence: 0.3
- Unit Mass: 50 g
- Addition Amount: 2.69 $\log_{10}$ cfu (500 cfu) per unit
- Addition Likelihood: 0.001

The resulting pathway states would be:

Table 2_5. Increase-by-Addition (microbial): changes in concentration and prevalence (example data)

State (Pathway)	Concentration After Addition Process	Prevalence After Addition Process	Probability of Pathway for Any Given Eating Occasion
No addition	2 $\log_{10}$ cfu/g (100 cfu/g)	0.3	0.999
Addition, previously contaminated	2.04 $\log_{10}$ cfu/g (110 cfu/g)	1	0.0003
Addition, previously uncontaminated	1 $\log_{10}$ cfu/g (10 cfu/g)	1	0.0007

Mass is unaffected: $M_i = M_{i-1}$.

For microbial (acute) hazards, FDA-iRISK assumes any individual might consume a serving from any given pathway in proportion to its relative frequency of occurrence. Therefore, the

probability of illness is computed downstream for each model pathway separately, and then the frequency-weighted average over all of the pathways is taken as the final probability of illness.

If pooling occurs downstream from an addition process, food units are pooled within each pathway and are not pooled across pathways. For information about how pooling is implemented with regard to other process types, such as addition, see *Section 2.3.9 Pooling-Microbial*.

2.3.4 Increase by Cross Contamination (Amount) - Microbial

This Increase by Cross Contamination (Amount) process type adds contamination to a unit using a defined pool of organisms and transfer rate.

The user is asked to input the following:

- The likelihood of sufficient contact to cause the transfer (i.e., set to 1 in the case where transfer of some degree always occurs)
- The amount of contamination (i.e. number of cells or microorganisms) in the environmental pool (e.g., Log CFU). This amount remains unchanged after transfer and no update is made following a cross-contamination event. The amount can be expressed as a distribution.
- The transfer rate from the pool, expressed as either percentage or Log10 percentage.

The amount added by this process type is computed by multiplying the number of cells or microorganisms in the environmental pool by the transfer rate. Once the amount is determined, the same logic described above for the Increase by Addition process type is applied.

2.3.5 Increase by Cross Contamination (Concentration) - Microbial

The Increase by Contamination (Concentration) process type adds contamination to a unit using a pool with a defined concentration and amount of material, and a transfer rate.

The user is asked to input the following:

- The likelihood of sufficient contact to cause the transfer (i.e., set to 1 in the case where transfer of some degree always occurs)
- The concentration (Log CFU per ml or g) in the environmental pool. This concentration remains unchanged after transfer and no update is made following a cross-contamination event. The concentration can be expressed as a distribution.
- The amount of material in the pool (ml or g). This amount remains unchanged after transfer and no update is made following a cross-contamination event. The amount can be expressed as a distribution.
- The transfer rate from the pool, expressed as either percentage or Log10 percentage.

The amount added by this process type is computed by multiplying the concentration by the amount of material in the environmental pool and by the transfer rate. Once the amount is determined, the same logic described above for the Increase by Addition process type is applied.

2.3.6 Decrease-Microbial

Assumption: The Decrease process type is not capable of the complete removal of the hazard from the system. There is always a non-zero chance of survival of the inactivation process.

Where:

- d_i is the user-specified $\log_{10}$ reduction for a microbial hazard at stage i.
- M_i is the unit mass at stage i.

Three cases are defined as follows:

Case 1: When $C_{i-1} = 0$ or when $P_{i-1} = 0$, then the new concentration and prevalence following the stage are also 0.

Case 2: When the user-specified distribution includes the possibility of some values where $d_{i,n} \leq 0$ in $\log_{10}$ units, then no decrease is applied for these specific iterations *n,* and the concentration and prevalence are unchanged from the previous values (e.g., effectively implementing a 0-log decrease, or no change).

Case 3: When the user-specified decrease is $d_{i,n} > 0$ in $\log_{10}$ units, then the new concentration is a random variable drawn from the binomial distribution (under the assumption that each organism has the same, independent probability of survival), conditioned upon there being at least one surviving organism in a contaminated unit. The fact that some units may become completely de-contaminated with respect to this hazard is addressed by adjusting the prevalence value associated with this unit.

$$C_i \sim pos\_binomial(N_{i-1}, \rho_1) / M_{i-1}$$

Equation 8

where:

- the probability of survival of an individual organism is:

$$\rho_1 = 10^{-d_i}$$

Equation 9

and:

- the microbial load prior to the decrease is (rounded to be an integer for use in the binomial calculation):

$$N_{i-1} = round(C_{i-1} \times M_{i-1})$$

Equation 10

The positive binomial function returns random samples from the binomial distribution, conditional upon the value being non-zero. For a description of the positive binomial function, see *Section 6 Positive Only Binomial and Poisson Distributions*.

To account for the probability that removal of contamination will result in individual units becoming uncontaminated (i.e., less than one cfu or pfu per unit), the prevalence after the decrease stage is given by adjusting the prevalence value from the previous stage by the probability that some contamination will remain after the decrease in this stage.

The probability of survival of one or more organisms, given an individual survival probability, ρ_1, and a starting microbial load of N_{i-1} is given by:

$$\rho_s = 1 - (1 - \rho_1)^{N_{i-1}}$$

Equation 11

Therefore, the final probability that this unit is contaminated is given by multiplying the probability that it was previously contaminated by the probability that one or more organisms will survive the *decrease* process:

$$P_i = P_{i-1} \times \rho_s$$

Equation 12

The unit mass is unaffected: $M_i = M_{i-1}$.

2.3.7 Decrease by Inactivation Model-Microbial

The log decrease is calculated from the Weibull model $L = \left(\frac{t}{D}\right)^{\alpha}$, when shape = 1 the Weibull becomes the familiar linear model. The user first defines inactivation models for the hazard (see Section 3

Using Pre-run Process Models in Multifood Risk Scenarios

For chronic exposure, the final mean concentration of a hazard in the food is used in combination with consumption models to compute a lifetime average daily dose (refer to Sections 564.3 Chronic Exposure and 4.4 Consumption Model for Chronic Exposure). A multifood risk scenario contains two or more process models to estimate the final prevalence-weighted concentration

of a specific hazard (e.g., a chemical) in each food. The process models often do not change while analysts explore other elements such as the amount consumed under different diet scenarios. To gain efficiency and allow for a larger number of foods and associated process models to be included in a multifood scenario, FDA-iRISK enables users to pre-run the process models and store the final prevalence-weighted concentration from a converged process model simulation (refer to Section 8 Evaluation of the Convergence for the Monte Carlo Simulation). If this option is used, the process model will not be re-simulated when generating a risk estimate report. Instead, the stored final prevalence-weighted concentration result will be used for all convergence batches of the exposure model, reducing the time required to estimate the risk through the simulation process.

Predictive Models). When defining this process type, the user selects one of the pre-defined inactivation models then assigns values for time, temperature and other parameters required by the model (e.g., z-value). FDA-IRISK uses these models compute a D-Value (the D-value is time for reduction of one $\log_{10}$ in the linear model or the first $\log_{10}$ in the Weibull model when shape ≠ 1), and then applies the Weibull model to compute the amount of decrease. From this point, FDA-iRISK applies the same logic as the Decrease process type described earlier.

2.3.8 Mass Change – Microbial

The Mass Change function addresses both Pooling and Partitioning process types. The function compares the previous unit size (by mass or volume) with the new unit size for each iteration, and selects pooling when the new size is larger and partitioning when the new size is smaller. In the case where the new size is the same as the previous size, no change is made. When $C_{i-1} = 0$ or when $P_{i-1} = 0$, then the new concentration and prevalence following the stage are also 0. Otherwise, the pooling or portioning functions are applied as appropriate.

Assumption: Microbial hazards are distributed randomly in each simulated unit of food and their presence in a sub-sample of the unit of food follows a Poisson distribution.

2.3.9 Pooling-Microbial

The Pooling process type addresses the possibility that the new unit size may result from the combination of both contaminated and uncontaminated units.

First, the number of portions that need to be constituted to determine the new unit mass is determined by dividing the new unit mass by the previous unit mass. This will result in *X* whole units and a fraction *f* ($0 \leq f < 1$) of one unit.

Prevalence is determined by using the previous prevalence to compute the probability that one or more of the *X* whole units is contaminated:

$$P_{unit} = 1 - (1 - P_{i-1})^X$$

Equation 13

and the probability that the fraction is contaminated is calculated by the probability that the unit from which it was drawn is contaminated, multiplied by the probability that the fraction contains one or more of the original microorganisms in the unit:

$$N_{i-1} = round(C_{i-1} \times M_{i-1}),$$

Equation 14

which are assumed to have a probability of being in the fraction equal to the fractional mass of the fractional unit:

$$P_{frac} = P_{i-1} \times (1 - (1 - f)^{N_{i-1}}).$$

Equation 15

The final new prevalence is computed by calculating the probability that none of the inputs (neither the whole units nor the fractional units) to the new mass (the "pool") are contaminated, and subtracting this value from one:

$$P_i = 1 - (1 - P_{unit}) \times \left(1 - P_{frac}\right)$$

Equation 16

The new concentration is determined by randomly sampling from the three possibilities:

- Only one or more of the whole units is contaminated, with probability $P_{unit} \times \left(1 - P_{frac}\right)$. The resulting concentration is $C_i = posPoisson\left(posBinom\left(X, P_{i-1}\right) \times N_{i-1}\right) / M_i$.
- Only the fractional unit is contaminated, with probability $P_{frac} \times (1 - P_{unit})$. The resulting contamination is $C_i = \frac{posBinomial(N_{i-1}, f)}{M_i}$.
- Both are with probability $P_{frac} \times P_{unit}$. The resulting contamination is:

$$C_i = \left(posPoisson(posBinomial(X, P_{i-1}) \times N_{i-1}) + posBinomial(N_{i-1}, f)\right)/M_i.$$

Equation 17

The mass M_i is set equal to the new unit (or "pool") size.

It should be noted that the process of pooling is not a completely random recombination of all of the simulated units within the total Monte Carlo simulation. Rather, it is assumed that the pooling occurs among units within a given iteration and addition pathway that have the same probability of contamination and level of contamination as the previous unit simulated. An alternative concept of pooling, where every simulated unit across all iterations has the potential to be included in a pool, is not applied.

The Poisson distribution is applied to simulate some variability in the actual contamination levels between the units being pooled, with the expected value being the product of the number of contaminated units and the expected number of micro-organisms in each. The positivePoisson function, in particular, is used to ensure that the number of organisms returned from the Poisson distribution is greater than zero. The pooling process stage assumes no clumping/clustering of organisms.

2.3.10 Partitioning-Microbial

As the microbial hazard is assumed to be randomly distributed in the food, the new prevalence is the probability that at least one micro-organism is present in the new, smaller unit size. The

starting number of micro-organisms available for partitioning to sub-units is $N_{i-1} = round(C_{i-1} \times M_{i-1})$.

The probability that a micro-organism is in the smaller unit size, given a previously contaminated larger unit, is equal to the fraction of the previous unit mass that the new unit mass represents. For example, if partitioning 100 liters of product into 4-liter bags, there is a 4% chance of any single organism ending up in a randomly selected new 4-liter unit. The prevalence is then adjusted by the probability that one or more organisms will end up in a random smaller unit.

$$Psmall = \left[1 - \left(1 - \left(\frac{M_i}{M_{i-1}}\right)\right)^{N_{i-1}}\right]$$

$$P_i = P_{i-1} \times Psmall$$

Equation 18

The new contamination level is determined by using the positive binomial to sample the number of micro-organisms that are in the new unit size using that probability.

The new concentration is the new randomly generated contamination count divided by the new unit size.

$$C_i = pos\_binomial\left(N_{i-1}, \frac{M_i}{M_{i-1}}\right) / M_i$$

Equation 19

The mass M_i equals the new, smaller unit size.

2.3.11 Evaporation/Dilution-Microbial

Assumption: Evaporation and dilution occur on a unit-by-unit basis and neither process adds or removes contamination from the system.

The user specifies a value (fixed or variability distribution) representing the factor change in concentration, resulting in one of two possible cases:

Case 1: When $C_{i-1} = 0$ or when $P_{i-1} = 0$, then the new concentration and prevalence following the stage are also 0.

Case 2: The new concentration is given by:

$$C_i = C_{i-1} \times \varepsilon_i$$

Equation 20

where ε_i is the user-specified concentration change due to evaporation ($\varepsilon_i > 1$) or dilution ($0 < \varepsilon_i < 1$) for a microbial hazard at stage i.

The mass is also adjusted, such that:

$$M_i = M_{i-1} / \varepsilon_i$$

Equation 21

The prevalence is unchanged $P_i = P_{i-1}$.

2.3.12 Partial Redistribution-Microbial

The user specifies a value, $\emptyset$ defining the redistribution factor (i.e., the number of units among which the contamination from one (previously contaminated) unit is spread).

Case 1: When $C_{i-1} = 0$ or when $P_{i-1} = 0$, then the new concentration and prevalence following the stage are also 0.

Case 2: If the product of the redistribution factor and the previous prevalence equals or exceeds 1, this stage becomes a total redistribution and that function is called instead (see below).

When the product of the redistribution factor and the previous prevalence is less than 1, the concentration of a microbial hazard among contaminated units following a partial redistribution step is given by:

$$C_i = \begin{cases} \dfrac{1}{M_i}, N_{i-1} \le \omega \\ \dfrac{C_{i-1}}{\omega}, N_{i-1} > \omega \end{cases}$$

Equation 22

where:

- $\emptyset$ is the user-specified redistribution factor at stage i.
- N_{i-1} is the available microbial load, defined as:

$$N_{i-1} = C_{i-1} \times M_{i-1}$$

Equation 23

The first case above refers to the case where there are not enough micro-organisms per unit to spread the contamination as widely as the user-specified value suggests.

The prevalence of contaminated units is given by:

$$P_i = \begin{cases} P_{i-1} \times N_{i-1} & N_{i-1} \le \omega \\ P_{i-1} \times \omega & N_{i-1} > \omega \end{cases}$$

Equation 24

Mass is unaffected: $M_i = M_{i-1}$.

2.3.13 Total Redistribution-Microbial

If the user specifies the Total Redistribution process type, no parameters are required to quantify this process.

Case 1: When $C_{i-1} = 0$ or when $P_{i-1} = 0$, then the new concentration and prevalence following the stage are also 0.

Case 2: The contaminated units are redistributed as widely as possible, subject to the availability of sufficient numbers of organisms. For example, if the current prevalence is 1%, and the contaminated units contain only 10 organisms, there will not be enough contamination to bring the prevalence up to 100%. The final prevalence will be 10%, with 1 organism in each unit. The new concentration following a total redistribution is given by:

$$C_i = \begin{cases} \frac{1}{M_i}, N_{i-1} \le \omega \\ \frac{C_{i-1}}{\omega}, N_{i-1} > \omega \end{cases}$$

Equation 25

where the redistribution factor is calculated as:

$$\omega = \frac{1}{P_{i-1}}$$

Equation 26

and the microbial load is:

$$N_{i-1} = C_{i-1} \times M_{i-1}$$

Equation 27

The new prevalence is given by:

$$P_i = \begin{cases} P_i \times N_{i-1} & N_{i-1} \le \omega \\ 1 & N_{i-1} > \omega \end{cases}$$

Equation 28

Mass is unaffected: $M_i = M_{i-1}$.

2.3.14 Sampling (OC Curve) - Microbial

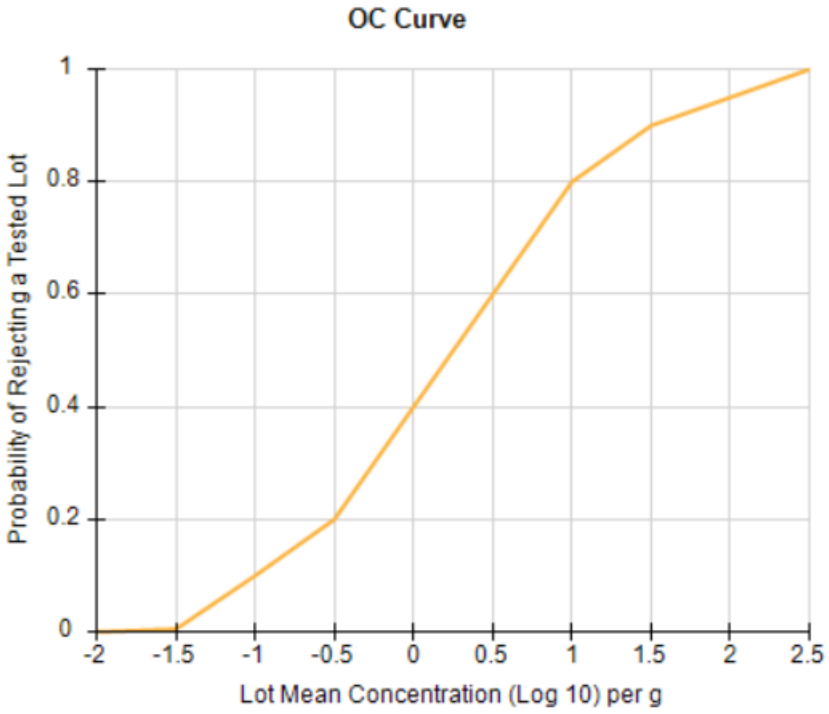
This sampling process type estimates the impact on the concentration and prevalence that results from sampling plans designed to detect microbiological contamination. There are two options for the sampling process type to choose from: Sampling (OC Curve) and Sampling (Simple Poisson). (For information about Sampling (Simple Poisson), see the next section.) For both sampling process type options, the probability of rejection (for a given concentration value) is estimated, and this probability is used to adjust the prevalence weights for the concentration values, essentially resulting in an adjustment in the concentration distribution.

For the Sampling (OC Curve) process type the user enters or loads a file containing a set of data-points that correspond to points on an OC curve with the Y-axis representing Probability of Rejection (P_reject) and the X-axis representing concentration on the log10-scale. The P_reject curve must be monotonically increasing. The probability of rejection of the sample from a food unit will be based only on the concentration in the food unit and will be linearly interpolated between the provided data points. Concentrations below and above the minimum and maximum concentrations will be assigned minimum P_reject and maximum P_reject, respectively. The series of data points provided by the user data will have increasing concentrations, and the corresponding probability can be user-specified as descending (P_accept) or increasing (P_reject).

The user has the option of uploading the data-points that correspond to the OC Curve points in the form of a “comma-separated values” (.CSV), text (.txt) or an Excel (.xls) file. For example, the following user specified data-points that correspond to the OC Curve points:

	Concentration	Preject
1	-2	0
2	-1.5	0.05
3	-1	0.1
4	-0.5	0.2
5	0	0.4
6	0.5	0.6
7	1	0.8
8	1.5	0.9
9	2	0.95
10	2.5	1

Results in the following OC Curve:



FDA-iRISK provides a file-based linkage to OC curve results generated by the FDA-OC App (Pouillot et al., 2024), which builds upon and extends beyond sampling acceptance methodologies published by the International Commission on Microbiological Specifications for Foods (ICMSF, 2020). The web application exports OC curve files for 2-class and 3-class sampling plans in a format that can be readily uploaded at a process stage. The user has the option to upload the OC curve data-points from CSV exports from the FDA-OC App (https://foodsafetyrisk.org/OC).

If the probability of contamination in the prior process stage for the given food unit is *PrevPrior*, the probability of contamination at the end of the sampling process for the same food unit is:

$$PrevPost = PrevPrior \times (1 - Ptest \times Preject)$$

Equation 29

Where *Ptest* is the proportion of food units that are tested.

Of note, when a sampling step is implemented, the generic term “food unit” is treated as a lot in the test-and-reject process (as the Simple Poisson and OC-Curve process types are implemented). The proportion of food units tested is the proportion of lots tested. In FDA-iRISK, homogeneity of concentrations within a food unit (a lot) is assumed, and thus the initial

contamination conditions include inputs for the prevalence among lots, and the distribution of the mean concentration among lots (note: there is not an option to define within lot distribution to reflect variability of concentrations within a lot).

FDA-iRISK estimates the probability of rejection (for a given concentration value) and uses this probability to adjust the prevalence weights for the concentration values, essentially resulting in an adjustment in the between-lot concentration distribution. For example, contaminated "lots" will be rejected according to the performance of the sampling plan OC curve (and the most contaminated lots are more likely to be rejected). The contaminated lots rejected are "replaced" by a random lot (which may or may not be contaminated, and may or may not be tested), resulting in a decrease in the prevalence-weighted average concentration among the "surviving" lots and an overall reduction in the number of illnesses after a sampling step is implemented, given that other model inputs are not changed.

2.3.15 Sampling (Simple Poisson) - Microbial

This sampling process type estimates the impact on the concentration and prevalence that results from sampling plans designed to detect microbiological contamination. Similar to the option for Sampling (OC Curve) described in the previous section, the probability of rejection (for a given concentration value) estimated, and this probability is used to adjust the prevalence weights for the concentration values, essentially resulting in an adjustment in the concentration distribution.

The Sampling (Simple Poisson) process type employs a single Poisson sample of fixed mass/volume.

First, the mass (m)of the sample Poisson sample is determined by:

$$m = n \times s$$

Equation 30

where:

- n is the number of samples.
- s is the mass or volume of each sample.

The mass of the Poisson sample is intended to be the final analytical sampling size (i.e., the unit of mass or volume that is subject to enrichment such as 1 g, 10g, or 25g), rather than what may be a larger physical sample taken from the food product (e.g., 100 gram samples taken, then mixed, with 25g of mixed sample subject to enrichment). The sample size must be smaller than the current food unit size.

The probability that the sample will be positive is calculated using the simple Poisson function:

$$P(+\mid test) = 1 - e^{-C.m.pdetect}$$

Equation 31

where:

- C is the concentration in the food unit from the previous process stage
- m is the mass or volume of the total combined sample
- *pdetect* is the probability the testing method will detect a single organism given it is present in the sample.

The probability of rejecting the sample is then:

$$P(reject) = P(+\mid test) \times Ptest$$

Equation 32

where:

- *Ptest* is the proportion of food units that are tested.

By extension:

$$P(accept) = 1 - P(reject) = 1 - Ptest \times (1 - e^{-C.m.pdeted})$$

Equation 33

If the probability of contamination in the prior process stage for the given food unit is *PrevPrior*, the probability of contamination at the end of the sampling process for the same food unit is:

$$PrevPost = PrevPrior \times Paccept = PrevPrior \times (1 - Ptest \times (1 - e^{-10\hat{}C.m.pDetect}))$$

Equation 34

Note: In the Simple Poisson Sampling process type, there is no assumed within-lot standard deviation. The lot is assumed to be well-mixed with respect to contamination. If the user seeks to actively include within-lot variability in concentration, they could use another tool that considers this (such as the FAO/WHO tool at www.fstools.org, or the FDA-OC App at https://foodsafetyrisk.org/OC) and transfer the resulting OC curve using the user-specified Sampling (OC Curve) process type. Alternatively, the user could start with the Sampling (Simple Poisson) process type in order to quickly explore the potential impact of sampling, and can explore the importance of considering variations on within-lot variability using the FAO/WHO tool or the FDA-OC App.

2.3.16 Set Maximum Population Density (MPD) - Microbial

For microbial hazards, the user may specify the MPD as part of the initial conditions. If the MPD is specified, the concentration of the hazard in the food is compared with the MPD at each stage of the process model and prior to consumption. If the concentration exceeds the MPD, the concentration value is set to the MPD.

The Set Maximum Population Density process type allows the user to set a new MPD at a designated point in the process model. This change in MPD may correspond to the introduction of growth inhibitors, evaporation, or other material changes to the food matrix. The value specified for a stage of this process type will be used as the MPD value from that point forward in the process model.

For example, a value of 9 log cfu/g might be specified as the initial MPD value and then changed to a value of 7 log cfu/g at a later stage in the process model using this process type.

2.3.17 Threshold Exceedance Test

This process type is intended for reporting purposes only. It does not modify the concentration, prevalence, or unit size of any stage in the process model. For this process type, the user specifies a fixed concentration value (e.g., 5 $\log_{10}$ cfu/g) at a specified process stage. FDA-iRISK will compare the concentration for each iteration from the previous process stage against this threshold value. If the concentration exceeds the threshold, the stage returns a value of 1. If the concentration is less than or equal to the threshold, the stage returns a value of 0. FDA-iRISK will then compute the mean of this value to determine the proportion of contaminated food units that exceed the threshold. It will also use the prevalence values from the previous stage to determine the proportion of all food units (contaminated and uncontaminated) that exceed this threshold. Both values are included in the resulting Risk Estimates and Ranking report.

2.3.18 No Change-Microbial

The No Change process type is designed for situations when the user wants, for the sake of completeness and transparency, to include processing steps that have no effect on unit mass, hazard concentration, or prevalence.

2.3.19 Placeholder-Microbial

The Placeholder process type is included for convenience as a temporary designation, while the process model is being built but before the data necessary to populate it have been collected. This process type is the only type that can later be changed to another type. It is distinguished from the No Change process type in that it indicates that the effect of unit mass, hazard concentration, or prevalence has not yet been determined, and therefore the model should be considered incomplete.

2.4 Process Types for Chemical Hazards

Assumption: When present, and when added, chemical hazards are uniformly distributed throughout a given unit of food. Different units of food, corresponding to different iterations within a Monte Carlo simulation, may have different concentrations, but each unit is assumed to be very well-mixed with respect to the chemical hazard.

The process types include:

- Increase by Addition
- Decrease
- Mass Change
- Pooling
- Partitioning
- Evaporation/Dilution
- Redistribution (Partial)
- Redistribution (Total)
- Set Maximum Population Density
- Sampling (OC Curve)
- Threshold Exceedance Test
- No Change
- Placeholder

2.4.1 Increase by Addition-Chemical

The Increase by Addition process type is specified using two parameters: the amount (not concentration) of contamination added and the likelihood of that addition occurring. FDA-iRISK models increase by addition at the batch level. That is, likelihood is defined as the likelihood that the amount of contamination specified will be added to each unit in a batch.

Four states may result from the application of an increase by addition process:

i) A previously contaminated unit did not experience addition.
ii) A previously contaminated unit experienced addition.
iii) A previously uncontaminated unit experienced addition.
iv) A previously uncontaminated unit did not experience addition.

As the fourth state does not pose any health risk (no contamination), it is not considered separately. Instead, it is incorporated with the first state using prevalence (the proportion of contaminated units in a batch).

To allow for low likelihood values but still maintain an efficient simulation model, FDA-iRISK implements separate pathways to model each state and applies a weight to each pathway that

is used to re-integrate the pathways when computing risk downstream in the model. Each state will have a different net concentration and prevalence result. The following table summarizes how concentration and prevalence change for each state, and the weight associated with that state:

Table 2_6. Increase-by-Addition (chemical): changes in concentration and prevalence (definition)

State (Pathway)	Concentration After Addition Process	Prevalence After Addition Process	Probability of Pathway for Any Given Consumer
No addition	$C_i = C_{i-1}$	$P_i = P_{i-1}$	$(1-P_a)$
Addition, previously contaminated	$C_i = C_{i-1} + A / M_{i-1}$)	1	$P_{i-1} * P_a$
Addition, previously uncontaminated	A / M_{i-1}	1	$(1-P_{i-1}) * P_a$

where:

- A is the amount added per unit. It is not a concentration value.
- P_a is the probability of addition to any unit.

For example, assuming:

- Initial Contamination: 2 ng/g
- Initial Prevalence: 0.3
- Unit Mass: 50 g
- Addition Amount: 5 ng per unit
- Addition Likelihood: 0.001

The resulting pathway states would be:

Table 2_7. Increase-by-Addition (chemical): changes in concentration and prevalence (example data)

State (Pathway)	Concentration After Addition Process	Prevalence After Addition Process	Probability of Pathway for Any Given Consumer
No addition	2 ng/g	0.3	0.999

State (Pathway)	Concentration After Addition Process	Prevalence After Addition Process	Probability of Pathway for Any Given Consumer
Addition, previously contaminated	2.1 ng/g	1	0.0003
Addition, previously uncontaminated	0.1 ng/g	1	0.0007

Mass is unaffected: $M_i = M_{i-1}$.

For acute chemical hazards, FDA-iRISK assumes any individual might consume a serving from any given pathway in proportion to its relative frequency of occurrence. Therefore, the probability of illness is eventually computed downstream for each model pathway separately, and then the frequency-weighted average over all of the pathways is taken as the final probability of illness.

For chronic chemical hazards, FDA-iRISK assumes each pathway will contribute to an individual's daily average consumption in proportion to its relative frequency of occurrence. Therefore, the frequency-weighted average of concentrations over all of the pathways is taken to compute the final mean concentration used when computing lifetime daily average doses.

If pooling occurs downstream from an addition process, food units are pooled within each pathway and are not pooled across pathways. For information about how pooling is implemented with regard to other process types, such as addition, see *Section 2.4.4 Pooling-Chemical*.

2.4.2 Decrease-Chemical

Assumption: The Decrease process type is not capable of the complete removal of the hazard from the system and can be described by a proportional reduction This is in line with available data on common reduction processes for chemicals in food (see for example Kaushik and Naik, 2009 for the chemical decrease impact of common food processes).

A proportional reduction is applied to the previous concentration, specifically:

$$C_i = C_{i-1} \times (1 - d_i)$$

Equation 35

where d_i is the concentration change (expressed as a fraction of the chemical removed) for a chemical hazard at stage i. For example, if the user specifies a fractional removal of 0.1, the new concentration will be 90% of the previous concentration. In this process type, the mechanism that leads to chemical decrease is unspecified; only the magnitude of the decrease is described by the process step.

The prevalence of contaminated units remains the same; therefore: $P_i = P_{i-1}$.

Mass is unaffected: $M_i = M_{i-1}$.

2.4.3 Mass Change – Chemical

The Mass Change function incorporates both pooling and partitioning elements. The function compares the previous unit size (by mass or volume) with the new unit size and selects pooling when the new size is larger and partitioning when the new size is smaller. In the case where the new size is the same as the previous size, no change is made.

Assumption: Chemical hazards are distributed uniformly throughout a given unit of food, but units of food can have different levels of contamination.

2.4.4 Pooling-Chemical

The Pooling process type addresses the possibility that the new unit size may result from the combination of both contaminated and uncontaminated units.

First, the number of portions is determined by dividing the new unit size by the previous unit size. This will result in *X* whole units and a fraction *f* ($0 \leq f < 1$) of unit.

Prevalence is determined by using the previous prevalence to compute the probability that one or more of the *X* whole units is contaminated:

$$P_{unit} = 1 - (1 - P_{i-1})^X$$

Equation 35

and the probability that the fraction is contaminated:

$$P_{frac} = P_{i-1}$$

Equation 37

The final new prevalence is computed by:

$$P_i = 1 - (1 - P_{unit}) \times \left(1 - P_{frac}\right)$$

Equation 38

The new concentration is determined by randomly sampling from the three possibilities:

- Only one or more of the whole units is contaminated, with probability $P_{unit} \times \left(1 - P_{frac}\right)$. The resulting concentration is $C_i = pos\_binomial(X, P_{i-1}) \times C_{i-1} \times M_{i-1}/M_i$.
- Only the fractional unit is contaminated, with probability $P_{frac} \times (1 - P_{unit})$. The resulting contamination is $C_i = f \times C_{i-1} \times M_{i-1}/M_i$.

- Both are with probability $P_{frac} \times P_{unit}$. The resulting probability is $C_i = (pos\_binomial(X, P_{i-1}) + f) \times C_{i-1} \times M_{i-1}/M_i$.

The mass M_i equals the new unit size.

It should be noted that the process of pooling is not a completely random recombination of all of the simulated units within the total Monte Carlo simulation. Rather, it is assumed that the pooling occurs among units within a given iteration and addition pathway that have the same probability of contamination and level of contamination. An alternative concept of pooling, where every simulated unit across all iterations has the potential to be included in a pool, is not applied.

2.4.5 Partitioning-Chemical

Chemical hazards are assumed to be uniformly distributed in the food, therefore a portion of the initial unit size will have the same concentration and prevalence as the previous unit: $C_i = C_{i-1}$ and $P_i = P_{i-1}$ following partitioning at stage i. The new mass is as specified by the user.

2.4.6 Evaporation/Dilution-Chemical

Assumption: Evaporation and dilution occur on a unit-by-unit basis and neither process adds or removes contamination from the system. This is in line with simple dilution models (see van Leeuwen & Vermeire, 2007).

A proportional reduction is applied to the previous concentration, specifically:

$$C_i = C_{i-1} \times \varepsilon_i$$

Equation 39

where ε_i is the user-specified concentration change due to evaporation ($\varepsilon_i > 1$) or dilution ($0 < \varepsilon_i < 1$) for a chemical hazard at stage i.

The prevalence of contaminated units remains the same, therefore $P_i = P_{i-1}$.

The mass adjustment is applied as:

$$M_i = M_{i-1}/\varepsilon_i$$

Equation 40

2.4.7 Partial Redistribution-Chemical

The user specifies a value, $\emptyset$ defining the redistribution factor (i.e., the number of units among which the contamination from one unit is spread). If the product of the redistribution factor and

the previous prevalence equals or exceeds 1, this stage becomes a total redistribution and that function is called instead (see below).

The prevalence of contaminated units for a chemical hazard following a partial redistribution step is:

$$P_i = P_{i-1} \times \omega$$

Equation 41

The concentration of a chemical hazard at stage i following partial redistribution is given by:

$$C_i = \frac{C_{i-1}}{\omega}$$

Equation 42

2.4.8 Total Redistribution-Chemical

Assumption: Total cross contamination results in a prevalence of 1.

The concentration of a chemical hazard at stage i following total redistribution is given by:

$$C_i = C_{i-1} \times P_{i-1}$$

Equation 43

The prevalence following total redistribution is, by definition, $P_i = 1$. For example, if the prior stage's prevalence is 10%, the final concentration will be 10% of the previous concentration, but the prevalence will be 100%.

2.4.9 Sampling (OC Curve) - Chemical

This sampling process type estimates the impact on the concentration and prevalence that results from sampling plans designed to detect chemical contamination.

For the Sampling (OC Curve) process type, the user enters or loads a file containing a set of data-points that correspond to points on an OC curve with the Y-axis representing Probability of Rejection (P_reject) and the X-axis representing concentration units that the user selects.

The P_reject curve must be monotonically increasing. The probability of rejection of the sample from a food unit will be based only on the concentration in the food unit, and will be linearly interpolated between the provided data points and will use the minimum and maximum probability specified for concentration values outside the specified range, as appropriate. The series of data points provided by the user data will have increasing concentrations, and the

corresponding probability can be user-specified as descending (P_accept) or increasing (P_reject).

The user has the option of uploading the data-points that correspond to the OC Curve points in the form of a "comma-separated values" or .CSV file format.

If the probability of contamination in the prior process stage for the given food unit is *PrevPrior*, the probability of contamination at the end of the sampling process for the same food unit is:

$$PrevPost = PrevPrior \times (1 - Ptest \times Preject)$$

Equation 44

Where *Ptest* is the proportion of food units that are tested.

2.4.10 Threshold Exceedance Test

This process type is intended for reporting purposes only. It does not modify the concentration, prevalence, or unit size of any stage in a process model. For this process type, the user specifies a fixed concentration value (e.g., 5 ng/g) at a specified process stage. FDA-iRISK will compare the concentration for each iteration from the previous process stage against this threshold value. If the concentration exceeds the threshold, the stage returns a value of 1. If the concentration is less than or equal to the threshold, the stage returns a value of 0. FDA-iRISK will then compute the mean of this value to determine the proportion of contaminated food units that exceed the threshold. It will also use the prevalence values from the previous stage to determine the proportion of all food units (contaminated and uncontaminated) that exceed this threshold. Both values are included in the resulting Risk Estimates and Ranking report.

2.4.11 No Change-Chemical

The No Change process type is designed for situations when the user wants, for the sake of completeness and transparency, to include processing steps that have no effect on unit mass, hazard concentration, or prevalence.

2.4.12 Placeholder-Chemical

The Placeholder process type is included for convenience as a temporary designation while the process model is being built but before the data necessary to populate it have been collected. The Placeholder process type is the only type that can later be changed to another type.

2.5 Using Pre-run Process Models in Multifood Risk Scenarios

For chronic exposure, the final mean concentration of a hazard in the food is used in combination with consumption models to compute a lifetime average daily dose (refer to Sections 564.3

Chronic Exposure and 4.4 Consumption Model for Chronic Exposure). A multifood risk scenario contains two or more process models to estimate the final prevalence-weighted concentration of a specific hazard (e.g., a chemical) in each food. The process models often do not change while analysts explore other elements such as the amount consumed under different diet scenarios. To gain efficiency and allow for a larger number of foods and associated process models to be included in a multifood scenario, FDA-iRISK enables users to pre-run the process models and store the final prevalence-weighted concentration from a converged process model simulation (refer to Section 8 Evaluation of the Convergence for the Monte Carlo Simulation). If this option is used, the process model will not be re-simulated when generating a risk estimate report. Instead, the stored final prevalence-weighted concentration result will be used for all convergence batches of the exposure model, reducing the time required to estimate the risk through the simulation process.

3 Predictive Models

The Increase by Growth Model and Decrease by Inactivation Model process types required predefined predictive models to describe the growth / inactivation / lag response of a microorganism to environmental conditions. The user has the option to add one or more of predictive models for each microbial hazard in FDA-iRISK. A predictive model (for growth or inactivation) may be re-used in multiple process models. You select the predefined model for the specified microbial hazard when adding process stages to the process model.

3.1 Increase by Growth Model

With the exception of the Concentration vs. Time model (see table below), the increase by growth is calculated from $LogIncrease = GR_T t_G$ where GR_T is the growth rate at temperature T, and t_G is the time that growth can occur, given by $t - lag_T$, where lag_T is the lag time at temperature T.

For both growth rate and lag time, options are provided to either enter values directly in a primary model, or to estimate the growth rate and lag from a secondary model provided. The details of the primary and secondary models available in FDA-iRISK are shown in the tables below. Note that the predictive models available in FDA-iRISK assume constant environmental conditions for any single given iteration. While variability can be represented across different iterations, e.g., by using a temperature distribution, at each specific iteration, FDA-iRISK assumes constant temperature for the entire time length.

Growth models may be obtained by using a ComBase import option. FDA-iRISK provides linkage to data and results from ComBase, a web resource for predictive food microbiology (https://www.combase.cc). FDA-iRISK supports uploading Excel export files containing fit primary models or concentration vs. time models from individual ComBase records and the online DMFit; and uploading Excel export files containing the predicted growth from the ComBase Broth Models and the Perfringens Predictor. If a primary model form is in FDA-iRISK (e.g., linear), it will be created. Otherwise, a concentration vs. time model will be created. Table 3_1 shows equations for the corresponding model used for the ComBase import once it is created. As ComBase models typically account for the lag phase, the user will generally select the No Lag model option when creating the associated process stage.

Note with regard to minimum and maximum temperatures

T_{min} represents the theoretical or notional minimum temperature and is defined as “Conceptual temperature of no metabolic significance” (Ratkowsky et al., 1982) and is the temperature below which the rate of growth is zero or lag time is infinite. T_{max} represents the theoretical or notional maximum temperature and is the temperature above which the rate of growth is zero or lag time is infinite.

When defining a predictive model in FDA-iRISK, the Minimum Growth Temperature and the Maximum Growth Temperature are also required as inputs, in addition to T_{min} and T_{max}. The Minimum Growth Temperature represents the experimental minimum temperature observed. The Maximum Growth Temperature represents the experimental maximum temperature observed. Note that there may or may not be a difference in the temperature value between the experimentally observed versus the theoretical minimum or the theoretical maximum.

In FDA-iRISK, it is necessary to define the experimental and the theoretical temperature separately because the theoretical temperature is the parameter required for some of the predictive microbiology models selected. FDA-iRISK will not compute growth for temperatures outside the range defined by the Minimum Growth Temperature and the Maximum Growth Temperature. Users should specify Minimum and Maximum Growth Temperature values such that they are within the range of T_{min} and T_{max} if the predictive model requires those parameters, otherwise the growth model may generate unexpected and invalid growth rates. That is, if T_{min} is 5.2˚ C, then the user should specify the Minimum Growth temperature as >= 5.2˚ C.

Note with regard to importing growth models from ComBase export files

When available (e.g., for the ComBase broth growth models), FDA-iRISK will assign the model's T_{min} and T_{max} values to the Minimum Growth Temperature and the Maximum Growth Temperature parameters for the model on import by default. The user may override these values with alternative values, but the same recommendation applies as the imported models are only valid for that range. Note that, by default, it is assumed that the models imported use ˚C for temperature and are on the $\log_{10}$ scale for concentration. In the associated process stage where the model is applied, users should set lag to "No Lag" as the lag model for imported models which already have the lag phase incorporated in the ComBase prediction. Users should be aware when using imported ComBase growth models in a process stage, they will not be prompted for a temperature distribution. The imported model (e.g., a Baranyi and Roberts model based on a ComBase record; a Concentration vs. Time model based on a ComBase record, ComBase broth model, or ComBase food model) explicitly specifies the amount of growth for a specified temperature and other environmental conditions that are imported as Notes in FDA-iRISK; furthermore, on import the user is required to provide the Minimum Growth Temperature and the Maximum Growth as inputs (or verify such inputs if provided in the ComBase data or model). When the ComBase model is used in a process step to predict growth, FDA-iRISK does not require a temperature input when the process step is defined and will not consider temperature when computing the amount of growth, but will consider only time (i.e., the predicted amount of growth is based on the time input given the temperature and other conditions documented in the Notes for the ComBase model). For comparison, when the user defines a predictive model not by using ComBase import but by using a primary model such as that described in Table 3_1 Equation pm1, the Minimum Growth Temperature and the Maximum Growth are required as inputs along with the growth rate input; when this predictive model is used at a process step,

FDA-iRISK requires both time and temperature inputs, where the time is used to calculate the amount of growth based on the growth rate while the temperature input is used to determine if growth occurs. That is, FDA-iRISK will assume no growth if the temperature is outside of the Minimum and Maximum Growth Temperature range.

Note with regard to sensitivity analysis with predictive models

When using the sensitivity analysis feature of FDA-iRISK to modify parameters for process stages that include growth predictive models, the user should exercise caution when providing different values for temperature as the supplied temperature must be within the range of the Minimum Growth Temperature and the Maximum Growth Temperature for growth to occur. Furthermore, for primary growth model, providing alternative temperature would not result in a difference in prediction because, here, the provided temperature will only be used to test against the Minimum and Maximum temperatures to determine if growth occurs.

Table 3_1. Growth rate: primary and secondary models

Model	Parameters	Equation for Log Cycles*	References
PRIMARY MODEL			
Simple Growth	GR_T – growth rate at temperature T t_G – time that growth can occur lag_T – lag time at temperature T	$$Log\ Increase = GR_T t_G$$ Where: $$t_G = t - lag_T$$ **Equation pm1**	For example, Buchanan et al., 1997
Concentration vs. Time	This model is defined by entering or importing a time and concentration curve starting at time 0. When FDA-iRISK determines the time duration for the process stage that uses the concentration vs. time model, it uses linear interpolation on the defined curve to determine the concentration at that time. It then subtracts the initial concentration from the final concentration to determine the log increase.	$$Log\ Increase = C_{t_G} - C_{t_0}$$ Where: $$t_G = t - lag_T$$ C = Concentration **Equation pm2**	

Baranyi and Roberts	The user is expected to use the online or downloadable version of DMFit available as part of the ComBase tool. DMFit implements the "Baranyi model" as described in Baranyi and Roberts, 1994 but as re-parameterized in Baranyi 2000. Base – the base of the logarithm used to measure concentrations and rate parameters, default is $\log_{10}$, second option is $\log_e$. y_0 – initial concentration in $\log_{base}$ cfu/g from DMFIt y_{max} – final concentration in same units as y_0 μ_{max} – maximum growth rate in logbase cfu/g/hr lag_T – duration of lag phase as fit by DMFit in hr. y_{prev} – the concentration in the previous process stage in the process model.	$$y(t) = y_{prev} + \mu_{max}A(t) - \frac{y_{max} - y_{prev}}{M} Ln\left(1 - e^{-M} + e^{-M\frac{y_{max} - y_{prev} - \mu_{max}A(t)}{y_{max} - y_{prev}}}\right)$$ Where: $$A(t) = t - lag_T + \frac{Ln\left(1 - e^{-\nu t} + e^{-\nu(t - lag_T)}\right)}{\nu}$$ $$M = 10$$ $$\nu = \frac{9 \cdot \mu_{max}}{(y_{max} - y_0)}$$ **Equation pm3**	Baranyi and Roberts, 2000. Reparameterization of Baranyi and Roberts, 1994. Equation for ν from Baranyi, 2020.

Model	Parameters	Equation for Log Cycles*	References
	t – the duration of the growth stage for this process stage in the process model Note that when using DMFit, ensure that the model fit by the tool is the Baranyi and Roberts Model (complete)		
SECONDARY MODELS			
Gamma Square Root (Temperature Only)	μ_{opt}– Optimal growth rate $T_{\min}$ – Notional min temp T – Temp T_{opt} – Optimal temp	$$\mu = \gamma(T)\mu_{opt}$$ Where $$\gamma(T) = \left(\frac{T - T_{min}}{T_{opt} - T_{min}}\right)^2$$ **Equation pm4**	Variation of Gamma Parameterization of Square Root by Zwietering et al., 1996

Model	Parameters	Equation for Log Cycles*	References
Gamma Parameterization of Square Root	T – Temp $T_{\min}$ – Notional min temp T_{opt} – Optimal temp for growth a_w – Water activity $a_{w,min}$ – Min water activity for growth $a_{w,opt}$ – Optimal water activity for growth pH – pH pH_{min} – Min pH for growth pH_{opt} – Optimal pH for growth	$$\mu = \gamma(T)\gamma(pH)\gamma(a_w)\mu_{opt}$$ Where $$\gamma(T) = \left(\frac{T - T_{min}}{T_{opt} - T_{min}}\right)^2$$ $$\gamma(pH) = \frac{(pH - pH_{min})(2 \cdot pH_{opt} - pH_{min} - pH)}{(pH_{opt} - pH_{min})^2}$$ $$\gamma(a_w) = \frac{a_w - a_{w,min}}{a_{w,opt} - a_{w,min}}$$ **Equation pm5**	Zwietering et al., 1996
Polynomial Response Surface	Includes options to specify temp, pH, NaCl, NaNO2, and associated co-efficient	Standard polynomial response surface – Users can specify which response parameters to include. Setting a co-efficient to zero essentially removes a response variable from the equation.	For example, Buchanan et al., 1993

Model	Parameters	Equation for Log Cycles*	References
Square Root for biokinetic	b – Constant T_{min} – Notional min temp T – Temp T_{max} – Notional max temp c – Constant	$\mu = \left(b(T - T_{min}).\left\{1 - e^{c(T - T_{max})}\right\}\right)^2$ **Equation pm6**	McMeekin et al., 1993a
Square Root	b – Constant T_{min} – Notional min temp T – Temp	$\mu = \left(b(T - T_{min})\right)^2$ **Equation pm7**	McMeekin et al., 1993b
Square Root with a_W	b – Constant $a_{w_{min}}$ – Min water activity for growth a_w – Water activity T_{min} – Notional min temp T – Temp	$\mu = \left(b\sqrt{(a_w - a_{w_{min}})}(T - T_{min})\right)^2$ **Equation pm8**	McMeekin et al., 1993c
Square Root with pH	b – Constant pH_{min} – Min pH for growth pH – pH T_{min} – Notional min temp T – Temp	$\mu = \left(b\sqrt{(pH - pH_{min})}(T - T_{min})\right)^2$ **Equation pm9**	McMeekin et al., 1993d

Model	Parameters	Equation for Log Cycles*	References
Square Root Modified	b – Constant T_{min} – Notional min temp T – Temp T_{max} – Notional max temp c – Constant	$\mu = [b(T - T_{min})]^2\{1 - exp[c(T - T_{max})]\}$ **Equation pm10**	Zwietering et al., 1991 (eqn. 4)

**Models and parameters can be in either Log_e or log_{10}. The conversion to log_{10} (specifically dividing by ln(10)) will be applied if the user specifies the model was fit using log_e.*

Table 3_2. Lag time models

Model	Parameters	Equation for Hours*	References
Hyperbola	*P* – Decrease in lag time when temperature increases *q* – Temperature where lag is infinite (for example < Tmin) *T* – Temp	$\log(Lag) = \frac{P}{T - q}$ **Equation pm11**	McMeekin et al, 1993e
Polynomial Response Surface	Includes options to specify temp, pH, NaCl, NaNO2, and associated co-efficient in a standard polynomial response surface.	Standard polynomial response surface. Setting a co-efficient to zero essentially removes a response variable from the equation.	For example, Buchanan et al., 1993

Model	Parameters	Equation for Hours*	References
Relative Lag	k – constant G – generation time	$lag = Gk$ **Equation pm12**	Ross & McMeekin, 2003
Square Root	b – Constant $T_{\min}$ – Notional min temp T – Temp	$\text{Lag} = \frac{1}{[\text{b}(\text{T} - T_{\min})]^2}$ **Equation pm13**	Zwietering et al., 1991 (inverse of eqn. 2)
Square Root Extended	b – Constant $T_{\min}$ – Notional min temp T – Temp $T_{\max}$ – Notional max temp c – Constant	$Lag = \left(b(T - T_{min}).\{1 - e^{c(T-T_{max})}\}\right)^{-2}$ **Equation pm14**	Zwietering et al., 1991 (eqn. 10)

***** *Log(Lag) will be converted to Lag as appropriate depending on user specification of $\log_e$ or $\log_{10}$*

3.2 Decrease by Inactivation Model

With the exception of the Concentration vs. Time model (see table below), the log decrease is calculated from the Weibull model $L = \left(\frac{t}{D}\right)^{\alpha}$, when shape = 1 the Weibull becomes the familiar linear model. There are 3 options presented to the user to incorporate the D value:

- Direct user input of the D value (as fixed value or distribution)
- Calculation from the linear model LogD = mT+b
- Calculation from log-linear model with user specified Z value, Dref and Tref

Inactivation models may be obtained by using a ComBase import option. FDA-iRISK supports uploading Excel export files containing fit primary models or concentration vs. time models from individual ComBase records and the online DMFit; and uploading Excel export files containing the predicted thermal inactivation and non-thermal survival from the ComBase Broth Models. If a primary model form is in FDA-iRISK (e.g., linear), it will be created. Otherwise, a concentration vs. time model will be created.

1147 **Table 3_3. Inactivation: primary and secondary model**

Model	Parameters	Equation	References
PRIMARY MODEL			
Log Reduction	t – time D – D-value α – shape parameter	$LR = \left(\frac{t}{D}\right)^{\alpha}$ **Equation pm15**	Peleg and Cole, 1998 Van Boekel, 2002
Concentration vs. Time	This model is defined by entering or importing a time/concentration curve starting at time 0. When FDA-iRISK determines the time duration for the process stage that uses the concentration vs. time model, it uses linear interpolation on the defined curve to determine the concentration at that time. It then subtracts that value from the initial concentration to determine the log reduction.	$LR = C_t - C_{t_0}$ C = Concentration **Equation pm16**	
SECONDARY MODELS			
Specification of D from Linear	T – temp m, b – constants	$D = 10^{mT+b}$ **Equation pm17**	
Specification of D from Z-valued	D_{ref} – D value at reference temp T_{ref} – reference temp z – z value T – temp	$D = 10\text{^}(log_{10}D_{ref} - \left(\frac{T - T_{ref}}{z}\right))$ **Equation pm18**	Bigelow, 1921

1148

4 Estimation of the Extent of Consumption: Consumption Models

4.1 Acute Exposure

In FDA-iRISK, acute exposure to a hazard refers to exposure during a single eating occasion, after which illness can ensue. The dose is calculated on a per-eating-occasion basis, so that the amount of food consumed during a particular eating occasion (i.e., a "serving"), along with the concentration of hazard in the food on that eating occasion, determines the applied dose. Each eating occasion is considered an independent opportunity to become ill.

FDA-iRISK uses this structure for risks due to microbial pathogens, and risks due to acute exposures to chemical hazards.

The concentration (and prevalence) of a hazard in the food at consumption is calculated in the process model using inputs from the user. The amount of food consumed is based on user-inputs comprising the consumption model (see Figure 2). In cases where the serving size differs from the final unit mass output of the processing stages, the mass change function (pooling or partitioning, according to the relative size of the unit and the serving) is used to determine concentration and prevalence values in servings. Monte Carlo simulation is employed to combine these inputs in a stochastic manner to capture variability in hazard concentration and in amount of food consumed. To increase the efficiency of the simulation, the dose response model uses doses from contaminated servings of food only, and provides an estimate for the risk of illness per contaminated serving. The prevalence is then incorporated to determine the risk of illness for any serving (Figure 2) for each iteration.

The mean risk of illness per serving across all iterations, is then multiplied by the user-specified annual number of servings consumed (again from the consumption model) to predict the number of cases per year. Each case is assigned a value for burden (in DALYs, COI, or QALY). In this way, the overall burden for the exposure is calculated. (This value for annual burden is the basis of the rankings.)

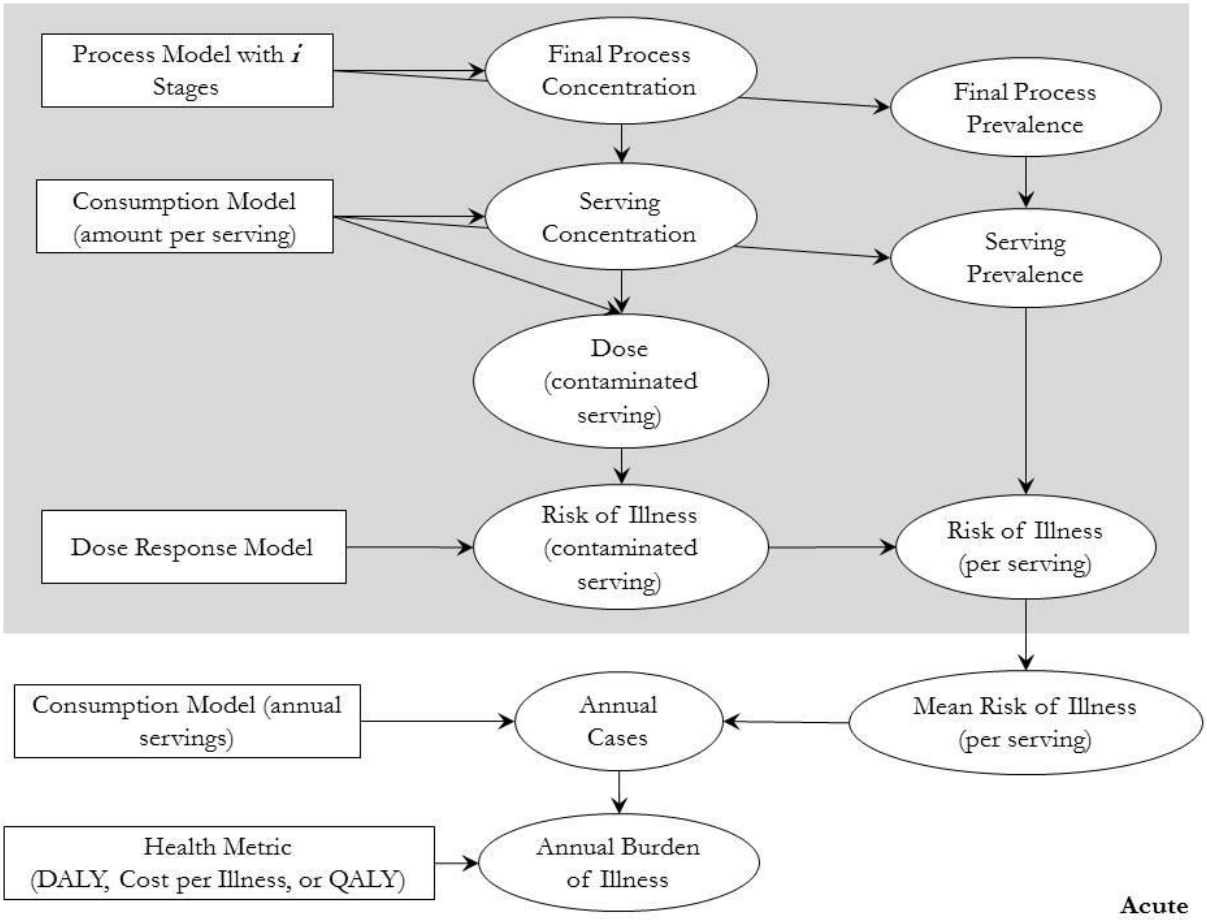


Figure 2. Schematic representation of the mathematical structure of a risk scenario for acute exposure (i.e., microbial hazards and acute exposures from chemical hazards). Rectangles represent user input and ovals represent FDA-iRISK results.

4.2 Consumption Model for Acute Exposure

4.2.1 Population Groups for Acute Exposure

The risk scenario for acute exposure assumes that illness results from a single exposure to a certain amount of microbial pathogen or chemical. The effect of this dosage can depend on the individual consuming the food, both in terms of the probability of becoming ill, and in terms of the severity or type of illness. FDA-iRISK therefore permits the user to define various mutually exclusive population groups for consideration in a risk scenario for a single acute exposure.

For example, pregnant women and the elderly are more likely to become ill than middle-aged non-pregnant consumers given the same dose of the bacterium, *Listeria monocytogenes*. In addition, such illness in a pregnant woman can affect the newborn child, whereas illness in the elderly is more likely to result in death than is illness in the general population. Therefore, when

creating a risk scenario involving acute exposure to hazards, such as *L. monocytogenes,* the user can define population groups of i) pregnant women, of ii) the elderly, and of iii) the general population. The user can also define consumption models, dose response models, and health metrics that are specific to each group.

The population groups in a risk scenario for acute exposure represent *mutually exclusive segments* of the population of interest that differ in terms of one or more of: consumption pattern, susceptibility to infection/illness, and type or severity of health impacts resulting from infection or illness. The sum of the eating occasions per year across the groups must account for all annual eating occasions of the food in the greater population.

4.2.2 Risk Per Person per Year

Users can also optionally define a distribution for the number of servings consumed by a single person for microbial hazards (e.g., Normal(200,30)), to obtain the risk per person per year - a measure used to describe risk by Lindqvist et al., 2019. If included, FDA-iRISK will multiply this distribution by the final mean risk per serving (e.g., 0.003 DALYs) to generate a distribution of annual risk for individuals. This result will then be included in the scenario's Risk Estimates and Ranking report as a chart. This feature requires the scenario to be run as variability-only.

Note that this calculation is based on the final mean risk per serving computed by FDA-iRISK which is a fixed value derived by simulating a stochastic FDA-iRISK scenario model. This calculated risk per person per year is obtained under the assumption that exposures from servings to servings are independent, and given that the mean risk per serving for the population and the DALY per case can be applied to the individual. There may be situations where it is not appropriate to assume that the risk per person per year reported captures the full potential for interindividual variability given that other individual risk factors (e.g., consumer behavior in storing, cooking, and serving sizes) will increase the variability in risks faced by individuals, and these should be considered independently of and in addition to the variability in the number of servings consumed by a single person in a year.

4.2.3 Calculation of Amount Consumed per Eating Occasion

The outputs of the consumption model for a risk scenario for acute exposure, are the mass of the food consumed per eating occasion (may be a distribution), and the number of eating occasions per year across the population of interest. These are explicitly defined by the user.

4.3 Chronic Exposure

FDA-iRISK uses a chronic exposure structure for those chemicals that may occur in food in levels too low to pose an immediate risk of illness, but that can cause illness after a long period of regular exposure at these low levels. The FDA-iRISK methodology has built upon reported

probabilistic exposure models in the assessment of dietary exposure to chemicals (Paoli et al., 2025).

In a risk scenario for chronic exposure, the consumption model is used to generate a value for the average amount of the food consumed per day (on a per unit body weight basis) over a "lifetime" of exposure. Lifetime is the entire length of lifespan when it is defined using life expectancy data, such as in averaging consumption over the lifetime for assessing cancer risks. Alternatively, users may define and document an exposure window less than lifetime, such as in assessing lead exposure during childhood that can result in chronic health effects. In either scenario, it takes into account the different daily amounts that may be eaten at different life stages, the body weight during those stages and the duration of those life stages relative to the entire lifespan or the shorter exposure window. This amount is then multiplied by the average concentration of hazard in the food, a value that represents all servings consumed in a lifetime and that is determined by both the average concentration of the hazard and the prevalence of contamination. The result is the Lifetime Average Daily Dose (LADD), a term used in FDA-iRISK for a risk scenario considering the entire lifespan or one considering a shorter exposure window (in the latter, the user documents the time window "lifetime" indicates). These individual LADDs are used in the dose response model to obtain a mean risk of illness per consumer. For a full discussion of the LADD and the role of the LADD in exposure (and risk) assessment see EPA (1992).

The mean risk of illness per consumer is then multiplied by the user-specified number of consumers (again from the consumption model) to predict the total number of cases over the user-defined duration of exposure. Each case is assigned a value for burden (in DALYs or COI) and in this way the overall burden for the exposure is calculated (see Figure 3). The overall burden is divided by the duration of the exposure to arrive at a value for annual burden. (This value for annual burden is the basis of the rankings.)

Where consumption or body weight is expressed as a probability distribution, daily consumption and body weight are sampled anew at each life stage. Thus, it is possible that a 10-year old weighing 50 kg and consuming 10 g of the food a day in one life stage (ending at 10 years of age) will be simulated as an 11 year old weighing 40 kg and consuming 20 g of the food a day in the subsequent life stage.

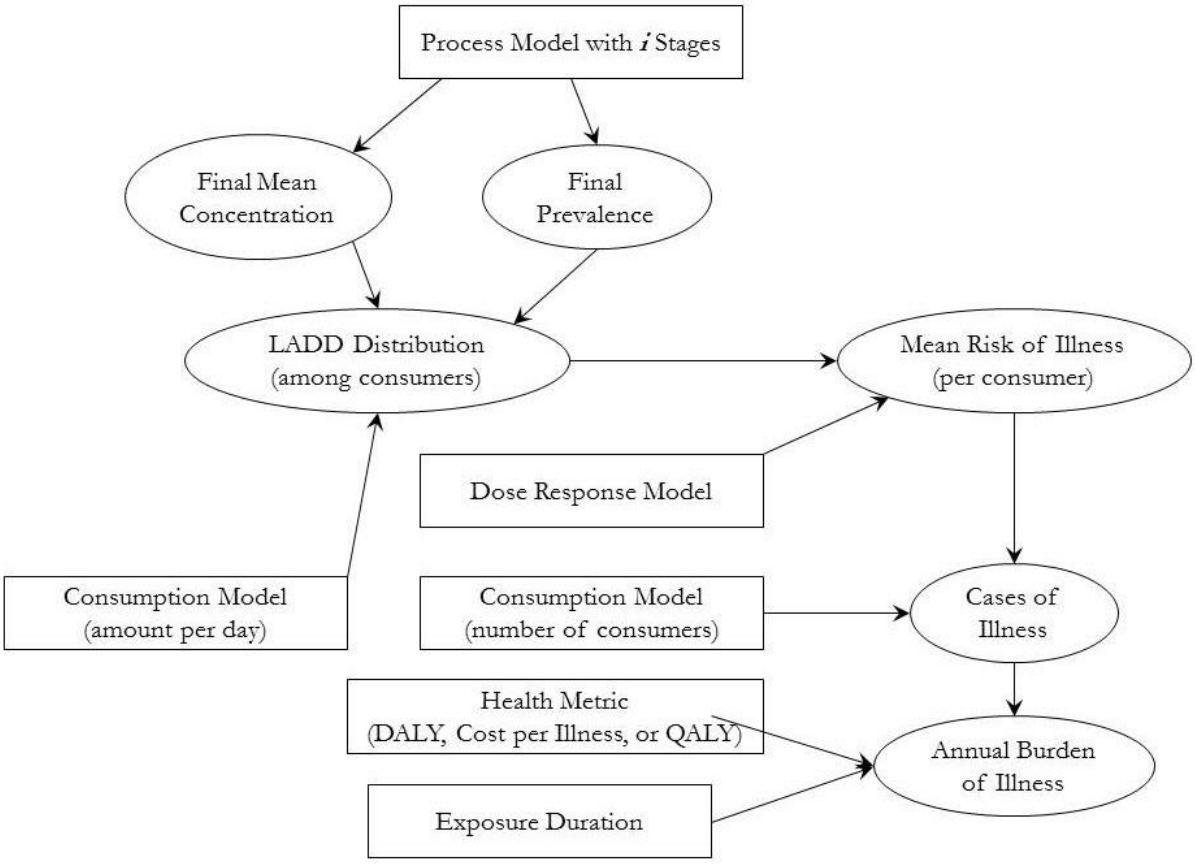


Figure 3. Schematic representation of the mathematical structure of a risk scenario for chronic exposure to chemical hazards. Rectangles represent user input and ovals represent FDA-iRISK results. LADD refers to Lifetime Average Daily Dose.

4.3.1 Chronic Exposure for Multifood Scenarios

FDA-iRISK introduces the concept of multifood chronic scenarios in which the exposure from multiple food sources is aggregated to compute a LADD over all food types prior to applying the dose-response model to compute the mean risk of illness per consumer.

A multifood scenario is otherwise very similar to a single food scenario (see Figure 4). A per capita consumption distribution is typically used for each of the foods in the FDA-iRISK method of aggregation of chronic exposure from multiple foods to approximate exposure overtime at the population level. It should be noted that not all foods would be consumed by an individual on a given day and thus, the consumption estimate in a multifood scenario (especially with a large number of foods across different dietary patterns) could overestimate the consumption. A more nuanced intake assessment may be conducted outside of FDA-iRISK, e.g., as in a study reported by Santillana Farakos et al. (2021), to consider nationwide food consumption survey data at the individual level to derive joint consumption of multiple foods; the resulting consumption

estimates from such a study may be imported into FDA-iRISK as empirical consumption distributions.

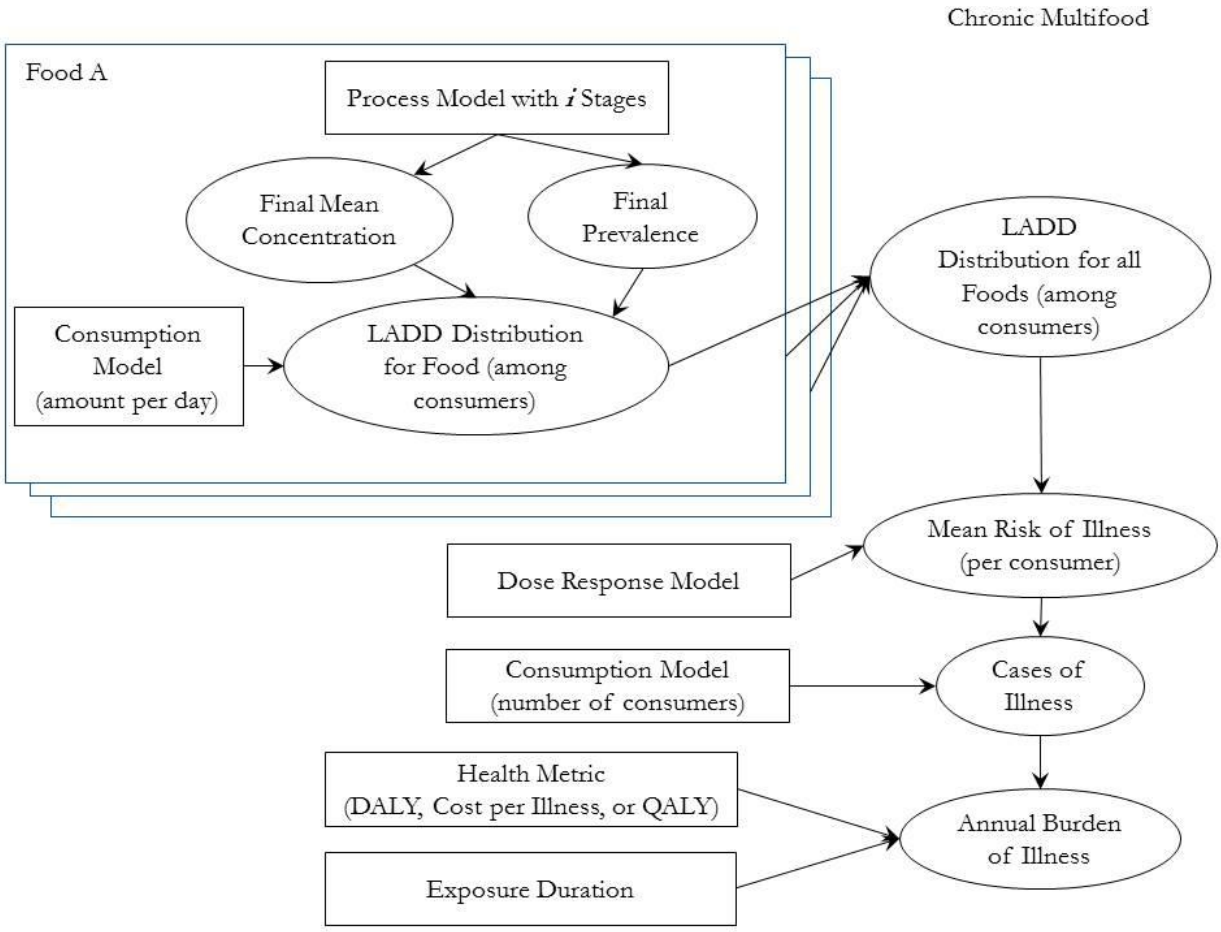


Figure 4. Schematic representation of the mathematical structure of a multifood chronic exposure scenario. Rectangles represent user input and ovals represent FDA-iRISK results. LADD refers to Lifetime Average Daily Dose.

4.3.2 Chronic Exposure for Multihazard - Multifood Scenarios

FDA-iRISK introduces the concept of multihazard - multifood chronic scenarios for risk benefit and tradeoff analysis. The exposure from each hazard is reported separately and aggregated to compute a LADD over all food types prior to applying the dose-response model to compute the mean risk of illness per consumer.

4.4 Consumption Model for Chronic Exposure

4.4.1 Life Stages for Chronic Exposure

The risk scenario for chronic exposure, typically assumes that illness results from lifetime exposure to low levels of a particular chemical. As the relevant dose units are expressed per unit

body weight, the actual dose can change over the course of a lifetime as the individual's weight changes with age. In addition, an individual may consume more or less of the food on an absolute basis over the course of a life. For these reasons, the effective dose is calculated as the LADD described above. Age of consumer is not explicitly considered in FDA-iRISK but can be captured 2 ways, 1) using the population groups life stages (described below). Should a specific age be of interest then a single population group of that age can be created, or 2) using the DALY templates to encompass age through the duration of the health outcomes, for example years of life lost would consider age at death.

Calculation of the LADD necessitates user-input of a daily average consumption amount and body weight (may be distributions) for each life stage defined, as well as the time span covered by each stage.

The life stages in a risk scenario for chronic exposure represent *sequential stages* experienced by the group of individuals enumerated in the user-defined "Number of Consumers".

4.4.2 Life Stages for Multifood Chronic Exposure

Life stage consumption data for multifood scenarios differs slightly from single food scenarios. For multifood scenarios, consumption from different food sources is aggregated over a common population. Therefore, the consumption data provided for each food must be for the common population and not just consumers of that specific food. As different foods will be consumed by different fractions of the population, the distribution used to describe the consumption values will necessarily include a proportion of consumers with zero consumption.

To achieve this, consumption data are parameterized using two elements: the cumulative empirical consumption distribution (for eaters only) and the percentage of consumers included in that distribution. These two elements are then combined to produce the per capita consumption distribution for the entire population, which includes both eaters and non-eaters. For example, assume a percentage of consumers of 40% and the following cumulative distribution of the amount consumed for those consumers:

Table 4_1. Example of consumers only distribution

Probability	Consumption (g/kg-day)
0	0
0.1	0.886
0.2	2.86
0.3	4.464
0.4	10.46
0.45	15.86
0.65	22.2
0.8	29.39
0.9	37.04

Probability	Consumption (g/kg-day)
0.99	47.57
1	90.06

Given 40% of the population are consumers, then 60% are not. Using this data, a per capita distribution can be created by starting consumption at a cumulative probability of 0.6 and then shifting and scaling the per consumer probabilities:

Table 4_2. Example of calculated per capita distribution

Probability Scaling	Per Capita Probability	Consumption (g/kg-day)
0	0	0
= 1-0.4	0.6	0
= (1-0.4) + (0.4*0.1)	0.64	0.886
= (1-0.4) + (0.4*0.2)	0.68	2.86
= (1-0.4) + (0.4*0.3)	0.72	4.464
= (1-0.4) + (0.4*0.4)	0.76	10.46
= (1-0.4) + (0.4*0.45)	0.78	15.86
= (1-0.4) + (0.4*0.65)	0.86	22.2
= (1-0.4) + (0.4*0.8)	0.92	29.39
= (1-0.4) + (0.4*0.9)	0.96	37.04
= (1-0.4) + (0.4*0.99)	0.996	47.57
= (1-0.4) + (0.4*1)	1	90.06

This results in the following per consumer (i.e., consumers only) and per capita cumulative distributions (Figure 5a). For the per capita distribution, the amount consumed increases from 0 starting at 0.6 and eventually matches the amount for the per consumer distribution. If the percentage of consumers is set to 100%, then both distributions will be the same.

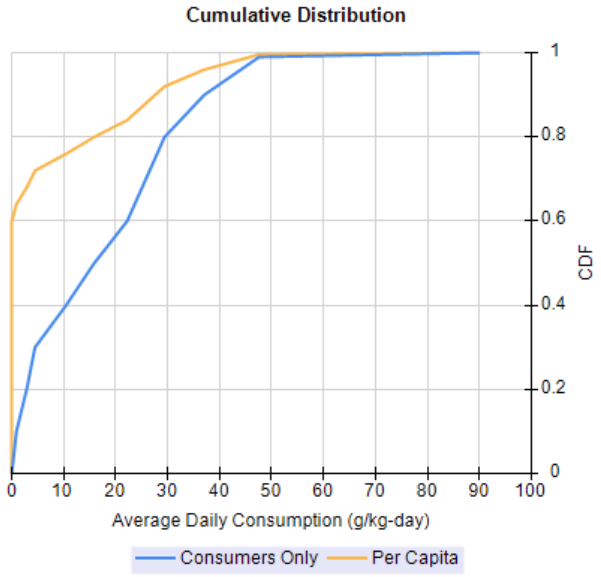


Figure 5a. Cumulative consumers only and per capital consumption distributions

4.4.3 Calculation of Lifetime Average Daily Consumption (LADC)

The outputs of the Consumption Model for a risk scenario for chronic exposure, are the average daily amount of the food consumed (may be a distribution) and the number of consumers within the population of interest. The time-weighted average daily amount of food consumed per unit bodyweight is termed the Lifetime Average Daily Consumption (LADC) of the food, and is calculated as:

$$LADC = \frac{\sum_{i=1}^{n} A_i \times y_i}{\sum_{i=1}^{n} y_i}$$

Equation 45

Where:

- n is the number of life stages defined by the user.
- i refers to the i^{th} life stage.
- A_i is the average daily amount consumed per kg of body weight during the life stage i.
- y_i is the duration of the life stage i in years.

As described above, the term “lifetime” can be defined as the entire lifespan or as an exposure window less than a lifetime. Thus, the calculated LADC may represent the average daily consumption over the entire lifespan, or the shorter time window based on the duration of the life stages as defined and documented by the user. FDA-iRISK provides flexibility that enables the user to define a critical window of exposure for relevant health endpoint and calculate average daily dose during that critical window (not necessarily lifetime). The use of the generic terms “lifetime” and “LADC” in the equations and Web interface of FDA-iRISK is for simplicity, while the

user chooses the entire lifespan or a shorter time window in chronic exposure assessment and document the scenarios accordingly.

4.4.3.1 Example of Lifetime Average Daily Dose (LADD) Calculation

FDA-iRISK models chronic exposure by simulating a large number of individual lifetime exposure patterns that are possible within the population of consumers. One pattern is simulated for each iteration of the model and the patterns may vary from iteration to iteration. For example, assuming no correlation in life stage consumption, the first iteration might represent a lifetime exposure pattern characterized by very high childhood exposure, followed by very low subsequent exposure, while the second iteration might represent a pattern featuring high exposure in childhood, youth, and old age but low exposure in middle age. **Note:** see Section 4.4.4 for alternative assumption for correlated life stage consumption.

The overall dose assigned to each of these lifetime exposures is the LADD. That is, the daily dose of the chemical ingested by the consumer (attributable to the food in question) averaged over the lifetime. The contribution of each life stage to this average is proportional to the length of the life stage. In this way, the changing exposure over the course of the lifetime is condensed into a single value representing the lifetime average daily exposure. The two iterations described above, for example, might both be represented by the same LADD, regardless of the timing of the different exposure peaks. (While average daily dose (µg/kg) in childhood is often larger than the adult average daily dose due to the lower body weight in childhood, the comparatively shorter period associated with this age compensates to some extent.) The LADD is then provided to the dose response model to obtain a mean risk of illness per consumer.

Required Inputs

Consider the calculation of the LADD of inorganic arsenic (iAs) in apple juice. The inputs required by FDA-iRISK for chronic exposure are food consumption in grams per day and body weight for the life stages, as well as the mean level of the contaminant in the food. The mean is appropriate for widely sourced foods consumed on a regular basis, and is computed by FDA-iRISK from the process model. Users may specify the concentration as a fixed value or a distribution in the process model but, for purposes of risk assessment for chronic exposure, FDA-iRISK will compute a mean concentration value from the final stage of the process model for use in computing the LADD.

For the purposes of this example, it is assumed the mean level of iAs in apple juice is 4.43 ng/g (ppb).

Basic Calculation for One Iteration

Body weight and consumption data populate rows 1 and 2 of Table 1, with row 3 generated by dividing the consumption in g/day by the body weight in kg.

The columns represent the age ranges associated with the user-defined life stages of the population under evaluation. In each iteration of the simulation, FDA-iRISK uses data from each age group to build a single "lifetime exposure".

Note: For Table 4_3, it is assumed that FDA-iRISK has sampled a random value from each of the consumption distributions associated with the life stages. (see *Table 4.4* for results of a different sample.)

Row 4 displays the time span occupied by each of the user-defined population age groups, and row 5 represents the fraction of the total exposure period contributed by each of these groups. This fraction can then be used to "weight" the row 3 values to obtain the component (average daily consumption, ADC) of lifetime average daily consumption that is contributed by each age (row 6).

Summing over the values in row 6 produces the LADC (of apple juice), shown in row 7.

The LADC in g/kg-day is then multiplied by the mean iAs occurrence in apple juice (4.43 ng/g) in order to calculate the LADD in ng/kg-day (row 8):

Table 4_3. Calculation of the LADD for inorganic arsenic from apple juice – iteration 1

Age Range	**2 to 10**	**11 to 17**	**18 to 64**	**65 to 85**
1) Body weight (kg)	26	57	80	80
2) Consumption (g/day)	53	20	10	15
3) Cons. by wt. (g/kg-d)	2.04	0.35	0.13	0.19
4) Time Span (years)	9	7	47	21
5) Fraction of total span	0.107	0.083	0.560	0.25
6) ADC (g/kg-day)	0.218	0.029	0.070	0.047
7) LADC (g/kg-day)	0.364			
8) LADD (ng/kg-day)	1.61			

This example represents a random iteration in which random samples have been drawn from the consumption (g/day) distribution for each life stage. (**Note:** Some values are rounded for presentation purposes.) See section 4.4.4 for the methodology on using correlated samples.

Basic Calculation for a Second Iteration

When consumption data are provided as a distribution, for example by using the Cumulative Empirical option to input percentile consumption data, FDA-iRISK samples a single value from each distribution in each iteration of the simulation. In other words, for one simulated lifetime exposure that the tool builds from the input data, a consumption value from the high end of the distribution might be selected to represent consumption in the youngest age group, while a consumption value from the low end of the distribution might be selected to represent every other age group. All combinations are possible. Table 4_4 illustrates the LADD calculation resulting from an iteration in which random samples have been drawn from the consumption distributions corresponding to high consumption at a young age and lower consumption at older ages.

Table 4_4. Calculation of the LADD for inorganic arsenic from apple juice – iteration 2

Age Range	**2 to 10**	**11 to 17**	**18 to 64**	**65 to 85**
1) Body weight (kg)	26	57	80	80
2) Consumption (g/day)	130	5	0.6	0
3) Cons. by wt. (g/kg-d)	5.00	0.088	0.0075	0
4) Time Span (years)	9	7	47	21
5) Fraction of total span	0.107	0.083	0.560	0.25
6) ADC (g/kg-day)	0.536	0.0073	0.0042	0
7) LADC (g/kg-day)	0.547			
8) LADD (ng/kg-day)	2.42			

This example represents random samples that have been drawn from the consumption distributions corresponding to high consumption at a young age and lower consumption at older ages. (**Note:** Some values are rounded for presentation purposes.)

Variations on the Basic Calculation

Note: If the user provides single (fixed) values to represent consumption for the different age groups, and the mean body weight (per age group) and inorganic arsenic level are also fixed (as shown here), then all iterations of the simulation will produce the same estimated LADD value.

If, on the other hand, distributions are used to represent body weight, rather than the fixed values used in this example, there will be a wider range of values possible for each cell in row 3, and by extension for the LADD estimate. FDA-iRISK does not enforce correlation between body weight and consumption, so when building a single simulated lifetime exposure, a body weight from the low end of the distribution can be combined with a consumption value from the high end of that distribution.

For situations where a food is not consumed for the entire lifespan, for example food for infants, an entry for the remaining time span when the food is not consumed is entered in the same way as a food which is consumed (as illustrated in the table above), but with an associated consumption entered as 0 grams for that time span. This ensures the LADD is calculated based on the complete lifetime time span.

4.4.4 Correlated Life stage Consumption for Multifood Chronic Exposure

Users have the option to specific positive or negative correlation coefficients for the life stage consumption distributions defined as part of a multifood chronic consumption model. Each coefficient is defined in relation to the previous life stage's distribution. For example, an individual that consumes a significant amount of rice (e.g., more than average amount) for their body weight during the childhood life stage might also consume rice more than average during their next life stage.

Consider a consumption model defined with three life stages. If high consumption at each stage correlates with higher consumption at later stages, the user might specify a positive correlation as follows:

Table 4_5. Example of correlation coefficient for consumption at three life stages

Life Stage	Start Age	End Age	Correlation Coefficient
Infants less than 1 year	0yr 0mo	0yr 11mo	N/A
Children 1 to 6 years	1yr 0mo	6yr 11mo	0.75
Persons 7 to under 50 years	7yr 0mo	49yr 11mo	0.75

The correlation coefficient is not specified for the first life stage distribution as correlation is only applied at each subsequent stage.

For a case where a strong positive correlation is modeled, a shift in exposure values across the entire distribution should occur compared to not applying correlation. This effect will be more evident in the leading tail (low end) and trailing tail (high end) of the sampled distribution. For

example, consider the illustrative example shown below (Table 4_6). Consumers that do not eat the food (i.e., consumption = 0 g/kg-day) are correlated at each life stage such that the percentile of consumers that do not eat any of the food increases from 0.07 to 0.16 when correlated. Similarly, consumers that eat larger amounts of the food are correlated such that average daily consumption increases at the higher end of the percentiles (e.g., at the 99th percentile, correlated consumption is 8.4 g/kg-day vs. 7.0 g/kg-day when life stages are not correlated).

Table 4_6. Example of consumption without or with correlation among life stages

Low End			High End		
Percentile	**No Correlation**	**Correlated**	**Percentile**	**No Correlation**	**Correlated**
0.01	0	0	0.80	1.52118	1.48946
0.02	0	0	0.81	1.57895	1.59735
0.03	0	0	0.82	1.65290	1.67827
0.04	0	0	0.83	1.75250	1.75009
0.05	0	0	0.84	1.81514	1.86394
0.06	0	0	0.85	1.92233	1.95599
0.07	0	0	0.86	2.07043	2.06928
0.08	0.00009	0	0.87	2.20697	2.23968
0.09	0.00022	0	0.88	2.31444	2.38177
0.10	0.00062	0	0.89	2.42056	2.50234
0.11	0.00127	0	0.90	2.54765	2.73274
0.12	0.00213	0	0.91	2.68098	2.87751
0.13	0.00299	0	0.92	2.91852	3.09444
0.14	0.00435	0	0.93	3.10636	3.27776
0.15	0.00549	0	0.94	3.35887	3.65281
0.16	0.00652	0	0.95	3.61912	4.04518
0.17	0.00740	0.00003	0.96	4.21902	4.82602
0.18	0.00838	0.00009	0.97	4.72492	5.50018
0.19	0.00967	0.00015	0.98	5.55042	6.66142
0.20	0.01210	0.00020	0.99	7.03720	8.43169

The following chart (Figure 5b) illustrates a second example where positive correlation was applied (Grey Line) compared to not applied (Red Line) showing the shift in the aggregate Lifetime Average Daily Dose. Note that for correlated life stage consumption the result is a wider distribution of exposure with a greater number of the population with lower doses and a higher dose in the higher percentiles.

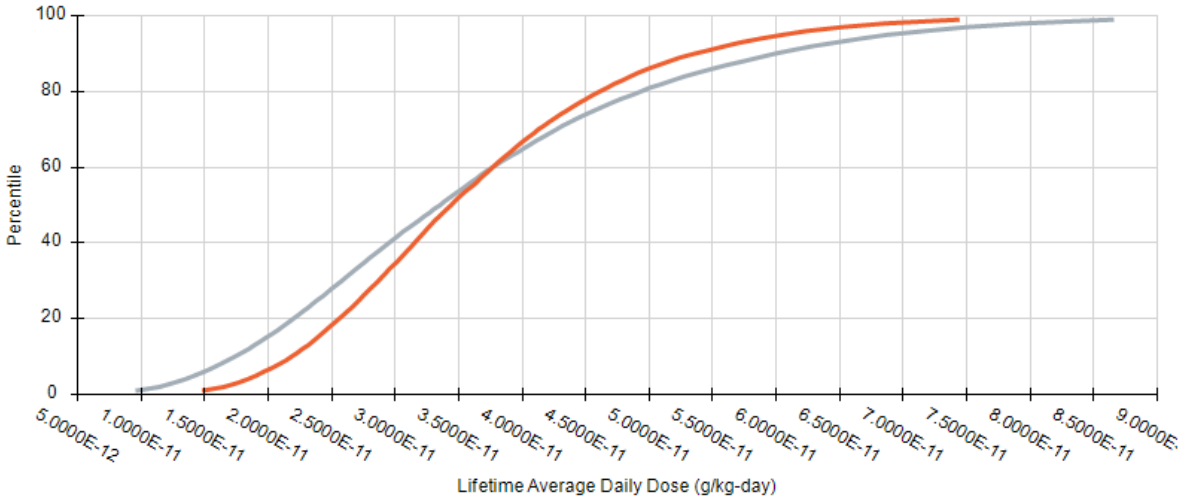


Figure 5b. Example of consumption distribution with correlation among life stages, with positive correlation applied (grey line) compared to not applied (red line).

Note: Users should carefully consider the correlation values as they represent the correlation between distributions of the integrated metric of per kg-day consumption and not the separate elements of total consumption or body weight. As such, by default, the correlation coefficient value is set to 0 for no correlation and the coefficients are not applied unless enabled by the user for a given scenario.

4.4.4.1 Correlation Method

The correlation coefficient is applied to consumption distributions in the simulation model using the built in function "Correlate_With" in Analytica™ to create a chain of correlated distributions. This function uses Rank (Spearman) Correlation instead of Pearson Correlation. Rank Correlation is appropriate in this context because the correlation in consumption is not a measure of linearity and it does not involve a Gaussian joint distribution. According to Analytica™ documentation (Lumina, 2023):

> For non-Gaussian distributions, it is not necessarily possible for two distributions to have a desired Pearson correlation. However, we can ensure a given Rank Correlation, also called Spearman correlation.

Additional details on the application of correlation are described in statistical publications such as for Pearson correlation (Pearson, 1895), Spearman correlation (Spearman, 1904), correlation in non-Gaussian distributions (Joe, 1989), and applications and interpretation (Cohen et al., 2003).

4.4.5 Diet Shifts for Chronic Multifood Consumption Models and Scenarios

FDA-iRISK includes a diet feature. This feature allows users to define consumption models for a set of common foods included in different diets (e.g. Mediterranean, Paleo, Vegetarian, Vegan).

Multifood risk scenarios that use diets will simulate the different consumption patterns to produce a risk estimate for each diet based on a common set of process models (for hazard concentration), dose-response models, and health impacts.

Chronic multifood consumption models are parameterized using two elements: the cumulative empirical consumption distribution and the percentage of consumers included in that distribution, as described in section **4.4.2 Life Stages for Multifood Chronic Exposure.**

The diet feature also allows users to define a change in the consumption distribution for a life stage under a diet, referred to as a diet shift. Users have the option to select one of six diet shift types for each life stage in each consumption model for each food included in the diet. These shifts are applied to the default consumption distribution defined for the life stage.

- Use Default.
- Set to Zero
- Percent Consumers Only
- Linear Amount Adjustment
- Exponential Percentile Adjustment
- Custom Distribution

Use Default

The default distribution for the life stage will be used for the diet.

Set to Zero

The consumption amount for this food and life stage will be set to zero for the diet.

Percent Consumers Only

The underlying cumulative per consumer distribution (i.e., the Consumers Only distribution) remains unchanged but the percentage of consumers used to create the per capita distribution can be increased or decreased.

Linear Amount Adjustment

The percentage of consumers used to create the per capita distribution can be increased or decreased or set to the same as the default value.

A linear amount adjustment factor is defined as a value greater than 0. This value is used to increase or decrease the amount consumed as defined in the default Consumers Only consumption distribution. With this option, minimum and maximum consumption values are also modified.

For example, assume the following default consumption model based on a percentage of consumers of 40%, with the following per consumer and per capital distributions:

Table 4_7. Example of consumption distribution: per consumer and per capita

Per Consumer		Per Capita	
Probability	**Value**	**Probability**	**Value**
0	0	0	0
0.1	0.886	0.6	0
0.2	2.86	0.64	0.886
0.3	4.464	0.68	2.86
0.4	10.46	0.72	4.464
0.45	15.86	0.76	10.46
0.65	22.2	0.78	15.86
0.8	29.39	0.86	22.2
0.9	37.04	0.92	29.39
0.99	47.57	0.96	37.04
1	90.06	0.996	47.57
		1	90.06

If the percentage of consumers is left at 40% but an amount adjustment factor of 1.7 is defined, the following distributions result:

Table 4_8. Example of shifted consumption distributions with linear scaling

Original Per Consumer		Shifted Per Consumer		Shifted Per Capita	
Probability	**Value**	**Probability**	**Value**	**Probability**	**Value**
0	0	0	0	0	0
0.1	0.886	0.1	1.5062	0.6	0
0.2	2.86	0.2	4.862	0.64	1.5062
0.3	4.464	0.3	7.5888	0.68	4.862
0.4	10.46	0.4	17.782	0.72	7.5888
0.45	15.86	0.45	26.962	0.76	17.782
0.65	22.2	0.65	37.74	0.78	26.962
0.8	29.39	0.8	49.963	0.86	37.74
0.9	37.04	0.9	62.968	0.92	49.963
0.99	47.57	0.99	80.869	0.96	62.968
1	90.06	1	153.102	0.996	80.869
				1	153.102

The probability values remain unchanged, but the amount of consumption is linearly increased in the shifted per consumer distribution by a factor of 1.7 for each row and this is also reflected in the final per capita distribution.

Exponential Percentile Adjustment

The percentage of consumers used to create the per capita distribution can be increased or decreased or set to the same as the default value.

An exponential percentile adjustment factor, *w*, is defined as a value greater than 0. The adjustment factor is applied to each probability value of the default distribution, *p*, as p^w. When *w* is less than one, the distribution shifts to lower consumption. When *w* is greater than one, consumption is increased. When *w* equals one, the distribution is unchanged. In this option, regardless of changes in w, the minimum and maximum consumption values remain unchanged. This type of shift is designed to explore scaling consumption without changing the minimum and maximum values. Although a published reference is not yet available for food consumption use case based on an exponential percentile adjustment factor, this is nonetheless an option to quickly shift a consumption distribution to various shapes (e.g., symmetrical, light-tailed, heavy-tailed) for exploratory analysis.

For example, assume the same default distribution as used in the linear scaling example and an exponential adjustment factor of 2. The following distributions result:

Table 4_9. Example of shifted consumption distributions with distribution warping

Original Per Consumer		Shifted Per Consumer		Shifted Per Capita	
Probability	**Value**	**Probability**	**Value**	**Probability**	**Value**
0	0	0	0	0	0
0.1	0.886	0.01	0.886	0.6	0
0.2	2.86	0.04	2.86	0.604	0.886
0.3	4.464	0.09	4.464	0.616	2.86
0.4	10.46	0.16	10.46	0.636	4.464
0.45	15.86	0.2025	15.86	0.664	10.46
0.65	22.2	0.4225	22.2	0.681	15.86
0.8	29.39	0.64	29.39	0.769	22.2
0.9	37.04	0.81	37.04	0.856	29.39
0.99	47.57	0.9801	47.57	0.924	37.04
1	90.06	1	90.06	0.99204	47.57
				1	90.06

For the per consumer (consumers only) distribution, the consumption values remain unchanged, but the cumulative probability is scaled exponentially by the adjustment factor value of 2 (p^2) in the shifted per consumer distribution. This change is also reflected in the shifted per capita distribution.

Custom Distribution

The percentage of consumers used to create the per capita distribution can be increased or decreased or set to the same as the default value. The user provides a new cumulative consumption distribution for consumers, which overrides the default distribution.

4.5 Variability Distributions for Amount per Eating Occasion and Body Weight

The following variability distributions (Table 4_10) are available in FDA-iRISK for users to define "Amount per eating occasion" for single food consumption models (for acute exposure to both microbial and chemical hazards) and to define Body Weight (kg) for chronic single food consumption models (for chemical hazards). Two of the options, Empirical (cubic) and Empirical (linear) apply to chronic multifood consumption models, where the user can define or import empirical consumption distribution considering both amount consumed and body weight (g/kg-day).

Table 4_10. Distributions for consumption (amount per eating occasions and body weight)

Distribution	Parameters
Beta	Alpha, Beta
Beta General	Alpha, Beta, Lower bound, Upper bound
Beta PERT	Minimum, Mode, Maximum
Chance	Probabilities, Values (for probabilities)
Empirical (cubic)	Probabilities (must include 0 and 1), Values (for probabilities)
Empirical (linear)	Probabilities (must include 0 and 1), Values (for probabilities)
Fixed Value	Value
Gamma	Shape, Rate
Log10Uniform	Log10 Minimum, Log10 Maximum
Log10Uniform (Percentiles)	Log10 5^{th} Percentile, Log10 95^{th} Percentile
Lognormal	Arithmetic Mean of X, Standard Deviation
LogUniform	Ln Minimum, Ln Maximum
LogUniform (Percentiles)	Ln 5^{th} Percentile, Ln 95^{th} Percentile
Normal	Mean, Standard Deviation
Normal (Truncated)	Mean, Standard Deviation, Lower bound, Upper bound
Triangular	Minimum, Mode, Maximum
Triangular (Percentiles)	5^{th} Percentile, Mode, 95^{th} Percentile
Triangular (Truncated)	Minimum, Mode, Maximum, Lower bound, Upper bound
Uniform	Minimum, Maximum
Uniform (Percentiles)	5^{th} Percentile, 95^{th} Percentile

5 Estimation of Cases of Illness: Dose Response Models

The process model produces a value (which may be a distribution) for the concentration of the hazard among contaminated units of the food at the point of consumption (i.e., including any consumer storage and/or cooking steps). It also yields a value for the prevalence of that contamination and the mass of each unit of food.

The consumption model provides a value (which may be a distribution) for the food consumed. The dose of the hazard to be applied in the dose response model is then determined by the mass of the food consumed and the hazard concentration in that food. The specific calculation of the dose depends on whether the exposure is acute or chronic.

5.1 Dose Calculation, Acute Exposure

The acute dose, AD, is calculated for servings of food in the case of acute exposure and is given by:

$$AD = C_s \times M_s \text{ or}$$
$$AD = \frac{C_s \times M_s}{BW}$$

Equation 46

where:

- C_s is the concentration by mass or volume unit of food (e.g., cfu/g, pfu/g, the number of oocysts or virus particles per unit mass of food, the number of microorganisms per ml) in contaminated units at consumption (the *s* subscript refers to *servings*).
- M_s is the serving size (mass or volume amount) consumed at an eating occasion.
- *BW* is the body-weight of the consumer.

The contamination of the serving of food at consumption is obtained from the output of the process model and using the Mass Change process type (not seen by the user), applying either pooling or partitioning, according to the relative size of the final mass of the unit of food from the process model, and the mass of the serving to convert from final processing unit size to serving size.

Note: The user has the option of specifying the dose units for acute exposures to a chemical as either {mass of substance}/{kg body-weight}, or simply {mass of substance}. This determines which of the two forms of the dose equation is applied. This allows for the option for acute exposures to cause illness at a probability that is dependent only on the amount of the substance consumed, and independent of the body-weight of the consumer.

5.2 Dose Calculation, Chronic Exposure

The dose applied in the case of chronic exposure is a LADD, which is equivalent to the weighted (by life stage duration) average of average daily doses (ADDs) for each life stage, across the duration of exposure, typically a lifespan. In FDA-iRISK, the weighting by life stage duration is implemented prior to calculating the dose, so the LADD is calculated as the weighted average daily consumption of the food, multiplied by the average concentration of hazard:

$$LADD = LADC \times mean(C_i \times P_i)$$

Equation 47

where:

- C_s is the mass of hazard per mass or volume unit of food in a contaminated serving of food at consumption (i.e., at the end of the process model).
- P_s is the prevalence of contaminated servings of food at consumption.
- *LADC* is the lifetime average daily consumption of the food, in mass or volume units per kg-day, calculated as described in the Consumption Model (see *Section 4.4.2).*
- *Life stage* consumption data for multifood scenarios differs slightly from single food scenarios. For multifood scenarios, consumption from different food sources is aggregated over a common population. Therefore, the consumption data provided for each food must be for the common population and not just consumers of that specific food. As different foods will be consumed by different fractions of the population, the distribution used to describe the consumption values will necessarily include a proportion of consumers with zero consumption. As such, only the cumulative empirical distribution is available in FDA-iRISK to describe consumption patterns for multifood scenarios. Other options are being considered (e.g., a discrete chance distribution).

5.2.1 Lifetime Daily Average Dose Threshold Test

Users can optionally specify a LADD threshold value for a single food, chronic chemical scenario (e.g., 0.3 ng/g). FDA-iRISK will compare the LADD computed for each iteration and return a value of 1 when the threshold is exceeded or 0 when the LADD is less than or equal to the defined threshold. FDA-iRISK then computes the mean of these results to determine the proportion of the population whose LADD exceeds the threshold and includes this result in the scenario's Risk Estimates and Ranking report.

5.3 Dose Response Models for Microbial Hazards (Acute Exposures)

All microbial hazards are assumed to act on an acute exposure basis. FDA-iRISK provides the following model options for acute exposures to microbial hazards:

- Beta-Poisson
- Empirical
- Exponential
- Non-Threshold Linear
- Threshold Linear
- Weibull

Doses for microbial hazards are expressed as cfu, pfu, or other as specified by the user.

In addition to the parameters listed for the dose response models described below, the user is required to provide a percentage value for probability of illness given response.

Note: For microbial hazards, FDA-iRISK uses a modeling approach that, for each iteration, tracks the prevalence (proportion) of contaminated food units and the number of bacteria in the contaminated food units. In each case, the contamination level is, by design, greater than or equal to 1 cfu (pfu) per food unit. This has implications for the formulations used for the Exponential and Beta-Poisson dose response models (Pouillot et al., 2015), as described below.

5.3.1 Beta-Poisson

The modeling approach used in FDA-iRISK generates individual doses. In keeping with the use of individual doses[1], the Beta-Poisson dose response model is implemented as a Beta-Binomial dose response function. This modeling approach corresponds to Method 2 as described in Pouillot *et al*., 2015. The specific implementation uses the Beta function alternative described by equation 11 of Haas, 2002:

$$P\ (dose, \alpha, \beta) = 1 - \left(\frac{\mathrm{B}(\alpha, \beta + dose)}{\mathrm{B}(\alpha, \beta)} \right)$$

Equation 48

where:

- *dose* is the dose on the non-log scale.
- α and β are parameters of the dose- response model.
- B(α, β) is the Beta function.
- $\alpha > 0$, $\beta > 0$

An example of a Beta-Poisson dose response model is shown in Figure 6.

[1] Note: The individual dose is required for this model, which differs from the classical Beta-Poisson model, P(dose, α, β) = 1 - (1 + dose/β)^α, where the dose used represents a mean dose ingested.

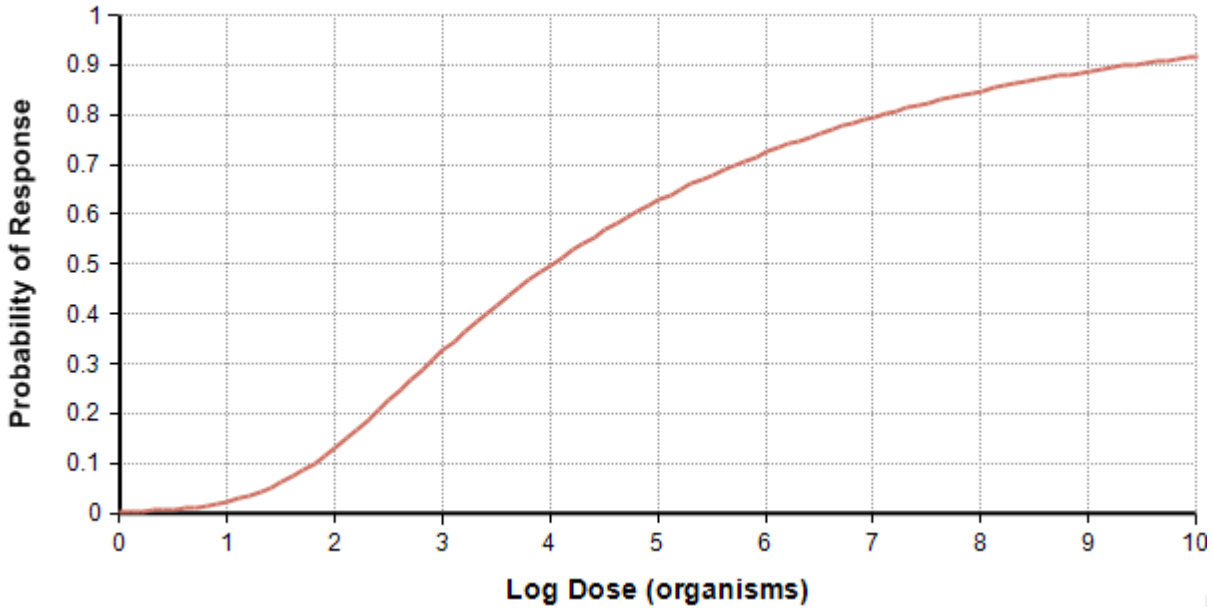


Figure 6. The Beta-Poisson dose-response relationship when α is 0.13 and β is 51. Note that the dose is shown on the log10 scale.

5.3.2 Empirical

The empirical dose-response model is used to create custom dose-response models using a set of concentration/probability of response data points. FDA-iRISK offers either linear or cubic interpolation to determine the probability of response between the specified doses.

5.3.3 Exponential

The modeling approach used in FDA-iRISK generates individual doses[2] . In keeping with the use of individual doses, the Exponential dose response model is implemented as a Binomial dose-response function. This modeling approach corresponds to Method 2 as described in Pouillot *et al*., 2015:

$$P\ (dose, r) = 1 - (1 - r)^{dose}$$

Equation 49

where:

- *dose* is the exposure dose (non-log scale).

[2] Note: The individual dose is required for this model, which differs from another form of the exponential model, P = 1 - exp(-rd), where the dose used represents a mean dose ingested.

- *r* is the probability that a single ingested organism is able to survive and initiate infection or illness (depending on how "response" is defined).

In the exponential model, the value of r is defined specifically for each pathogen (and, should the user so choose, for each population group). An example of an Exponential dose response model is shown in Figure 7.

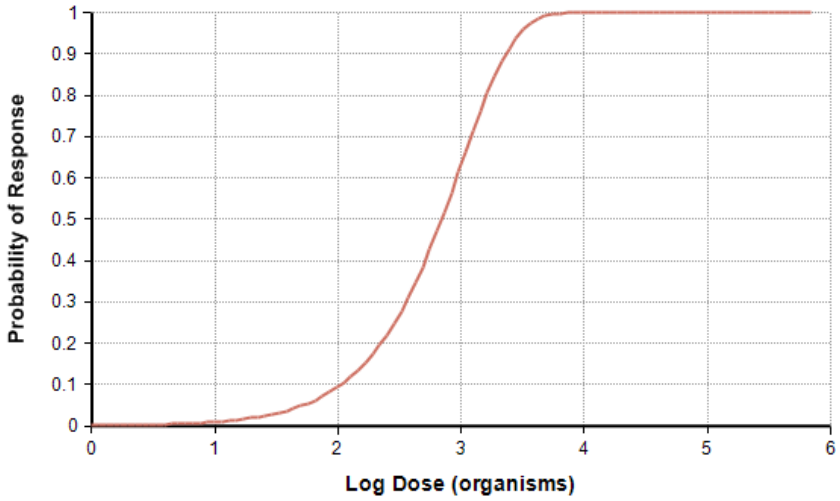


Figure 7: The Exponential dose-response relationship given an *r* value of 0.001. Note that the dose is shown on the $\log_{10}$ scale.

5.3.4 Non-Threshold Linear

Given a user-specified dose ("Reference Point") on the $\log_{10}$ scale and associated risk at that dose ("Risk at Reference Point"), combined with the assumption that the risk is zero at (and only at) zero exposure, a linear relationship is obtained. The probability of response in general can then be determined as:

$$P\ (dose, RfP) = dose \times \left(\frac{RiskatRfP}{10^{RfP}} \right)$$

Equation 50

where:

- *dose* is the exposure dose expressed in cfu or pfu.
- *RfP* is the user-specified dose ("Reference Point"), expressed in $\log_{10}$ units.
- *RiskatRfP* is the user-specified probability of response given exposure to dose *RfP* ("Risk at Reference Point").
- The probability of response is limited to not exceed 1.

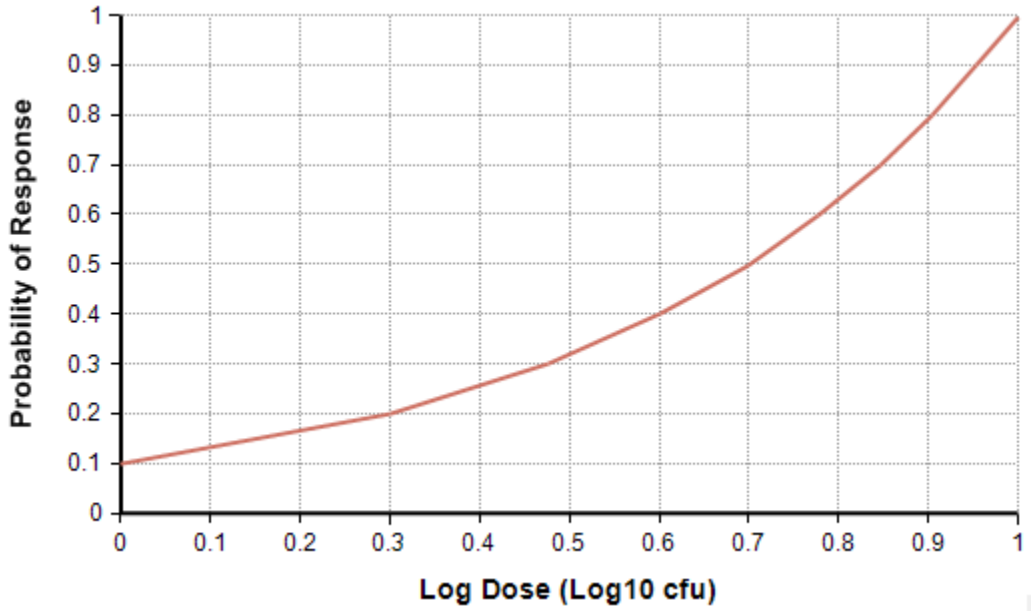


Figure 8. A hypothetical Non-Threshold Linear dose response relationship where the user specifies that the probability of response at a dose of 0.7 $\log_{10}$ cfu (5 cfu) is 0.5.

5.3.5 Threshold Linear

The Threshold Linear model assumes a linear relationship between the level of exposure (dose) and the probability of response. It also assumes that there is a threshold effect in this relationship such that for numbers of organisms below the threshold, there is a zero probability of response, but for numbers above the threshold, the dose response is linear, so that the probability of response in general is:

$$P(dose, T, RfP, RiskatRfP) = \begin{cases} 0 & d \le T \\ (dose - T) \times \left(\dfrac{RiskatRfP}{RfP - T} \right) & d > T \end{cases}$$

Equation 51

where:

- *dose* is the exposure dose.
- *RfP* is the user-specified dose ("Reference Point").
- *RiskatRfP* is the user-specified probability of response given exposure to dose *RfP* ("Risk at Reference Point").
- *T* is the user-specified threshold below which the probability of response is zero.
- The probability is limited not to exceed 1.

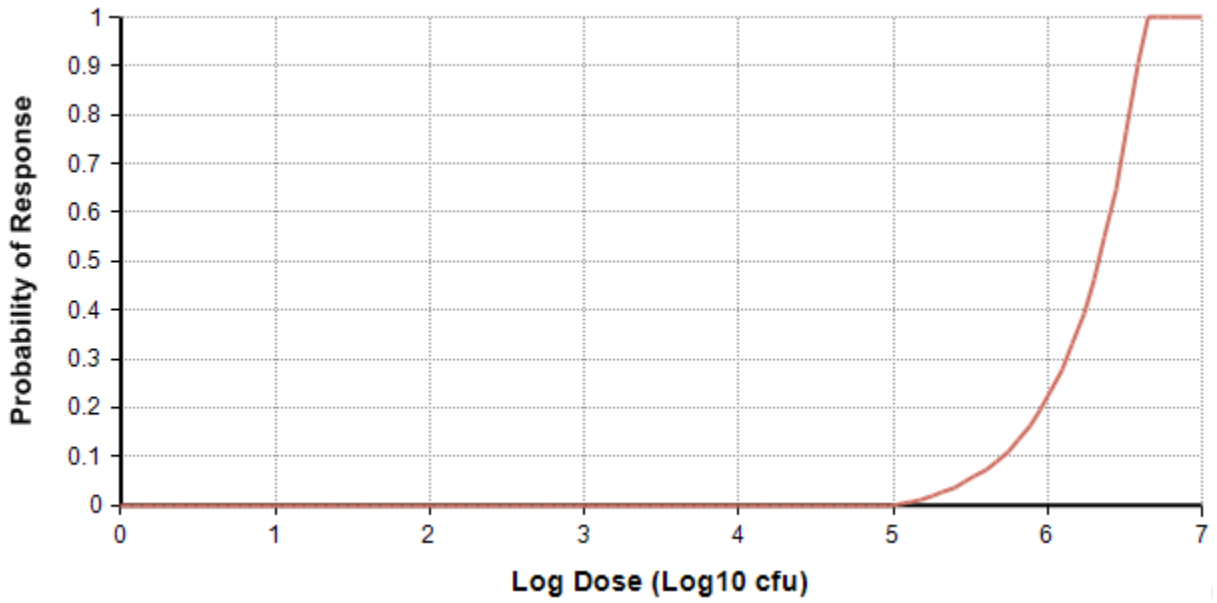


Figure 9. The Threshold Linear dose-response relationship given a Reference Point dose of 6 $\log_{10}$ cfu with associated probability of response of 0.22, and a threshold of 5 $\log_{10}$ cfu.

5.3.6 Weibull

The following formula is used (Haas, 1999):

$$P(dose.\alpha,\beta) = 1 - \exp(-\beta \times dose^{\alpha})$$

Equation 52

where:

- α (power) ≥ 1[3] and β (slope) > 0 are parameters of the dose response model.

An example of a Weibull dose response model is shown in Figure 10**.**

[3] Restricted to ≥ 1 based on EPA (2012).

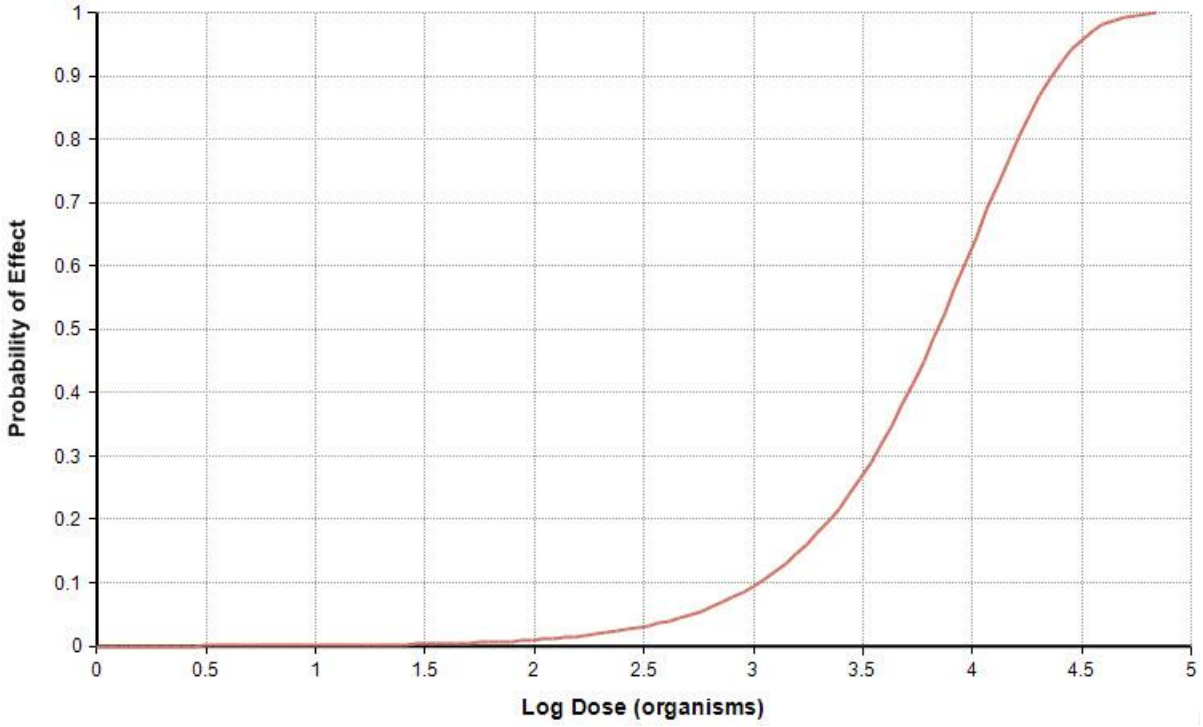


Figure 10. An example of a Weibull dose-response relationship with β = 0.0001 and α = 1. Note that the dose is shown on the log10 scale.

5.4 Dose Response Models for Chemical Hazards (Acute Exposures)

Chemical hazards may act based on an acute or a chronic exposure. The tool provides the following model options for acute exposures to chemical hazards:

- Cumulative Lognormal
- Empirical
- Linear by Slope Factor
- Non-Threshold Linear
- Step Threshold
- Threshold Linear
- Weibull

The units of dose for acute chemical exposures are expressed in terms of either mass or mass/(kg of bodyweight). This allows the user the option to model acute exposures as causing illness at a probability that is dependent only on the amount of the substance consumed, and independent of the body-weight of the consumer. The option to use mass rather than mass per unit body weight, may be appropriate for some substances that trigger a response (e.g., immediate and localized) that is independent of the mass of the consumer).

In addition to the parameters listed for the dose response models described below, the user is required to provide a percentage value for probability of illness given response. This allows the response to be a sub-clinical event (like a positive biomarker with or without illness), and the probability of illness to represent the fraction of sub-clinical events that result in a sufficiently adverse response as to be considered an illness.

5.4.1 Cumulative Lognormal

The dose response relationship is a re-parameterization of the Log-Probit model (described below), based on the cumulative distribution of the log-normal distribution or the normal distribution when using log-transformed values for the dose, the median effective dose (i.e. the ED_{50} is the dose causing the predicted effect in 50% of the exposed population) and geometric standard deviation (GSD).

$$P(dose, ED_{50}, GSD) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\ln(dose)} \exp\left(\frac{-(t-\mu)^2}{2\sigma^2}\right) dt$$

Equation 53

where

- Where $\mu = \ln(ED_{50})$ and $\sigma = \ln(GSD)$.
- ED_{50} is the dose causing a 50% probability of response.
- GSD is the geometric standard deviation.
- $\mu > 0$ and $\sigma > 0$; $ED_{50}>0$ and GSD>0.

An example of a Cumulative Lognormal dose response model is shown in Figure 11.

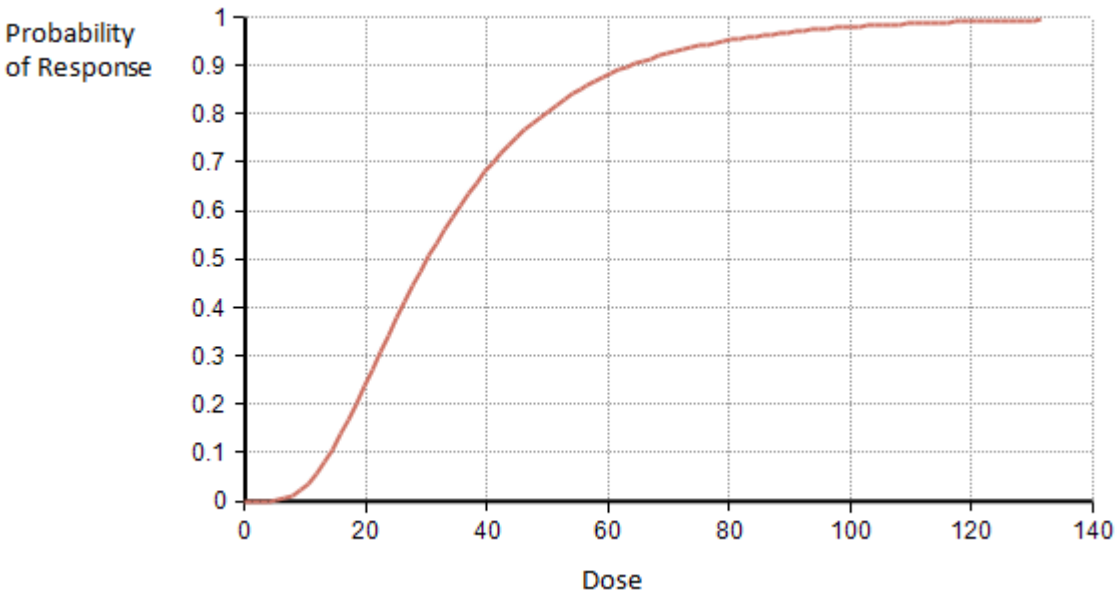


Figure 11. The Cumulative Lognormal dose-response relationship when the ED_{50} is 30 and the GSD is 1.8.

5.4.2 Empirical

The empirical dose-response model is used to create custom dose-response models using a set of concentration/probability of response data points. FDA-iRISK will use linear interpolation to determine the probability of response between the specified doses.

5.4.3 Linear by Slope Factor

This dose response is another parameterization of the non-threshold linear dose response. Given a user-specified slope factor, a linear relationship is obtained:

$$P(dose, c) = dose \times c$$

Equation 54

where:

- *dose* is the exposure dose.
- *c* is the slope.
- The probability is limited not to exceed 1.

An example of the Linear by Slope Factor dose response model is shown in Figure 12.

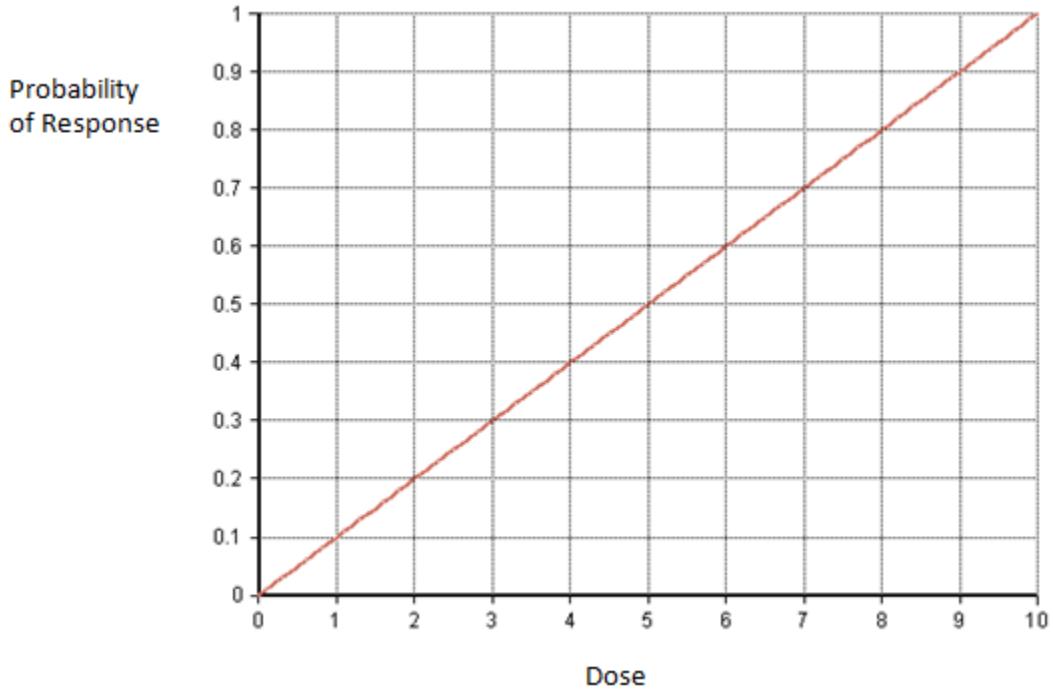


Figure 12. The Linear by Slope Factor dose response model where the slope is 0.1.

5.4.4 Non-Threshold Linear

Given a user-specified dose ("Reference Point") and associated risk at that dose ("Risk at Reference Point"), combined with the assumption that the risk is zero at (and only at) zero dose, a linear relationship is described. The probability of response in general can then be determined as:

$$P\left(dose, RfP, RiskatRFP\right) = dose \times \left(\frac{RiskatRfP}{RfP}\right)$$

Equation 55

where:

- *dose* is the exposure dose.
- *RfP* is the user-specified dose ("Reference Point").
- *RiskatRfP* is the user-specified probability of response given exposure to dose *RfP* ("Risk at Reference Point"). The probability of response is limited to not exceed 1.

An example of a Non-Threshold Linear dose response model is shown in Figure 13.

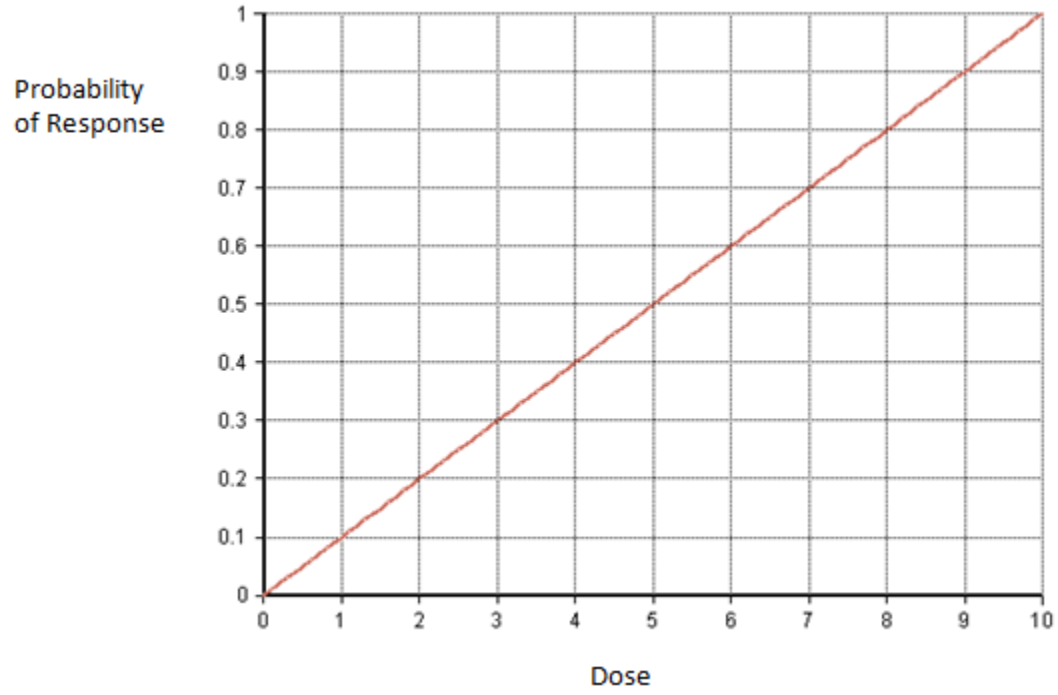


Figure 13. The Non-Threshold Linear dose-response relationship where the probability of response at a dose of 5 is 0.5.

5.4.5 Step Threshold

The Step Threshold model assumes that given a user-specified threshold, exposure at or below this threshold results in zero risk of health effects, and exceedance of this threshold results in a 100% probability of response, specifically:

$$P(d,T) = \begin{cases} 0 & d \le T \\ 1 & d > T \end{cases}$$

Equation 56

where:

- d is the dose ingested by the consumer.
- T is the user-specified threshold below which there is no response.

An example of a Step Threshold dose response model is shown in Figure 14.

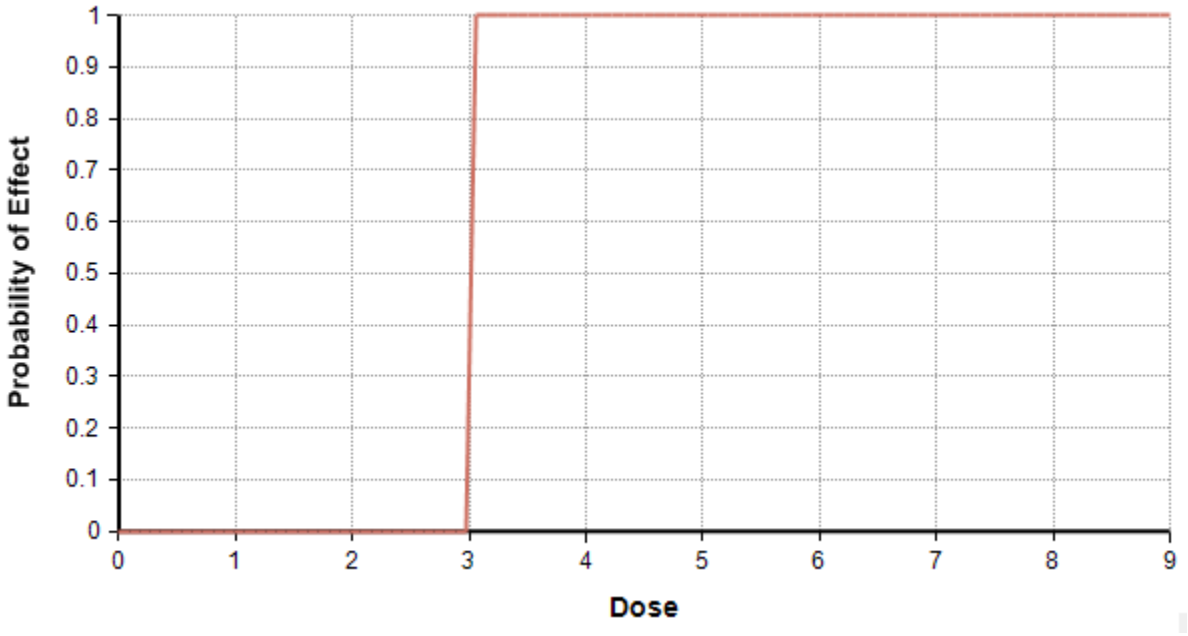


Figure 14. The Step Threshold dose-response relationship where the threshold is given as 3.

5.4.6 Threshold Linear

The Threshold Linear model assumes a linear relationship between the level of exposure (dose) and the probability of response. It also assumes that there is a threshold effect in this relationship such that below the threshold, there is a zero probability of response, but above the threshold the dose response relationship is linear, so that the probability of response in general is:

$$P(dose, T, RfP, RiskatRfP) = \begin{cases} 0 & d \le T \\ (dose - T) \times \left(\dfrac{RiskatRfP}{RfP - T} \right) & d > T \end{cases}$$

Equation 57

where:

- *dose* is the exposure dose.
- *RfP* is the user-specified dose (“Reference Point”).
- *RiskatRfP* is the user-specified probability of response given exposure to dose *RfP* (“Risk at Reference Point”).
- *T* is the user-specified threshold below which the probability of response is zero.
- The probability is limited not to exceed 1.

An example of a Threshold Linear dose response model is shown in Figure 15.

Figure 15. The Threshold Linear dose-response relationship given a Reference Point dose of 5 with associated probability of response of 0.5, and a threshold of 2.

5.4.7 Weibull

The following formula is used (based on USEPA, 2012):

$$P(dose.\alpha,\beta) = 1 - \exp(-\beta \times dose^{\alpha})$$

Equation 58

where:

- $\alpha \geq 1$[4] and $\beta > 0$ are parameters of the dose response model.

[4] Restricted to ≥ 1 based on EPA (2012).

An example of a Weibull dose response model is shown in Figure 16.

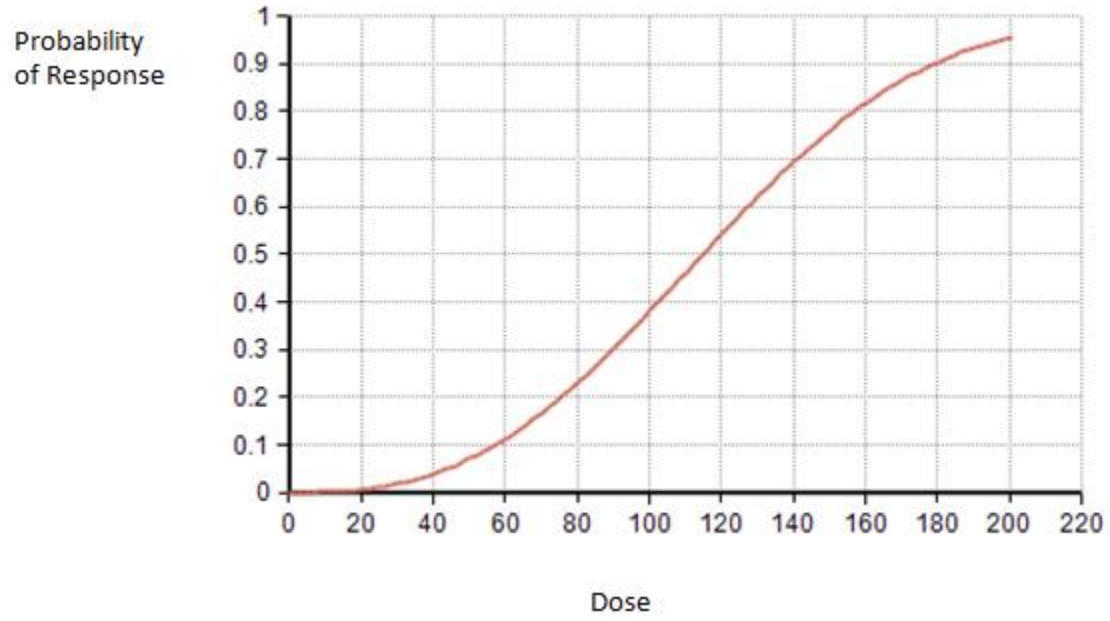


Figure 16. An example of a Weibull dose response relationship with β = 1.9E-6 and α = 2.7.

5.5 Dose Response Models for Chemical Hazards – Chronic Exposures

For chronic exposures to chemical hazards, FDA-iRISK provides the following model options for the dose response relationship:

- Cumulative Lognormal*
- Decreasing Log10-Logistic
- Decreasing Logistic
- Decreasing Log-Logistic
- Decreasing Probit
- Empirical*
- Gamma
- Linear by Slope Factor*
- Logistic
- Log-Logistic
- Log-Logistic with Background
- Multistage
- Non-Threshold Linear*
- Probit
- Restricted Log-Probit

- Restricted Weibull
- Step Threshold*
- Threshold Linear*
- Weibull*

*Described in *Section 5.4 Dose Response Models for Chemical Hazards (Acute Exposures)*.

The models are based on the notation employed in the Benchmark Dose Software (BMDS) from EPA (USEPA, 2012) for dose-response modeling. This allows dose response models developed using the BMDS software to be easily implemented in FDA-iRISK. Users need to ensure the correct dose-units are selected. Note that the dose-units do not need to be adjusted by the user to be the same as those used in the process model. FDA-iRISK adjusts the dose units from the process model to match those of dose response model (e.g., dividing doses expressed in μg by 1000, if the dose response model is expecting mg).

All doses for chronic chemical exposures are expressed in mass/kg-day, where kg refers to the body weight of the consumer.

In addition to the parameters listed for the dose response models described below, the user may provide a percentage value for probability of illness given response (100% is the default value). This is intended to allow for conversion from estimates of response (which may not result in illness) to estimates of a more adverse effect that would be considered a health outcome.

5.5.1 Decreasing Log10-Logistic

The 'nfDecreasingLog10Logistic' function uses the log-logistic model and requires intercept (β_0) and slope (β_1) parameters. The function assumes that data have been fit using the $\log_{10}$ value of the dose. Note that $\beta_1 < 0$.

The $\log_{10}$-logistic model is:

$$P(Effect|log_{10}Dose) = \frac{1}{1 + \exp\left(-(\beta_0 + \beta_1 * log_{10}Dose)\right)}$$

Equation 59

An example of a Log10-Logistic dose response model is shown in Figure 17.

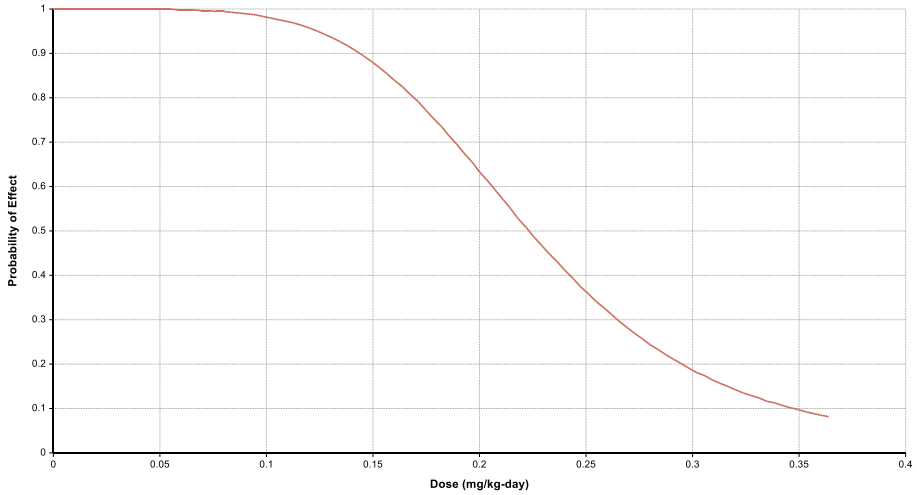


Figure 17. The Log10-Logistic dose response relationship where β_0 = -7.4758 and β_1 = -4.9874.

5.5.2 Decreasing Logistic

The 'nfDecreasingLogistic' function uses the logistic model and requires intercept (β_0) and slope (β_1) parameters. Note that $\beta_1 < 0$.

The logistic model is given by:

$$P(Effect \mid Dose) = \frac{1}{1 + \exp\left(-(\beta_0 + \beta_1 * Dose)\right)}$$

Equation 60

An example of a Logistic dose response model is shown in Figure 18.Figure 17.

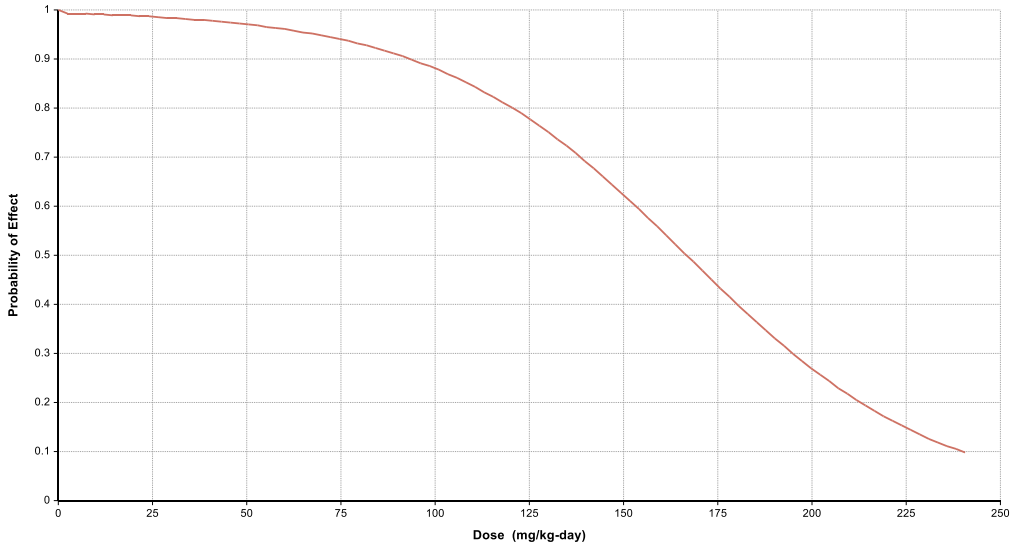


Figure 18. The Logistic dose response relationship where β_0 = 5 and β_1 = -0.03

5.5.3 Decreasing Log-Logistic

The 'nfDecreasingLogLogistic' function uses the log-logistic model and requires intercept (β_0) and slope (β_1) parameters. The function assumes that data have been fit using the natural log (ln) of the dose. Note that $\beta_1 < 0$.

The log-logistic model is given by:

$$P(Effect|logDose) = \frac{1}{1 + \exp\left(-(\beta_0 + \beta_1 * logDose)\right)}$$

Equation 61

5.5.4 Decreasing Probit

The 'nfDecreasingProbit' function uses the probit model and requires background (α) and slope (β) parameters. Note that $-\infty < \alpha < \infty$ and $\beta > 0$.

The probit model is given by:

$$P(Effect \mid Dose) = 1 - [\Phi(\alpha + \beta * Dose)]$$

Equation 62

where Φ is the cumulative distribution function of the standard normal distribution.

5.5.5 Gamma

The following formula is used (based on EPA, 2012, with background set equal to zero):

$$P(dose, \alpha, \beta) = \frac{1}{\Gamma(\alpha)} \int_0^{\beta \times dose} t^{\alpha-1} e^{-t} dt$$

Equation 63

where:

- $\alpha \geq 1$ is "power".
- $\beta \geq 0$ is "slope" (USEPA, 2012).

An example of a Gamma dose response model is shown in Figure 19.

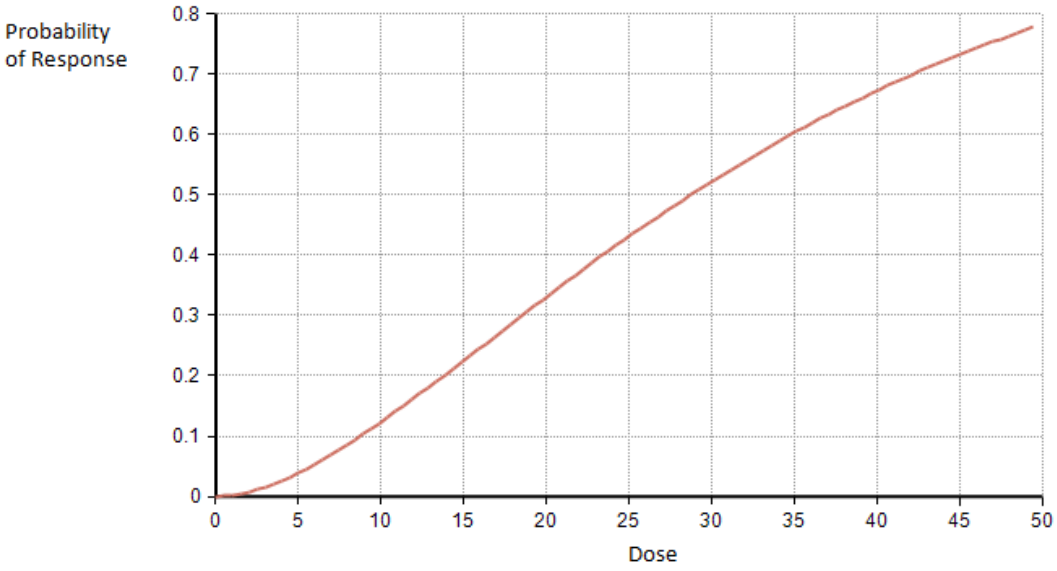


Figure 19. The Gamma dose-response relationship for a power of 1.9 and a slope of 0.055.

5.5.6 Logistic

The probability of response at a certain dose is given by:

$$P(dose, \alpha, \beta) = \frac{1}{1 + \exp(-(\alpha + \beta \times dose))}$$

Equation 64

where $\alpha > 0$ and $\beta > 0$ are parameters of the dose response model.

An example of a Logistic dose response model is shown in Figure 20.

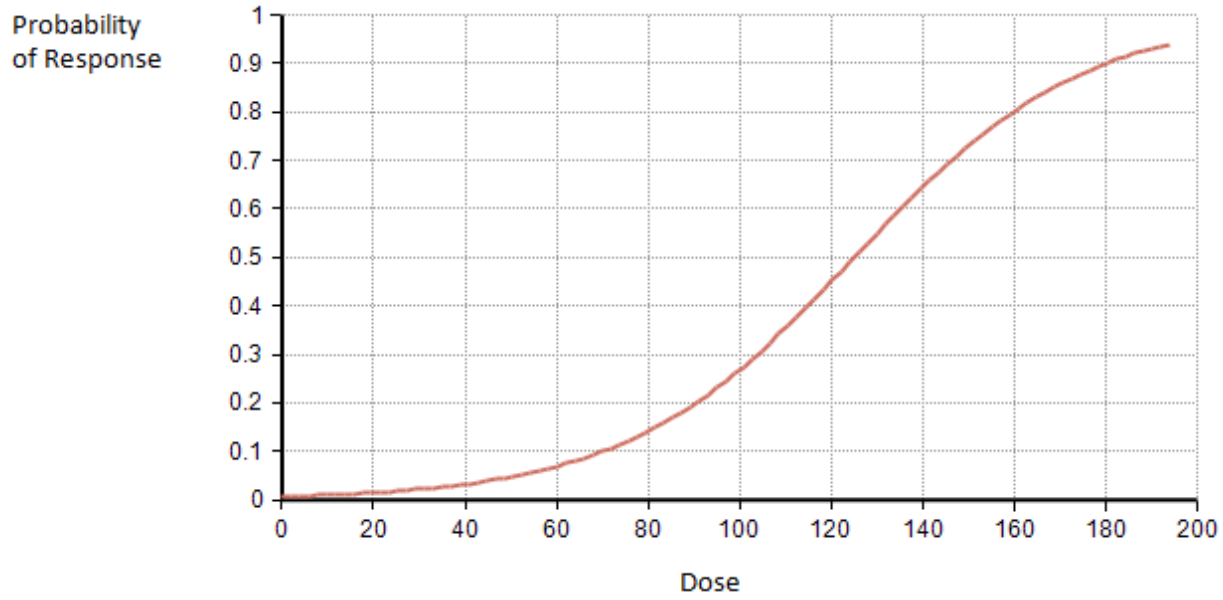


Figure 20. The Logistic dose response relationship where β = 0.04 and α = -5.

5.5.7 Log-Logistic

The following formula is used (based on USEPA, 2012):

$$P(dose, \alpha, \beta) = \frac{1}{1 + \exp(-(\alpha + \beta \times \ln(dose)))}$$

Equation 65

Where -∞< α < ∞ and β >= 1 are parameters of the dose response model.

An example of a Log-Logistic dose response model is shown in Figure 21.

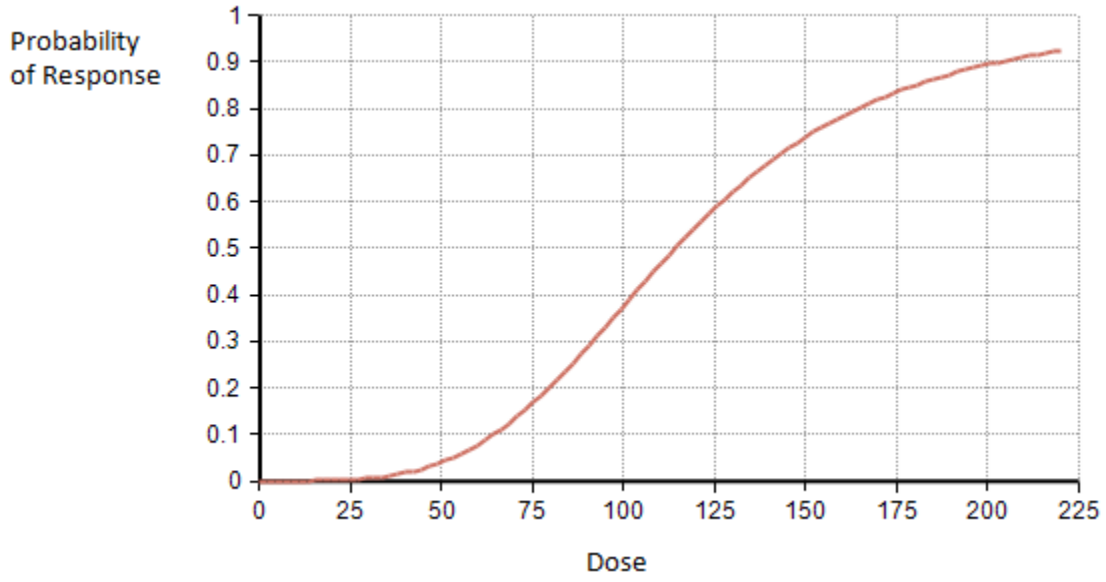


Figure 21. The Log-Logistic dose response relationship for α = -18 and β = 3.8.

5.5.8 Log-Logistic with Background

The Log-Logistic with Background dose response model introduces a background probability of response to the Log-Logistic dose response model described in *Section 5.5.7 Log-Logistic*. The standard formula for this dose response model is:

$$P\ (dose, \alpha, \beta) = \gamma + \frac{(1-\gamma)}{1+\exp(-(\alpha+\beta\times\ln(dose)))}$$

Equation 66

However, as FDA-iRISK is estimating additive risk, not including the background risk, the formula in FDA-iRISK removes the background risk:

$$P\ (dose, \alpha, \beta) = \frac{(1-\gamma)}{1+\exp(-(\alpha+\beta\times\ln(dose)))}$$

Equation 67

where:

- $-\infty < \alpha < \infty$ and $\beta \geq 1$ are parameters of the dose response model.
- $0 < \gamma < 1$ is the background probability of response.

An example of a Log-Logistic with background dose response model is shown in Figure 22.

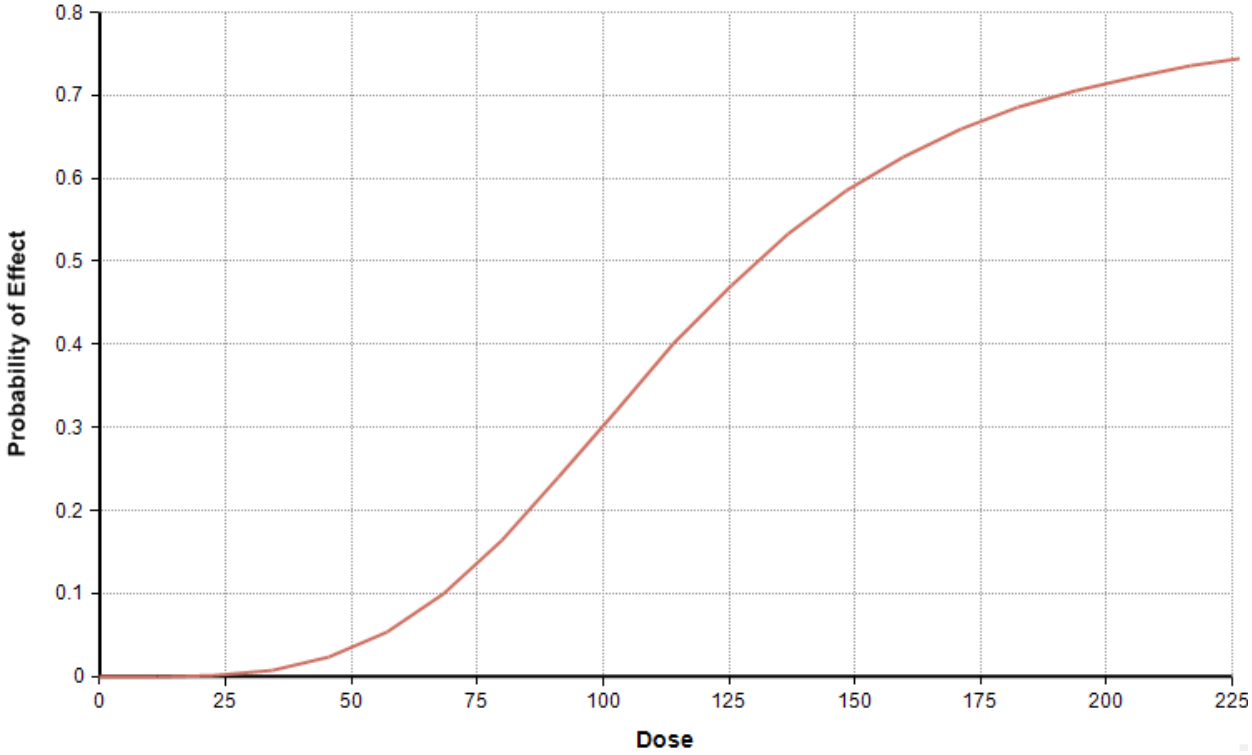


Figure 22. The Log-Logistic with Background dose response relationship for α = -18 and β = 3.8, and a background risk of 0.2.

5.5.9 Multistage

The following formula is used (based on USEPA, 2012):

$$P(dose, \beta_1, \beta_2, \beta_3) = 1 - \exp\left(-\sum_{j}^{3} \beta_j \times dose^j\right)$$

Equation 68

Where β_1, β_2, β_3 are parameters of the dose response model.

An example of a Multistage dose response model is shown in Figure 23.

Note: FDA-iRISK does not reproduce the equation from the appendix of the BMDS technical documentation (USEPA, 2012) exactly. The desired response is the additive risk, not including the background risk. The equation is therefore adjusted to reflect this by removing the background risk terms.

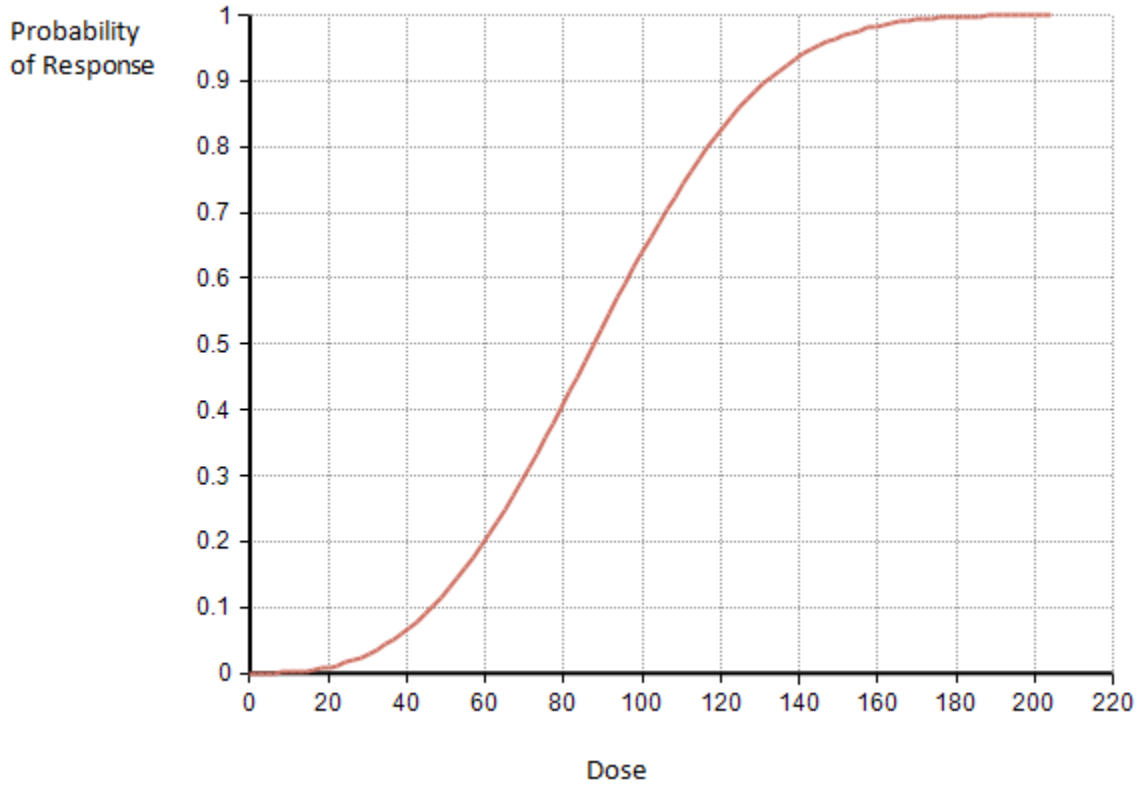


Figure 23. An example of a Multistage dose response relationship where the three parameter values are 2E-7, 2E-6, and 1E-6 for β_1, β_2, and β_3 respectively.

5.5.10 Probit Model

The Probit dose response relationship is based on the cumulative distribution of the normal distribution. The user specifies two parameters, *α and β:*

$$P(dose, \alpha, \beta) = \Phi(\alpha + \beta \times dose) - \Phi(\alpha)$$

Equation 69

where:

- Φ() is the Cumulative Distribution Function of the standard normal distribution (μ = 0, σ = 1).
- *α* is "intercept".
- *β* > 0 is "slope" (USEPA, 2012).

Note: FDA-iRISK does not reproduce the equation from the appendix of the BMDS technical documentation (USEPA, 2012) exactly, since the desired response is the additive risk, not including the background risk. This is achieved by subtracting the probability of response at zero dose, Φ(*α*). This ensures that the risk is zero when the dose is zero.

An example of a Probit dose response model is shown in Figure 24.

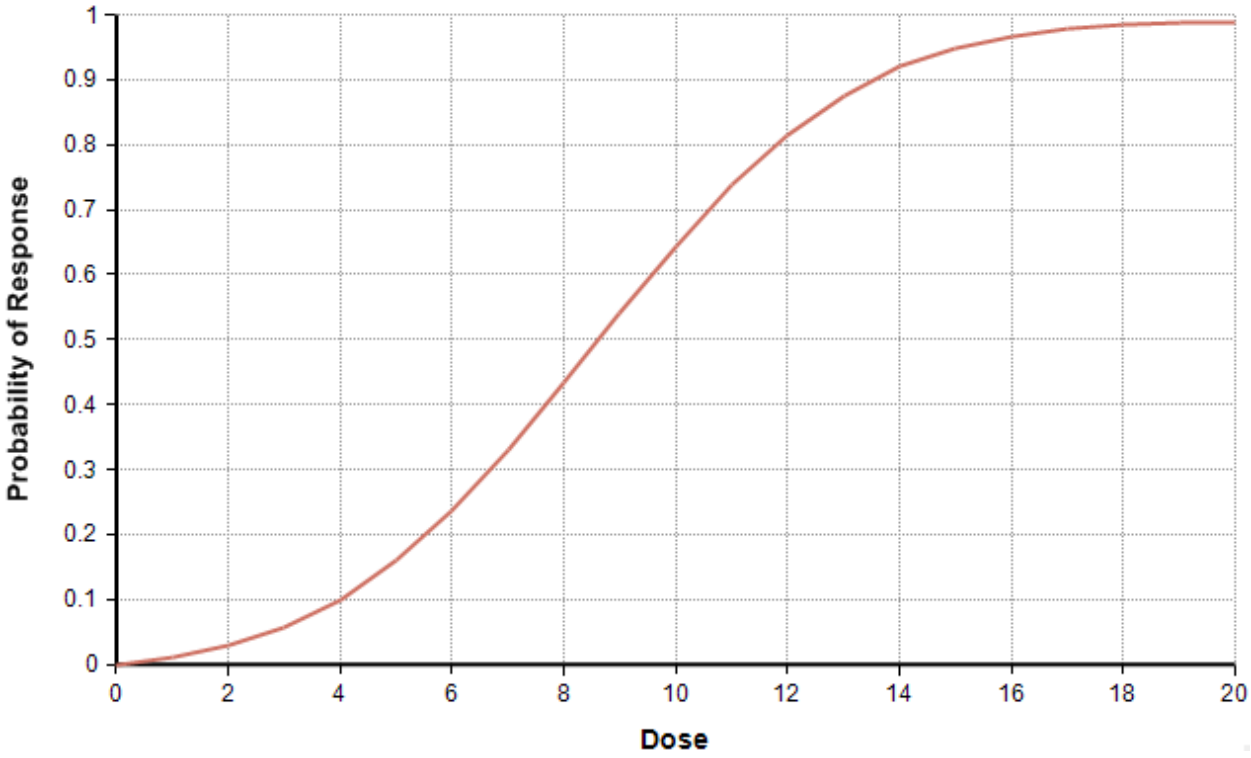


Figure 24. The Probit dose response relationship α is -2.3 and β is 0.27.

5.5.11 Restricted Log-Probit

The Log-Probit dose response relationship is a re-parameterization of the Cumulative Lognormal Distribution. The user specifies two parameters, *α and β:*

$$P(dose, \alpha, \beta) = (1-\gamma) \times \Phi(\alpha + \beta \times \ln(dose))$$

Equation 70

- where:Φ() is the Cumulative Distribution Function of the standard normal distribution (μ = 0, σ = 1).
- *α* is "intercept".
- *β* ≥ 1 is "slope" (USEPA, 2012).
- *0 < γ < 1* is the background probability of response (USEPA, 2012).

Note: FDA-iRISK does not reproduce the equation from the BMDS technical documentation (USEPA, 2012) exactly, since the desired response is the additive risk, not including the background risk. The background term may have been required to fit the BMDS data to some observational data. This ensures that the risk is zero when the dose is zero (i.e., since $\Phi(-\infty)=0$).

An example of a Log-Probit dose response model is shown in Figure 21.

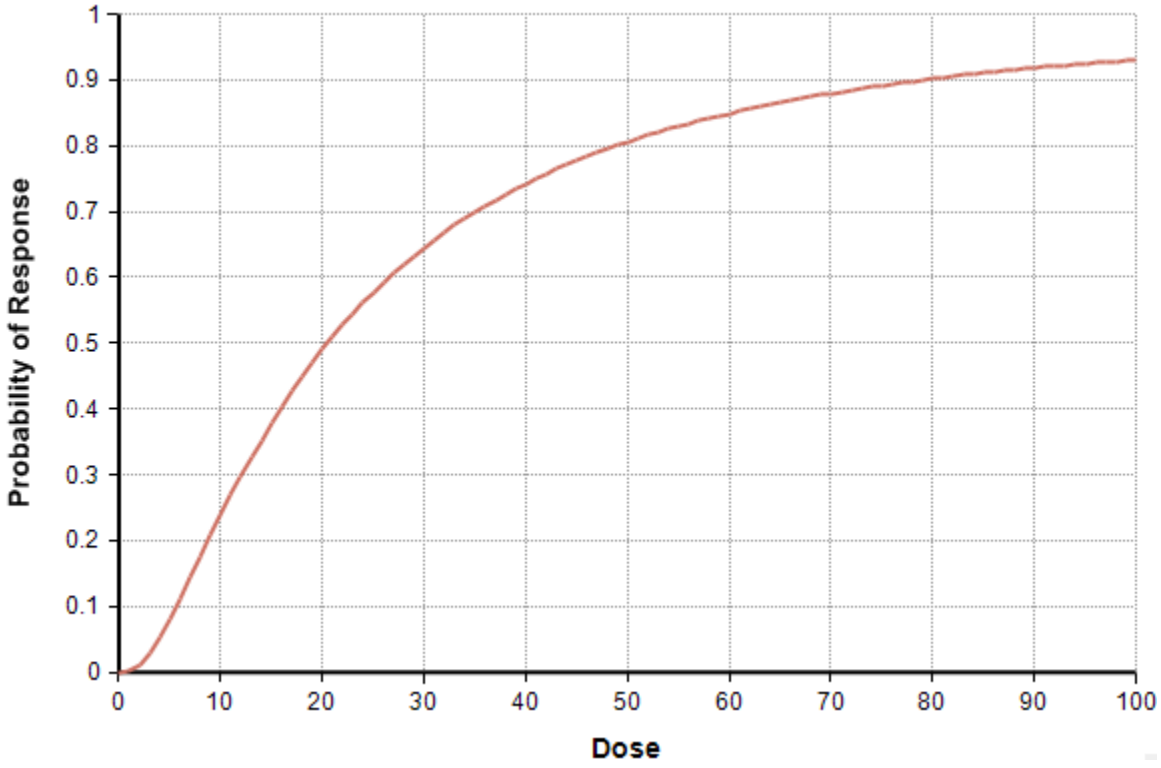


Figure 25. The Log-Probit dose response relationship where α is -3, β is 1, *γ* is 0.016.

5.5.12 Restricted Weibull

The Restricted Weibull dose response model includes a background probability of adverse outcome to the general Weibull dose response model. As with other forms in FDA-iRISK that include a background risk, FDA-iRISK removes the background risk from the formula:

$$P\ (dose.\alpha, \beta) = (1-\gamma) * (1-\exp(-\beta \times dose^{\alpha}))$$

Equation 71

where:

- α (power) ≥ 1 and β (slope) > 0 are parameters of the dose response model. α is restricted to ≥ 1 based on EPA (2012).
- *0 < γ < 1* is the background probability of response.

An example of a Restricted Weibull dose response model is shown in Figure 26.

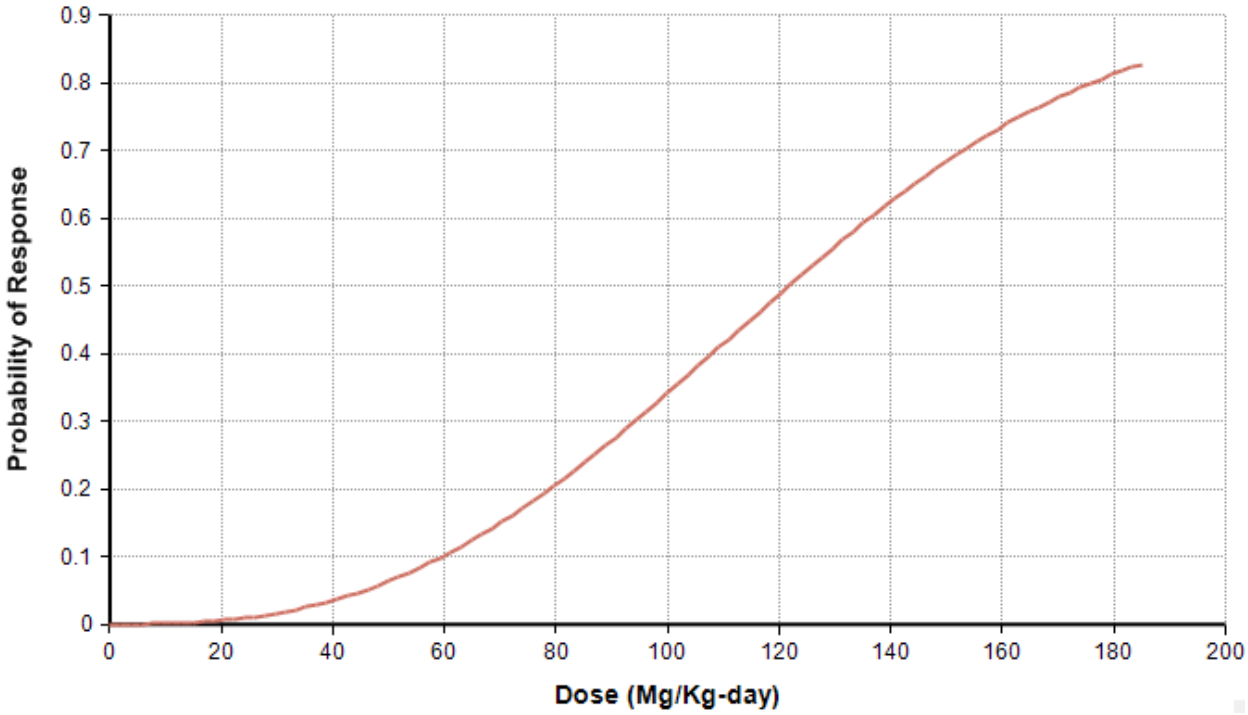


Figure 26. An example of a Restricted Weibull dose response relationship with β = 1.9E-6 and α = 2.7, with a background of 0.1

5.5.13 APROBA Lognormal

The APROBA Lognormal variation is based on the expectation that the user will use the APROBA spreadsheet tool (or variations which are available online) developed by the WHO/IPCS (WHO, 2018). The APROBA Tool provides the parameters necessary to specify the parameters of the cumulative lognormal distribution which serves as the dose-response function. It also provides parameters to describe the level of uncertainty in the parameters of the dose-response function. Due to the nature and the purpose of the APROBA tool, it is strongly recommended that the user use it in model runs that include uncertainty. An appropriate "variability" only parameterization based on the APROBA parameters has yet to be determined.

The lognormal model is given by:

$$P\ (dose, ED_{50}, GSD_H) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{ln(dose)} exp\left(\frac{-(t-\mu)^2}{2\sigma^2}\right) dt$$

Equation 72

where

- Where $\mu = \ln(ED_{50})$ and $\sigma = \ln(GSD_H)$.
- ED_{50} is the dose causing a 50% probability of response.

- GSD_H is the geometric standard deviation representing inter-individual variability among humans.

$\mu > 0$ and $\sigma > 0$; $ED_{50}>0$ and GSD>0.

The ED_{50} is provided by using the APROBA Tool, setting the Population Incidence user input to 50, and extracting the median (P_{50}) value for the uncertainty distribution of ED_{50}. The tool also provides the P_{95}/P_{50} ratio which can be used to establish the geometric standard deviation of the uncertainty distribution of the ED_{50} which is also, itself, lognormally distributed. The geometric standard deviation is calculated based on the P_{95} to P_{50} ratio representing 1.645 geometric standard deviations, so the GSD of the uncertainty distribution is:

$$GSD_{ED_{50}} = \left(\frac{P_{95}}{P_{50}}\right)^{\left(\frac{1}{1.645}\right)}$$

Equation 73

Because the FDA-iRISK model implements Monte Carlo simulation, it is not necessary to replicate the "APPROXIMATE" part of the APROBA model approach applied within the APROBA Excel tool. As such, the uncertainty distribution for GSD_H is such that the LOG(GSD_H) is itself log-normally distributed, rather than normally distributed as assumed in the approximate calculation. The WHO-IPCS Guidance document (Table 4.4 of WHO-IPCS, 2018) provides for a default uncertainty distribution for LOG(GSD_H) which is lognormal with a median value of 0.324, and a GSD of 1.5935. The user may override this with chemical-specific data based on toxicodynamic or toxicokinetic considerations that would modify the central estimate or increase or decrease the level of uncertainty in the level of human variability. For variability only runs, a default value of 0.4235 is used, which is the median value for the log-normal distribution applied in the approximation within the APROBA tool (see footnote to Table 4.5 in WHO-IPCS, 2018).

6 Positive-Only Binomial and Poisson Distributions

With the exception of a process model starting with zero initial concentration and prevalence, FDA-iRISK is structured to require that all units of food that are subject to the calculations in the process model are contaminated. FDA-iRISK uses the prevalence (weighting) value associated with each concentration value to account for units that are not contaminated.

For chemical hazards, this does not pose any special computational issues. However, for microbial hazards, this requires that each unit must have at least one cfu, pfu, or other specified count of a hazard. As such, FDA-iRISK uses two modified distribution functions to guarantee that the minimum value returned by the distribution is 1.

These functions are the called the Positive Binomial (pos_Binomial()) and the Positive Poisson (pos_Poisson()). These functions generate random numbers drawn from these two distributions but are conditional upon generating positive values. This is critical to efficient computation of risk, particularly where contamination becomes rare due to low concentrations in raw materials, or through reductions due to microbial inactivation. The purpose of the conditional random sampling is to avoid wasting significant computational effort in further simulating the fate of uncontaminated servings. The probability that the Binomial process or Poisson process being simulated will generate a value of zero, is taken into account by adjusting the corresponding estimate of prevalence.

7 Quantifying Uncertainty

In addition to specifying variability for FDA-iRISK model elements, users can also specify quantitative descriptions of uncertainty. This is achieved by specifying an uncertainty distribution for one or more model parameters on the variability dimension. These can be fixed (single value) parameters such as the initial prevalence of a process model, or parameters defined using a variability distribution such as an initial concentration defined as a Normal distribution.

When specifying uncertainty for a fixed parameter, the user assigns an uncertainty distribution directly to that parameter (e.g., a beta distribution for the point value of prevalence). When specifying uncertainty for a variability parameter or a dose-response model, the user must assign an uncertainty distribution to one or more of the distribution's or model's parameters (e.g., the mean of a Normal distribution or the beta value for a Beta-Poisson dose-response model). For dose-response models, users can specify uncertainty for each parameter independently, or define linked sets of uncertainty values. For microbial process models, users can also specify uncertainty separately for the initial prevalence, concentration and unit size, or define linked sets of uncertainty distributions.

When the risk scenario is simulated, FDA-iRISK adds an uncertainty dimension to the underlying Monte Carlo simulation. For each uncertainty loop, FDA-iRISK draws a random sample from each uncertainty distribution defined and assigns the values to the corresponding parameter in the model. FDA-iRISK will then simulate the variability dimension using these values and the variability distributions defined by the user.

Note that FDA-iRISK computes these uncertainty results on a per scenario basis and they cannot be aggregated across scenarios. For example, FDA-iRISK offers the user the option to group two or more scenarios together when generating a ranking report. If the scenarios do not include uncertainty, their results can be combined to produce an overall ranking. However, if they do include uncertainty, their results cannot be combined.

Uncertainty Example:

In this example, the initial prevalence of the process model is defined as uncertain with a uniform distribution of (0.1, 0.3). All other model parameters are fixed values or use only variability distributions.

For each uncertainty loop, FDA-iRISK will draw a random sample for the initial prevalence. The following table lists the first five such values:

Table 7_1. Uncertainty example: prevalence

Uncertainty Index	1	2	3	4	5
Initial Prevalence	0.15	0.11	0.21	0.27	0.19

FDA-iRISK will then execute the full Monte Carlo simulation of the model variability using each of these initial prevalence values.

If uncertainty distributions are defined for more than one parameter, FDA-iRISK will assign each distribution a value over the uncertainty index, not increase the number of uncertainty samples. For example, assume the amount of growth (log_{10}) in a process stage was assigned a uniform distribution of (3,5), then the following would result:

Table 7_2. Uncertainty example: prevalence and amount of growth

Uncertainty Index	1	2	3	4	5
Initial Prevalence	0.15	0.11	0.21	0.27	0.19
Amount of Growth	3.3	4.1	3.9	3.2	4.5

8 Evaluation of the Convergence for the Monte Carlo Simulation

Evaluating the convergence for the Monte Carlo simulation has three main purposes:

- To ensure the validity of the FDA-iRISK Monte Carlo simulations (strictly speaking FDA-iRISK uses Random Latin Hypercube Sampling). Monte Carlo simulation provides an approximation of statistical measures that improves as the number of iterations increases.
- To allow for adaptation of the number of iterations to the user-specified model, rather than attempting a one-size-fits-all number of iterations, which may be impossible to specify given the wide variation in potential applications of FDA-iRISK.
- To minimize the number of iterations while providing a specified level of convergence of selected outputs statistics.

To meet these purposes, FDA-iRISK implements a convergence analysis of each model simulated. This convergence analysis determines how many iterations of the Monte Carlo simulation are required. The user can configure many of the aspects of this convergence analysis, including the number of iterations use in each batch, the number of convergence tests required, and the maximum number of batches to run.

For a variability only model, the stability analysis uses the following algorithm:

- An endpoint is selected to test for stability. This endpoint is either the mean of the final risk measure (e.g., DALYs per year) resulting from the simulation or the mean of the exposure (e.g., cfu), depending on user preference. Note that for exposure-only scenarios, the exposure will always be used.
- Default settings are defined on the Simulation Settings tab, however the user should always consider if these are appropriate for their simulation and adjust as necessary. The default settings are not endorsed by FDA or RSI as suitable convergence settings for all applications of FDA-iRISK.
- FDA-iRISK executes an initial batch of 9000 iterations (default - configurable).
- FDA-iRISK executes a subsequent batch of 3000 iterations (default - configurable).
- FDA-iRISK tests the change in the selected metric between the batches. If the change in the running mean is less than a specified threshold (1%, default - configurable), that batch is flagged as having passed the convergence test.
- If any of the following conditions are met, the simulation ends:
    - If the total number of sequential passed tests equals the number of tests required (3, default - configurable), the model is considered to have converged and simulation ends.
    - If the total number of batches simulated is greater than the maximum allows (100, default - configurable), the simulation ends.

- Otherwise, the simulation continues and executes a new simulation batch
- If a test fails, the total number of sequential passed tests is reset to 0.
- If the simulation ends due to exceeding the batch limit, then the failure to converge is reported.

For models including uncertainty, an expanding algorithm is used:

- FDA-iRISK draws samples for each of the uncertainty distributions
- FDA-iRISK assigns these values to the corresponding parameters in the model
- FDA-iRISK executes the variability convergence algorithm outlined above for the current set of uncertainty values
- FDA-iRISK repeats the process for a batch of 100 uncertainty iterations (default - configurable)
- FDA-iRISK records the mean, median and interval (5^{th} to 95^{th} percentile, configurable) values of the endpoint selected for variability convergence. By default all 3 statistics are collected, this is configurable with the exception of the mean which is mandatory.
- If this is the first batch, FDA-iRISK repeats the process
- If this is the second or subsequent batch, FDA-RISK tests the change in the running values of the mean, median (optional) and interval (optional) against user-specified criteria (e.g., 5%, 5% and 10%). It also checks that all variability simulations in the batch converged.
- If the number of sequential required tests has passed, the simulation will stop. Otherwise, the process will repeat until the model converges or the maximum number of batches is exceeded.
- FDA-IRISK will report if the model has converged or not.

FDA-IRISK provides a convergence report for each simulation job that provides a summary of convergence testing, reporting results batch by batch.

9 References

Baranyi, J., and T. A. Roberts. 1994. A dynamic approach to predicting bacterial growth in food. International Journal of Food Microbiology 23:277-294.

Baranyi, J., and T. A. Roberts. 2000. Principles and application of predictive modelling of the effects of preservative factors on microorganisms. In B. M. Lund, T. C. Baird-Parker, and G. W. Gould (eds.) *The Microbiology Safety and Quality of Foods.* Ch. 18: 342-358. Ed. Aspen Publishers Inc., Gaithersburg, Maryland.

Baranyi, J. 2020. Personal communication with G. Paoli regarding replicating calculations of DM-Fit, by email on August 28, 2020.

Batz, M., S. Hoffman, and J. G. Morris Jr. 2014. Disease-outcome trees, EQ-5D scores, and estimated annual losses of quality-adjusted life years (QALYs) for 14 foodborne pathogens in the United States. Foodborne Pathogens and Disease 11:395-402. doi: 10.1089/fpd.2013.1658.

Bigelow, W. D. 1921. The logarithmic nature of thermal death time curves. Journal of Infectious Diseases 29:528-536.

Buchanan, R. L., R. C. Whiting, and W. C. Damert. 1997. When is simple good enough: a comparison of the Gompertz, Baranyi, and three-phase linear models for fitting bacterial growth curves. Food Microbiology. 14:313-326.

Cohen, J., P. Cohen, S. G. West, and L. S. Aiken. 2003. Applied multiple regression/correlation analysis for the behavioral sciences. 3rd ed. Lawrence Erlbaum Associates, Inc., Mahwah, New Jersey.

GBD 2019 Diseases and Injuries Collaborators. 2020. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. The Lancet. 396:1204-1222.

Haas, C. N., J. B. Rose, and C. P. Gerba. 1999. Quantitative Microbial Risk Assessment. John Wiley & Sons, Inc., New York, New York.

Haas, C. N. 2002. Conditional dose-response relationships for microorganisms: development and application. Risk Analysis 22:455-463.

Havelaar, A. H., J. A. Haagsma, M.-J. J. Mangen, J. M. Kemmeren, L. P. B. Verhoef, S. M. C. Vijgen, et al. 2012. Disease burden of foodborne pathogens in the Netherlands, 2009. International Journal of Food Microbiology 156:231–238.

ICMSF (International Commission on Microbiological Specifications for Foods). 2020. Microbiological sampling plans spreadsheet standard program 10 (v2.10): A tool to explore ICMSF recommendations. Available at: https://www.icmsf.org/wp-content/uploads/2020/10/standardprogram10.xlsm. Accessed March 17, 2023.

ILSI (International Life Sciences Institute). 2010. Impact of Microbial Distributions on Food Safety. ILSI Europe Report Series. ILSI Europe, Brussels.

Joe, H. 1989. Relative entropy measures of multivariate dependence. Journal of the American Statistical Association, 84:157-164.

Kaushik G., and S. N. Naik. 2009. Food processing a tool to pesticide residue dissipation – A review. Food Research International 42:26-40.

Lindqvist, R., T. Langerholc, J. Ranta, T. Hirvonen, and S. Sand. 2019. A common approach for ranking of microbiological and chemical hazards in foods based on risk assessment – useful but is it possible? Critical Reviews in Food Science and Nutrition 60:3461–3474. doi:10.1080/10408398.2019.1693957.

Lumina. 2023. Analytica documents: Correlate_With function. Available at: https://docs.analytica.com/index.php/Correlate_With. Accessed September 26, 2023.

McMeekin T. A. 1993a. Predictive Microbiology Theory and Application. Research Studies Press Ltd., Taunton, Somerset, England. See p. 94 (citing Ratkowsky et al., 1982).

McMeekin T. A., et al. 1993b. Predictive Microbiology Theory and Application p. 93 (citing Ratkowsky et al., (1982) Relationship between temperature and growth rate of bacterial cultures. Journal of Bacteriology 149:1-5.

McMeekin T. A. 1993c. Predictive Microbiology Theory and Application. Research Studies Press Ltd., Taunton, Somerset, England. See p. 168.

McMeekin T. A. 1993d. Predictive Microbiology Theory and Application. Research Studies Press Ltd. See p. 170.

McMeekin T. A. 1993e. Predictive Microbiology Theory and Application. Research Studies Press Ltd., Taunton, Somerset, England. See p. 128 (citing use by Zwietering et al. 1991).

Minor, T., A. Basher, K. Klontz, B. Brown, C. Nardinelli, and D. Zorn. 2015. The per case and total annual costs of foodborne illness in the United States. Risk Analysis 35:1125-1139. doi: 5 10.1111/risa.12316.

Nauta, M. J. 2002. Modelling bacterial growth in quantitative risk assessment: is it possible? International Journal of Food Microbiology 73:297-304.

Nauta, M. J. 2005. Risk assessment modelling of the food handling processes mixing and partitioning. International Journal of Food Microbiology 100:311-322.

Nauta, M. J. 2008. The modular process risk model (MPRM): a structured approach for food chain exposure assessment. In: D. W. Schaffner (ed.) *Microbial Risk Analysis of Foods*. pp. 99-136. ASM Press, Washington, D.C.

Paoli, G. M., E. Hartnett, and P. S. Price. 2025. Use of probabilistic exposure models in the assessment of dietary exposure to chemical. Food Additives & Contaminants: Part A, 42:7, 819-848. https://doi.org/10.1080/19440049.2025.2506104.

Pearson, K. 1895. Notes on regression and inheritance in the case of two parents. Proceedings of the Royal Society of London 58: 240-242.

Peleg, M., and M. B. Cole. 1998. Reinterpretation of microbial survival curves. Critical Reviews in Food Science and Nutrition 38:353-80.

Pouillot, R., M. Lubran, S. Cates, and S. Dennis. 2010. Estimating parametric distributions of storage time and temperature of ready-to-eat foods for US households. Journal of Food Protection 73:312-321.

Pouillot, R., Y. Chen, and K. Hoelzer. 2015. Modeling number of bacteria per food unit in comparison to bacterial concentration in quantitative risk assessment: Impact on risk estimates. Food Microbiology 45:245-253. http://dx.doi.org/10.1016/j.fm.2014.05.008.

Pouillot, R., Y. Chen, and J. M. Van Doren. 2024. Elucidating the influence of the lower and upper microbiological limits: when is a 3-class sampling plan useful to test for pathogens in food?. Food Control, 63, 110544. https://doi.org/10.1016/j.foodcont.2024.110544).

Ratkowsky, D. A., J. Olley, T. A. McMeekin, and A. Ball. 1982. Relationship between temperature and growth rate of bacterial cultures. Journal of Bacteriology 149:1-5.

Ross T. A., and T. A. McMeekin. 2003. Modeling microbial growth within food safety risk assessments. Risk Analysis 23:179-197.

Santillana Farakos, S.M., R. Pouillot, J. Spungen, B. Flannery, J.M. Van Doren, and S. Dennis. 2021. Implementing a risk-risk analysis framework to evaluate the impact of food intake shifts on risk of illness: a case study with infant cereal, Food Additives & Contaminants: Part A, 38:5, 718-730.

Spearman, C. 1904. The proof and measurement of association between two things. The American Journal of Psychology, 15, 72-101.

Teunis, P. F. M., and A. H. Havelaar. 2000. The beta Poisson dose-response model is not a single-hit model. Risk Analysis 20:513–520. doi: 10.1111/0272-4332.204048.

U. S. Environmental Protection Agency (EPA). 1992. Guidelines for Exposure Assessment. Published on May 29, 1992, Federal Register 57(104):22888-22938.

U. S. Environmental Protection Agency (EPA). 2011. Exposure Factors Handbook 2011. Available at: http://www.epa.gov/ncea/efh/pdfs/efh-complete.pdf. Accessed June 24, 2014.

U.S. Environmental Protection Agency (EPA). 2012. Benchmark Dose Technical Guidance. Available at: http://www.epa.gov/raf/publications/benchmarkdose.htm. Accessed February 2, 2014.

van Boekel, M. A. J. S. 2002. On the use of the Weibull model to describe thermal inactivation of microbial vegetative cells. International Journal of Food Microbiology 74:139–159.

van Leeuwen C. J., and T. G. Vermeire. 2007. Risk Assessment of Chemicals: An Introduction. Springer Science and Business Media. Dordrecht, The Netherlands.

World Health Organization International Programme on Chemical Safety (WHO/IPCS). 2018. Guidance Document on Evaluating and Expressing Uncertainty in Hazard Characterization, 2nd ed. World Health Organization, Geneva. Available at: https://www.who.int/publications/i/item/guidance-document-on-evaluating-and-expressing-uncertainty-in-hazard-characterization-2nd-ed. Accessed September 21, 2020.

Zwietering, M. H., J. C. de Wit, and S. Notermans. 1996. Application of predictive microbiology to estimate the number of *Bacillus cereus* in pasteurised milk at the point of consumption. International Journal of Food Microbiology 30:55-70.

Zwietering M. H., J. T. de Koos, B. E. Hasenack, J. C. de Witt, and K. van't Riet. 1991. Modeling of bacterial growth as a function of temperature. Applied and Environmental Microbiology 57:1094-1101.