DILIsym Services

A SIMULATIONS PLUS COMPANY

DILIsym User Training –

Using the DILIsym Optimization Feature

DILIsym Development Team

Goals for the Optimization Training Session

Participants should understand the following general concepts:

- Applications of the Optimization feature within DILIsym
- The critical components necessary for optimization within DILIsym
- Key concepts behind creating a fitness function
- Key concepts behind the genetic algorithm optimization routine
- The practical workflow for completing an optimization within DILIsym

DILIsym Optimization - General Introduction

- *Optimization is the process of making changes to DILIsym parameters so that one or more simulation outcomes better align with user-defined endpoints (optimization target data sets)*
- Optimization is an art form, with endless possibilities, so this is an introduction to the interface tool in general, but not a comprehensive review of all possible optimization setups
- Optimization of complex data sets and many parameters will take time to learn and time to simulate (iteration typically required)
- DILIsym (as of version 7A) contains an optimization interface, allowing users to set up and run complex optimizations without code manipulation

Some Applications of DILIsym Optimization

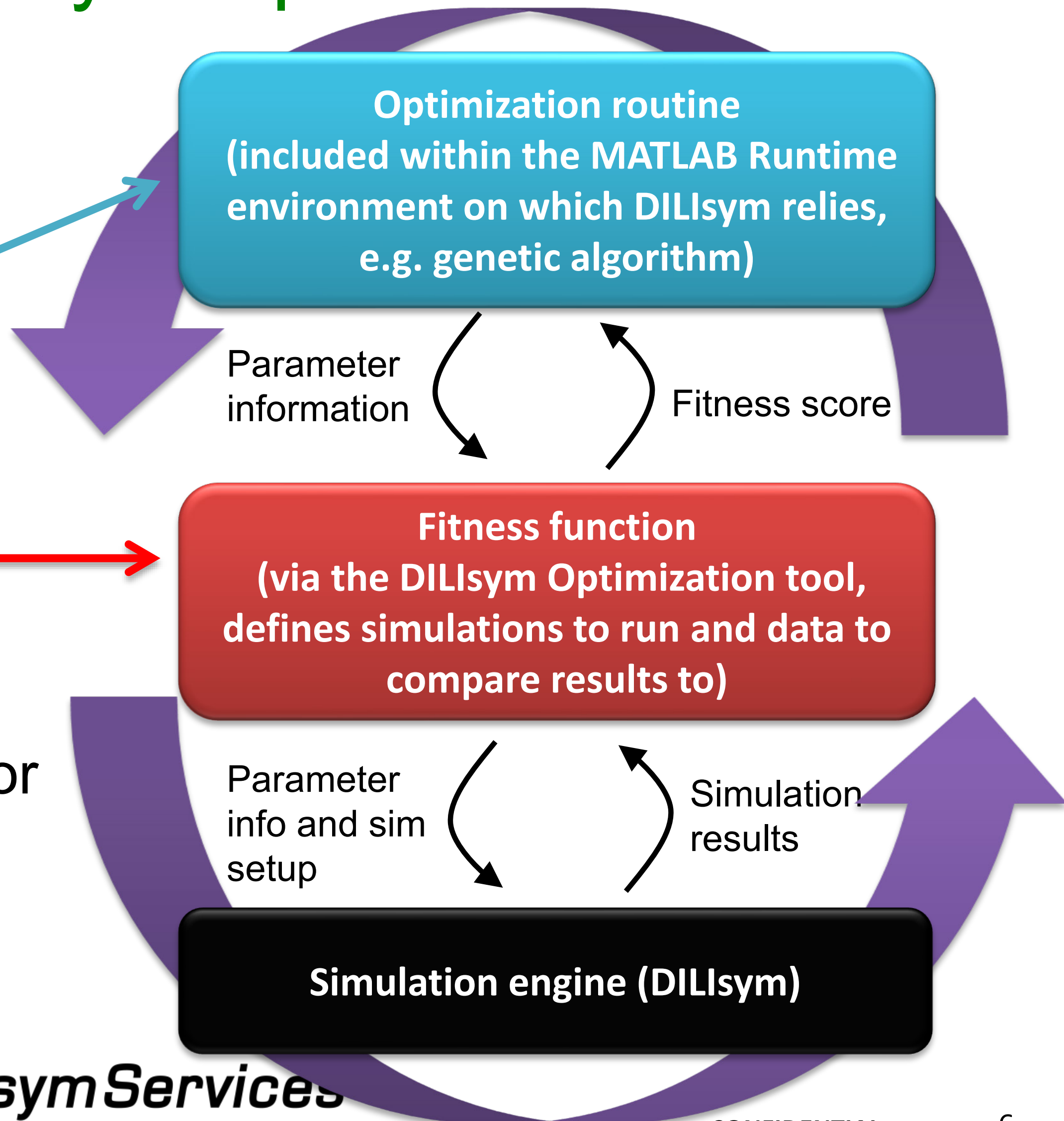
1. DILIsym toxicity parameter identification based on *in vitro* data
   - oxidative stress (or reactive oxygen species, ROS) production parameters
   - mitochondrial effects, such as V_{max} identification for saturable ETC inhibition pathways
2. Physiologically based pharmacokinetic (PBPK) model construction to fit user-defined data sets using the DILIsym PBPK model framework
   - *Note that when importing PK time courses from GastroPlus or other PBPK platforms to drive toxicity predictions, PBPK optimization for those models is done in the program of origin, not DILIsym*
3. Identification of new SimPops, or simulated humans or animals, to fit user-defined data sets
   - In this case, the user would typically be optimizing **Species** parameters related to fundamental processes, rather than **Drug** parameters related to exposure or toxicity pathways

DILIsym Optimization Tool – Critical Components

1. <u>SimSingles (simulation setups) within DILIsym via the home screen</u>
   – Parameter selections (dosing, species, time, etc.) are made from the primary DILIsym home screen
   – Simulations need to be set up for all scenarios pertaining to the optimization target data set(s)
2. <u>Constructing a fitness function using the DILIsym Optimization interface</u>
   – Goal is comparison(s) between a given simulation result and data
   – Tailored to fit the needs of each optimization by the user
   – Can be very complex or very simple
3. <u>Utilizing an optimization routine, the genetic algorithm, to conduct the comparisons and iterations toward best fits</u>
   – DILIsym version 7A only includes the genetic algorithm optimization routine option
   – Future versions may include different fitting algorithms, if desired by users

The Flow of Information Within a DILIsym Optimization

- The optimization routine begins the process with parameter information
- The fitness function is the intermediary between the optimization tool and the simulation engine
- The cycle runs iteratively for the *number of generations* defined by the user



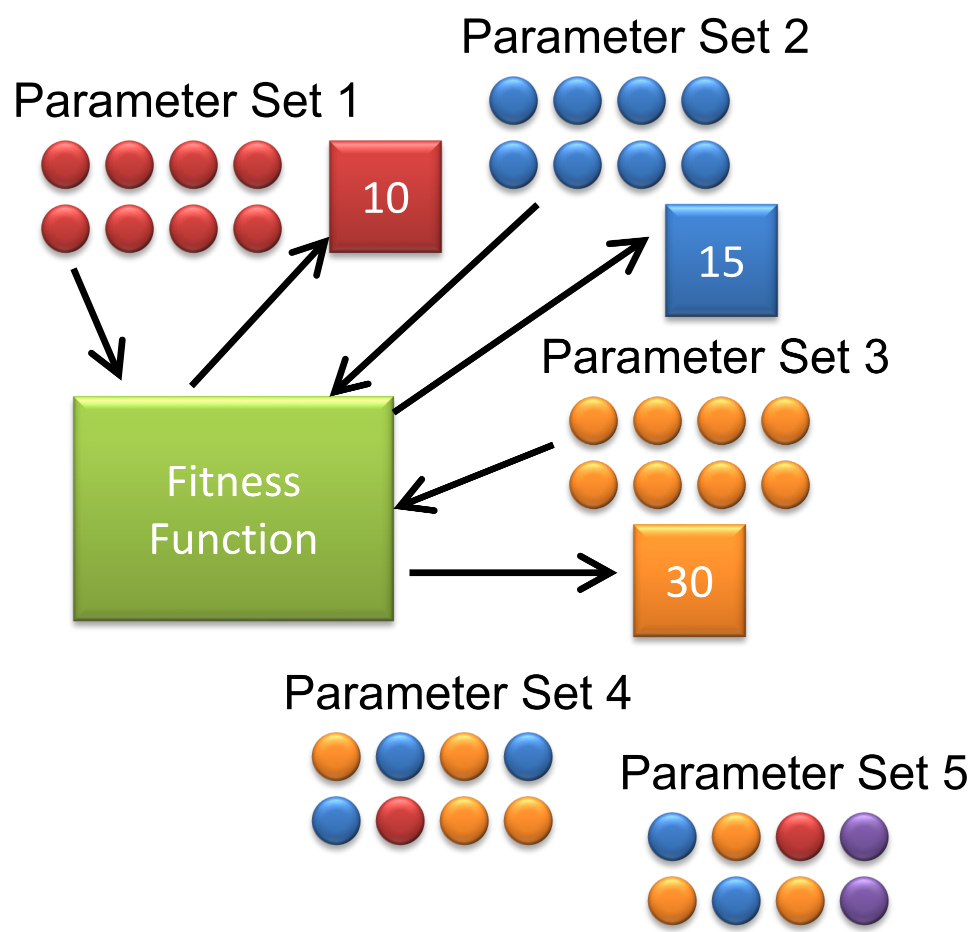
Genetic Algorithm Defined – MATLAB Documentation

- A genetic algorithm (GA) is a method for solving both constrained and unconstrained optimization problems based on a natural selection process that mimics biological evolution.
- The algorithm repeatedly modifies a population of individual solutions. At each step, the genetic algorithm randomly selects individuals from the current population and uses them as parents to produce the children for the next generation. Over successive generations, the population "evolves" toward an optimal solution.
- You can apply the genetic algorithm to solve problems that are not well suited for standard optimization algorithms, including problems in which the objective function is discontinuous, nondifferentiable, stochastic, or highly nonlinear.
- The genetic algorithm differs from a classical, derivative-based, optimization algorithm in two main ways, as summarized in the following table.

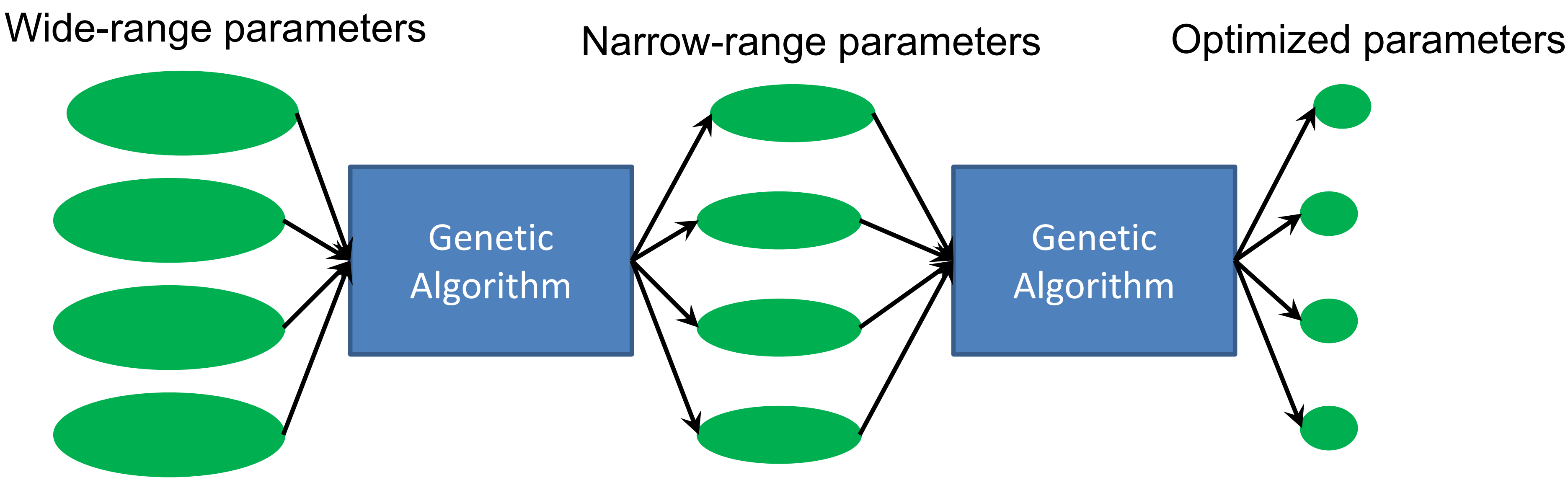
Classical Algorithm	Genetic Algorithm
Generates a single point at each iteration. The sequence of points approaches an optimal solution.	Generates a population of points at each iteration. The best point in the population approaches an optimal solution.
Selects the next point in the sequence by a deterministic computation.	Selects the next population by computation which uses random number generators.

The Genetic Algorithm Allows for Efficient Exploration of Parameter Space

- The genetic algorithm is an optimization method based on choosing the best parameter set fit to a specific fitness function
  - Best parameter sets are kept from generation to generation; other parameter sets created for each generation
- Genetic algorithms allow for the efficient exploration of a multi-dimensional parameter space
- DILIsym is utilizing the MATLAB global optimization toolbox for its genetic algorithm
- DILIsym contains nonlinear equations that can take time to solve, so the GA is a good fit



Choosing Parameter Ranges for the Genetic Algorithm



- There are many ways to choose parameter bounds and distributions
- Different consideration applied to “data-derived” vs “fitted” parameters
  - Prior information can be used to inform the optimized ranges of “data-derived” parameters (such as tissues’ PC, Vmax Km)
  - “fitted” parameters should be tested in wide-range to allows the optimization scheme to sweep out as much of the parameter space as possible in order to find the best fit
  - Parameter ranges can be narrowed iteratively if the first fit is not ideal

Fitness Functions Are Used During Optimization to Evaluate Relative Performance

- The fitness function is the tool used by the optimization routine to evaluate performance relative to a target
- The fitness function contains several critical components:
  - Information for one or more simulations
  - Data to be compared with simulation results, along with method for comparison
  - Parameter probability distribution scoring, if included
- Two primary ways of evaluating a given parameter solution within DILIsym
  - Difference between data point and simulation result, such as least squares fit, etc.
  - Combined probability of parameter values selected compared to assumed or known distributions, if applicable

Setting Up an Optimization in DILIsym – Step by Step Instructions

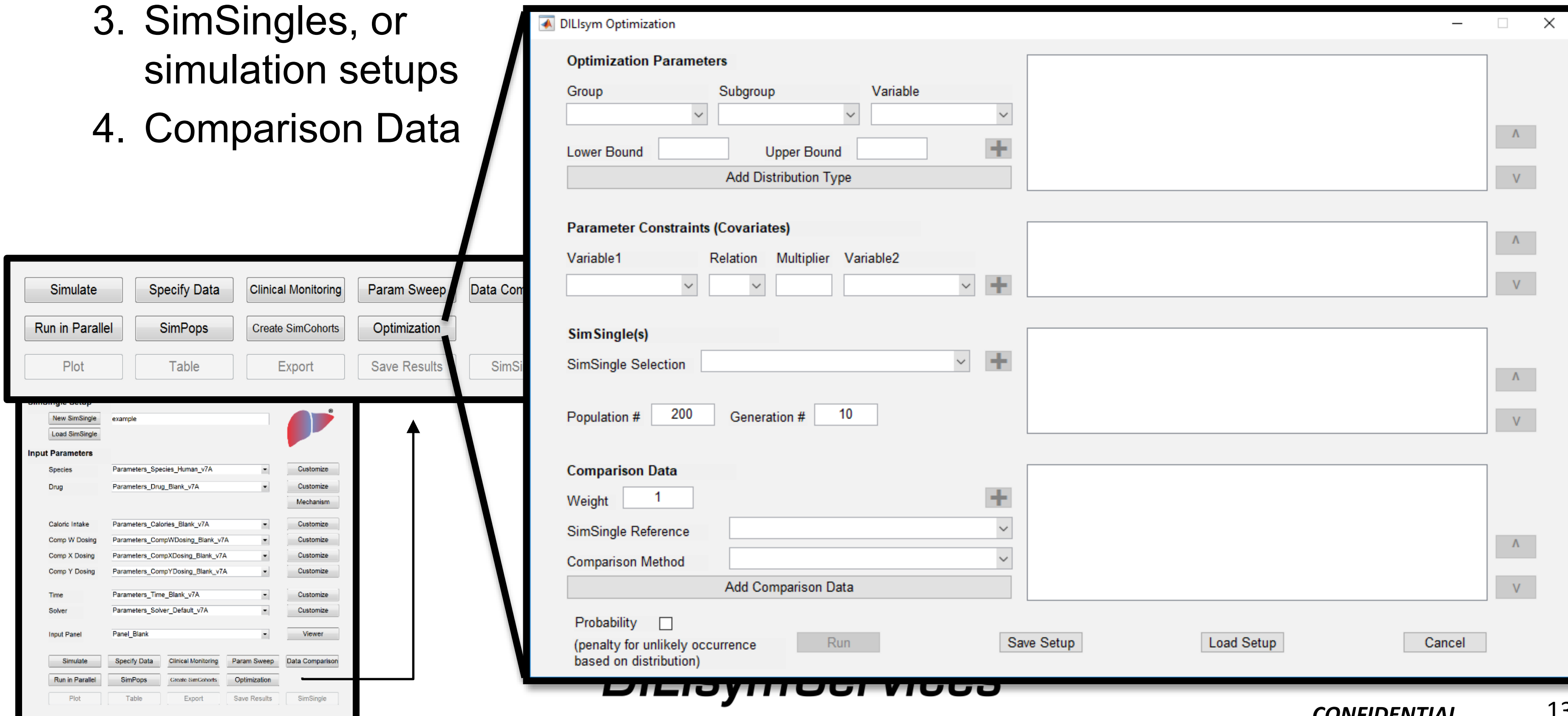
1. <u>Set up SimSingles (simulation setups) within DILIsym via the home screen</u>
   – Parameter selections (dosing, species, time, etc.) are made from the primary DILIsym home screen
   – Simulations need to be set up for each scenario from which each optimization target data set is provided
2. <u>Construct your fitness function using the DILIsym Optimization interface</u>
   – Goal is comparison(s) between a given simulation result and data
   – Tailored to fit the needs of each optimization
   – Can be very complex or very simple
3. <u>Run the optimization routine (which then utilizes the genetic algorithm in the background) to conduct the comparisons and iterations toward best fits</u>
   – DILIsym version 7A only includes the genetic algorithm optimization routine option
   – Future versions may include different fitting algorithms, if desired by users

Setting Up an Optimization in DILIsym – Step by Step Instructions

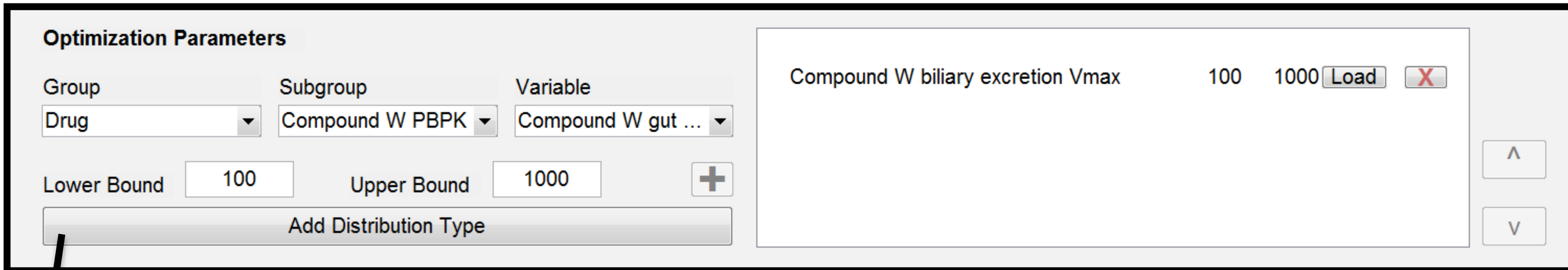
1. Set up SimSingles (simulation setups) within DILIsym via the home screen
   – Parameter selections (dosing, species, time, etc.) are made from the primary DILIsym home screen
   – Simulations need to be set up for each scenario from which each optimization target data set is provided
2. Construct your fitness function using the DILIsym Optimization interface
   – Goal is comparison(s) between a given simulation result and data
   – Tailored to fit the needs of each optimization
   – Can be very complex or very simple
3. Run the optimization routine (which then utilizes the genetic algorithm in the background) to conduct the comparisons and iterations toward best fits
   – DILIsym version 7A only includes the genetic algorithm optimization routine option
   – Future versions may include different fitting algorithms, if desired by users

The DILIsym Optimization Feature is Accessed from the DILIsym Home Screen

- The optimization window has four main sections:
  1. Optimization Parameters
  2. Parameter Constraints (Covariates)
  3. SimSingles, or simulation setups
  4. Comparison Data



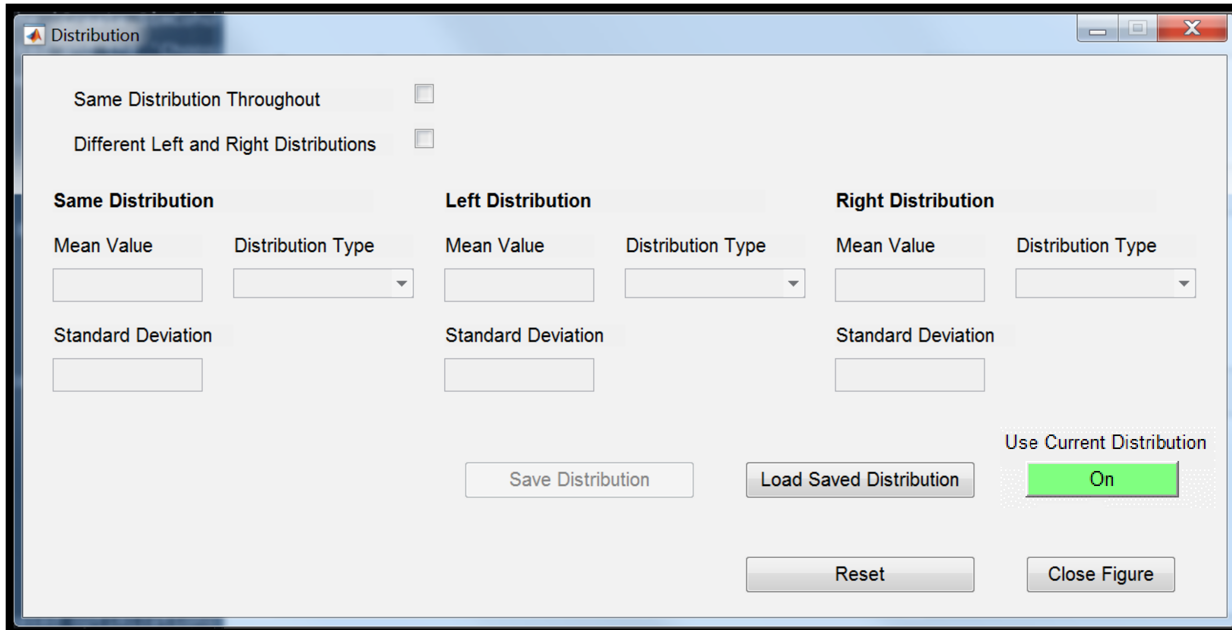
Add Your Intended Optimization Parameters to the Optimization Parameters Table



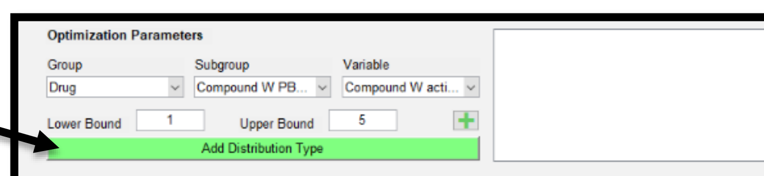
- The parameter name is input in the drop-down menus marked "Group", "Subgroup", and "Variable"
- Upper and lower bounds are input into the text box
- Note that dependent parameters, or covariates of other optimization parameters, need to be added as well
- The next step is adding the Distribution Type for the parameter, which must be done before the parameter can be added to the table of parameters being optimized

Set Up the Distribution Desired for Each Optimization Parameter

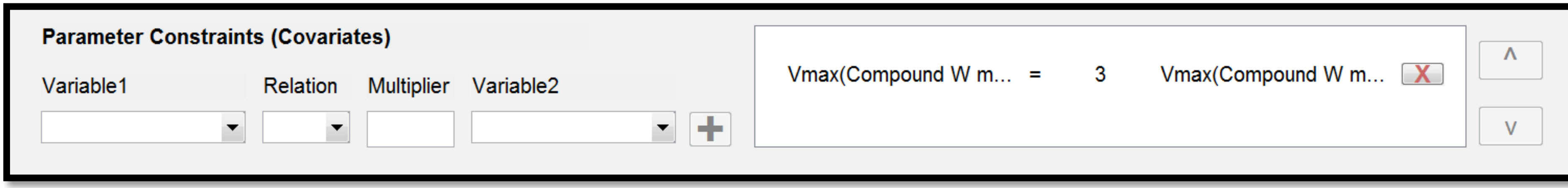
- The selected parameter value at each iteration will be selected from the distribution provided
- DILIsym includes normal, lognormal, Poisson, exponential, and uniform stochastic distributions
- Distribution can be symmetric or asymmetric
- Fill in the mean value, standard deviation (if applicable), distribution type



- Once a distribution has been created, it must be saved and then loaded
- Click on the “Load Saved Distribution” button, select a distribution, and load
- Click beneath the words “Use Current Distribution” to turn on or off loaded distribution
- Close the Distribution window and note that the “Add Distribution Type” button has now turned green, if turned on

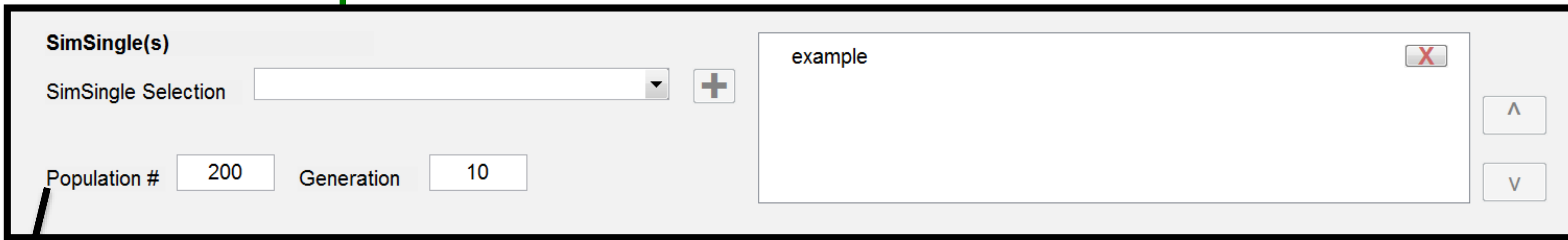


Parameter Constraints (Covariation) Can Be Added, if Desired



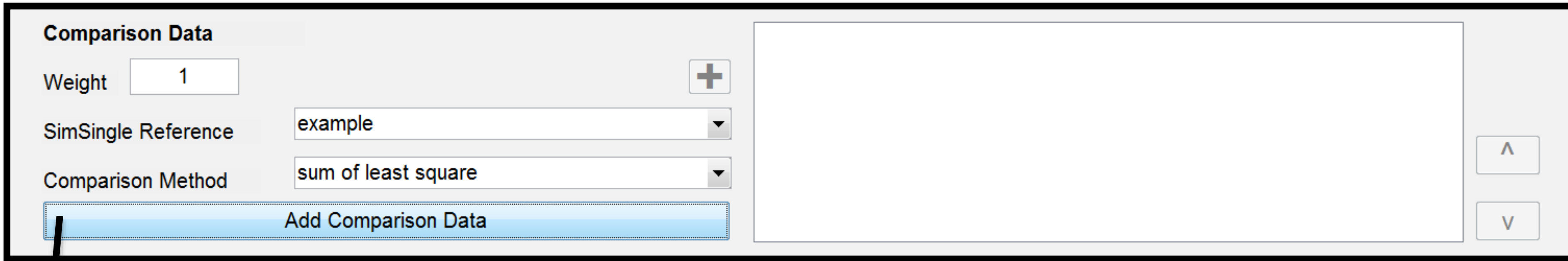
- Allows constraint of two parameters that should be covariates or otherwise related
- Choose the two parameters you wish to constrain with the two drop-down menus labeled “Variable 1” and “Variable 2”
- Choose a relationship (either =, <, or >) from the “Relation” drop-down menu
- The “Multiplier” input box allows constraint of a parameter to a multiple of another parameter’s value
- Click the green plus button next to the constraint inputs

Select the SimSingles to Include in the Optimization Process and the Population and Generation Sizes



- A SimSingle can be added to the optimization by selecting a SimSingle from the drop-down menu and clicking the green plus next to the drop-down menu
- The Population and Generation options are located beneath the SimSingles Selection menu
  - The population size represents the number of parameter sets, or simulations, conducted within each generation
    - <u>Population size is important, as larger populations will provide better coverage of the solution space</u>
  - The number of generations represents the number of times that each population will be generated and simulated after the initial condition simulations are complete (total generations is generations requested by the user plus one (e.g., the setup shown above would be 11 generations total))
    - <u>Generation number is important, as a sufficient number of generations is critical for the optimization routine to successfully progress</u>
  - ***We recommend at least 5-10 generations, with a population size of roughly 20-200, depending on the simulation time required***
  - ***Note that depending on the application, more or less individuals and generations may be needed***

Use the Comparison Data Section to Add the Data for Comparison to Simulation Results

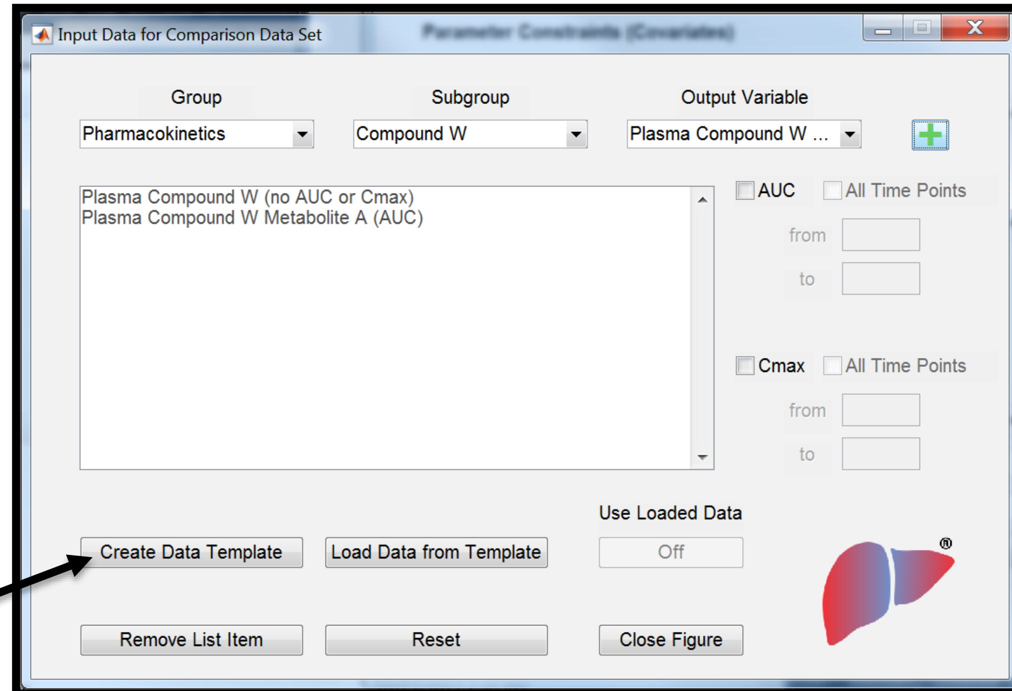


<u>Each comparison data set has four components:</u>

1. Weight
   - The weight given to a particular data set can be modulated by changing the input in the "Weight" input box
   - Each data set is normalized to the data results so that the inclusion of results that are in different orders of magnitude will not skew the comparison
2. SimSingle Reference, which must be chosen from the SimSingles added above
3. Comparison method
   - User can select absolute difference, least-squares, or logarithmic absolute comparison method for each data set
4. Data itself – click "Add Comparison Data" to add the data

Data Is Added by Clicking on the "Add Comparison Data" Button Which Opens the Template Generation Tool

- Using the drop down menus, add the DILIsym output that is analogous to the data to be added for comparison
- For AUC or C_{max} entries, the AUC and/or C_{max} check boxes should be highlighted and bounds for the AUC/C_{max} input into the appropriate input box
- For time-course data, leave both AUC and C_{max} check boxes empty
- Once variables are added to the window, click "Create Data Template"
- Note that a data template will be required for each SimSingle desired for comparison
    - e.g., if 4 SimSingles are desired with 4 unique simulation scenarios, at least 4 data templates will be required, each one corresponding to a particular SimSingle



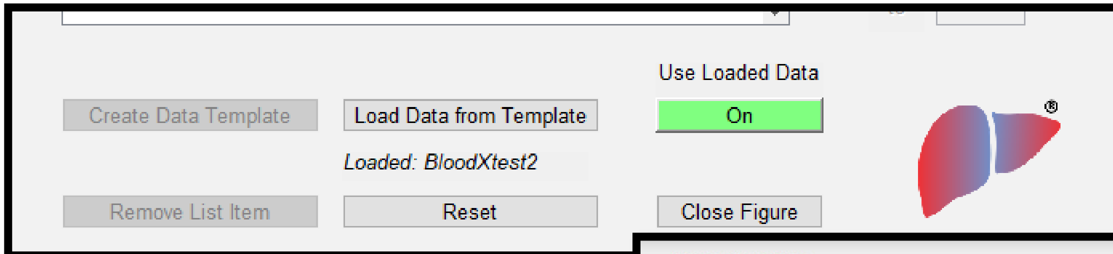
Add Your Data to the Template via the Time Course Tab (1st) or the AUC and/or C_{max} Tabs (next tabs)

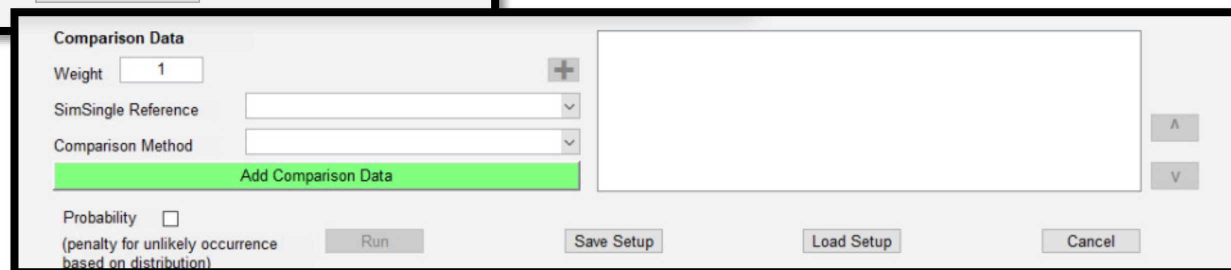
- Input the time vector into the column labeled "Time" and the time course data into their respective columns
- The AUC and C_{max} sheets are similar; input the AUC or C_{max} beneath the appropriate label
  - Located on separate tabs
- ***Do not*** create a template with both time course information and AUC / C_{max} information for the same output
- ***Do not alter*** the other areas of the template, as the optimization tool will use the template information

	A	B	C
1	Data Set to Compare	Data Set to Compare	
2	time	plasma_compound_W	
3	Time (hour)	Plasma Compound W (ug/mL)	
4	0	0	
5	1	0.05	
6	2	0.21	
7	3	0.14	
8	4	0.054	
9	5	0.043	
10	6	0.027	
11	8	0.017	
12	12	0.0064	
13	24	0.0048	

	A	B
1	AUC to Compare	AUC to Compare
2		plasma_CompW_MetA
3		Plasma Compound W Metabolite A (ug/mL)
4	AUC Value	45000
5	Time Range	All Time Points
6		
7		

Click the "Use Loaded Data" Button and Close the Template Generator Tool to See the "Add Comparison Data" Button Turn Green



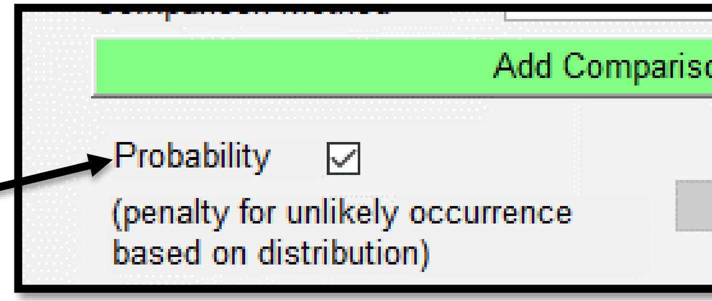


<u>Once a template has been fully completed with data:</u>

1. Click "Load Data from Template" to load in a fully specified data template
   – User should see "*Loaded: Template_Name*" beneath the button
2. Click the "Use Loaded Data" button (it should turn green)
3. Close the template generator tool
4. The Comparison Data section should now show the "Add Comparison Data" button as green and the comparison data set can be added to the list with the plus sign
5. Repeat for all desired comparison data sets

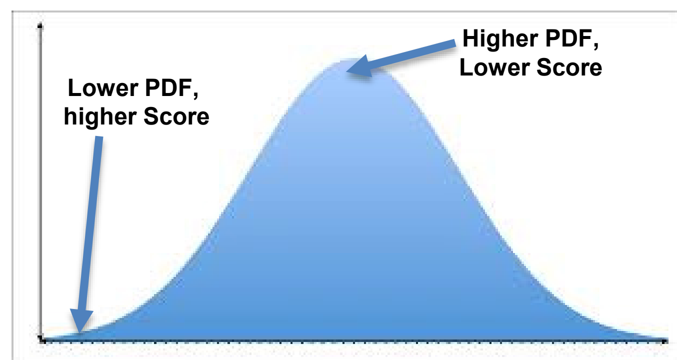
An Optional Penalty Can Be Assessed for the Likelihood of Parameter Value Occurrence Based on the Provided Distributions

- The selection of parameter values as the optimization routine progresses will depend on the fitness scores but distributions provided will initially bias the selection towards more likely values
- Checking the Probability box causes the fitness score to include an additional component not based on simulation comparison to data, but based on the likelihood of that value truly occurring based on the distribution
- Normalized to number of data points and units
- Formula used for probability component shown at right
  - *PDF* = probability density function value, computed for each parameter for each value of the parameter selected compared to the provided distribution, always between 0 and 1
  - *n* = number of parameters included in optimization
  - *Data Score Component* = value from comparison between simulation outcomes and data provided based on comparison method selected by user
    - Score is equal to Data Score Component alone if Probability box unchecked



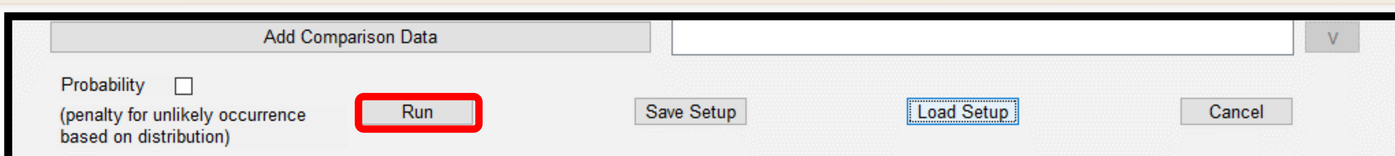
$$P = \prod_{i=1}^{n} PDF_i$$

$$Score = \left(\frac{1}{P}\right)^{\frac{1}{n}} (Data\ Score\ Component)$$



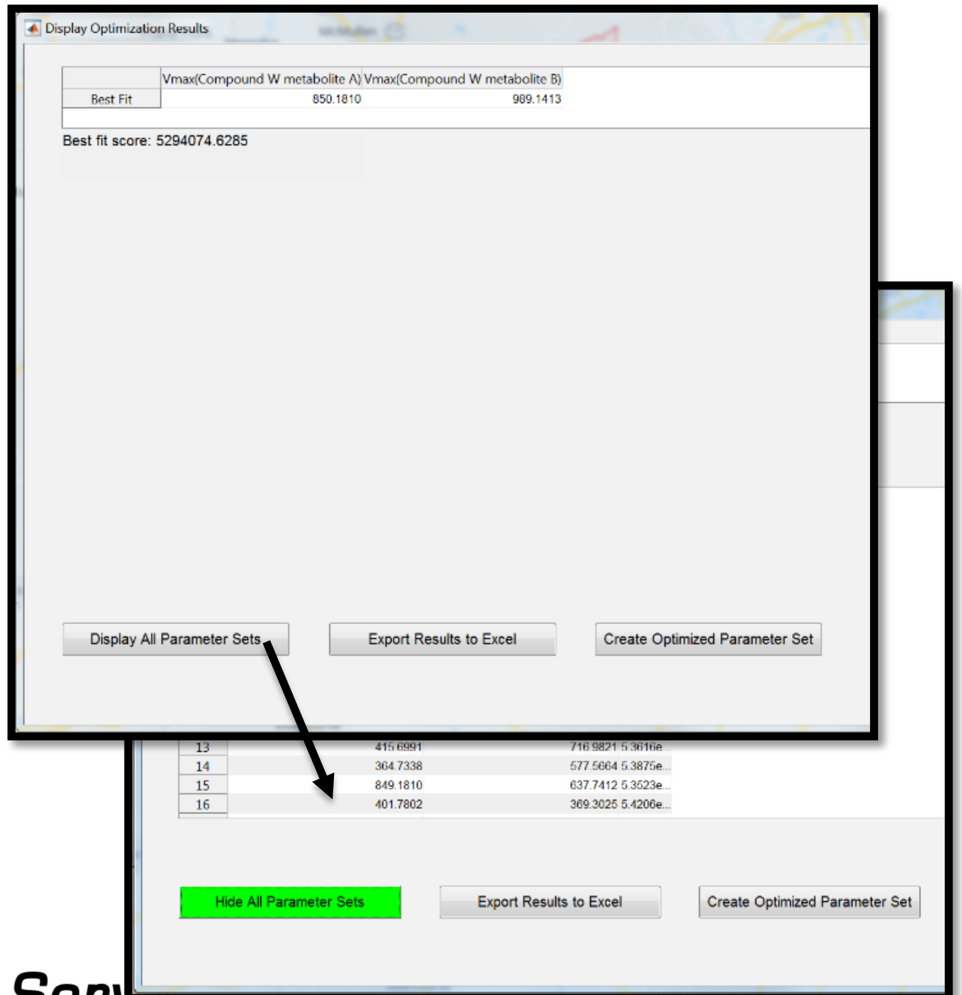
Setting Up an Optimization in DILIsym – Step by Step Instructions

1. <u>Set up SimSingles (simulation setups) within DILIsym via the home screen</u>
   - Parameter selections (dosing, species, time, etc.) are made from the primary DILIsym home screen
   - Simulations need to be set up for each scenario from which each optimization target data set is provided
2. <u>Construct your fitness function using the DILIsym Optimization interface</u>
   - Goal is comparison(s) between a given simulation result and data
   - Tailored to fit the needs of each optimization
   - Can be very complex or very simple
3. <u>Run the optimization routine (which then utilizes the genetic algorithm in the background) to conduct the comparisons and iterations toward best fits</u>
   - DILIsym version 7A only includes the genetic algorithm optimization routine option
   - Future versions may include different fitting algorithms, if desired by users



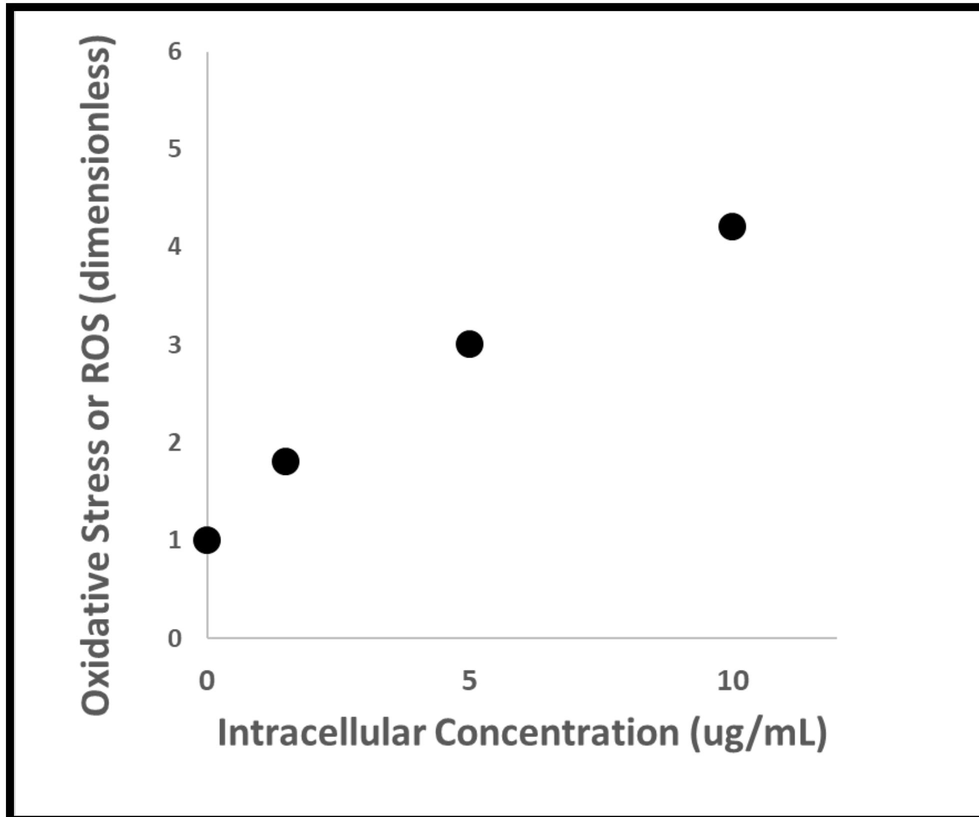
The Optimization Results Display Includes Parameter Values and Fitness Function Scores

- Optimization Results window initially shows two things
  1. the best-fit parameter set
  2. fitness function score for best-fit parameter set
- "Display All Parameter Sets" displays a table with all parameter sets run during optimization and all fitness function scores
- The "Export Results to Excel" button creates Excel spreadsheet containing
  1. parameter sets and their scores
  2. separate sheet with just the best-fit parameter set
  3. third sheet that contains the optimization settings
- "Create Optimized Parameter Set" button allows the user to create a parameter set using the parameters from the overall best fit
  – User selects an existing parameter set to modify; this is the parameter set into which the parameters will be placed (overwritten to) for the new parameter set
- Previously saved optimization results can be loaded from the "Results" drop-down menu on the main DILIsym screen



Hands-on Optimization Example – Optimizing to an Oxidative Stress Data Set

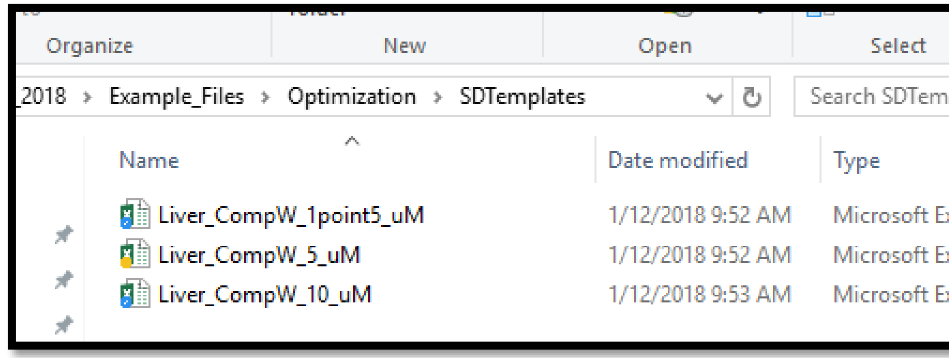
- Goal is to determine an RNS/ROS parameter value that will reproduce exposure-response shown at right
- Assume data was generated for human – will use human species for simulations
- Data was collected after 6 hours of incubation at concentrations leading to intracellular concentrations shown
- Example is simplified for the purposes of illustration

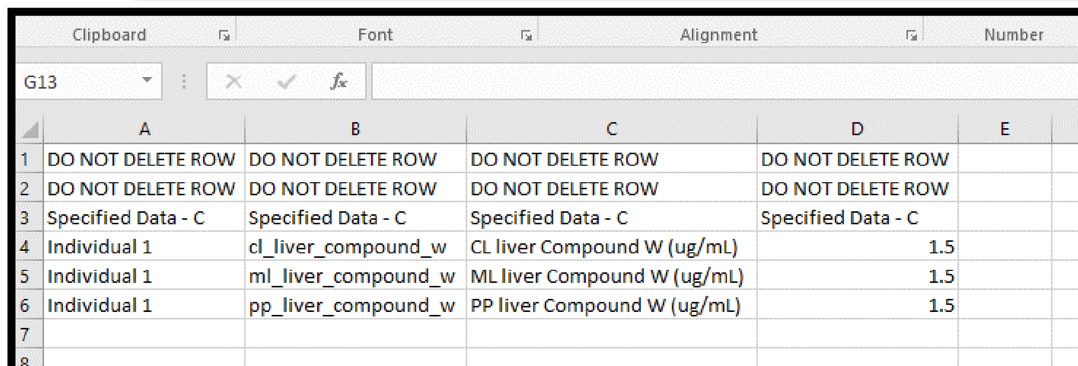


Theoretical Preclinical Data

Hands-on Optimization Example – *Step 1 – Review Specified Data Excel Files Provided to Understand Exposure Setup*

- Open Specified Data templates provided
  - **Liver_CompW_1point5_uM**
  - **Liver_CompW_5_uM**
  - **Liver_CompW_10_uM**
- Each template sets all 3 zones of liver to constant intracellular concentration value of 1.5, 5, or 10 ug/mL
- *In vitro* like Drug parameter file could also be used for this purpose
  - Both methods accomplish rapid steady state in liver tissue

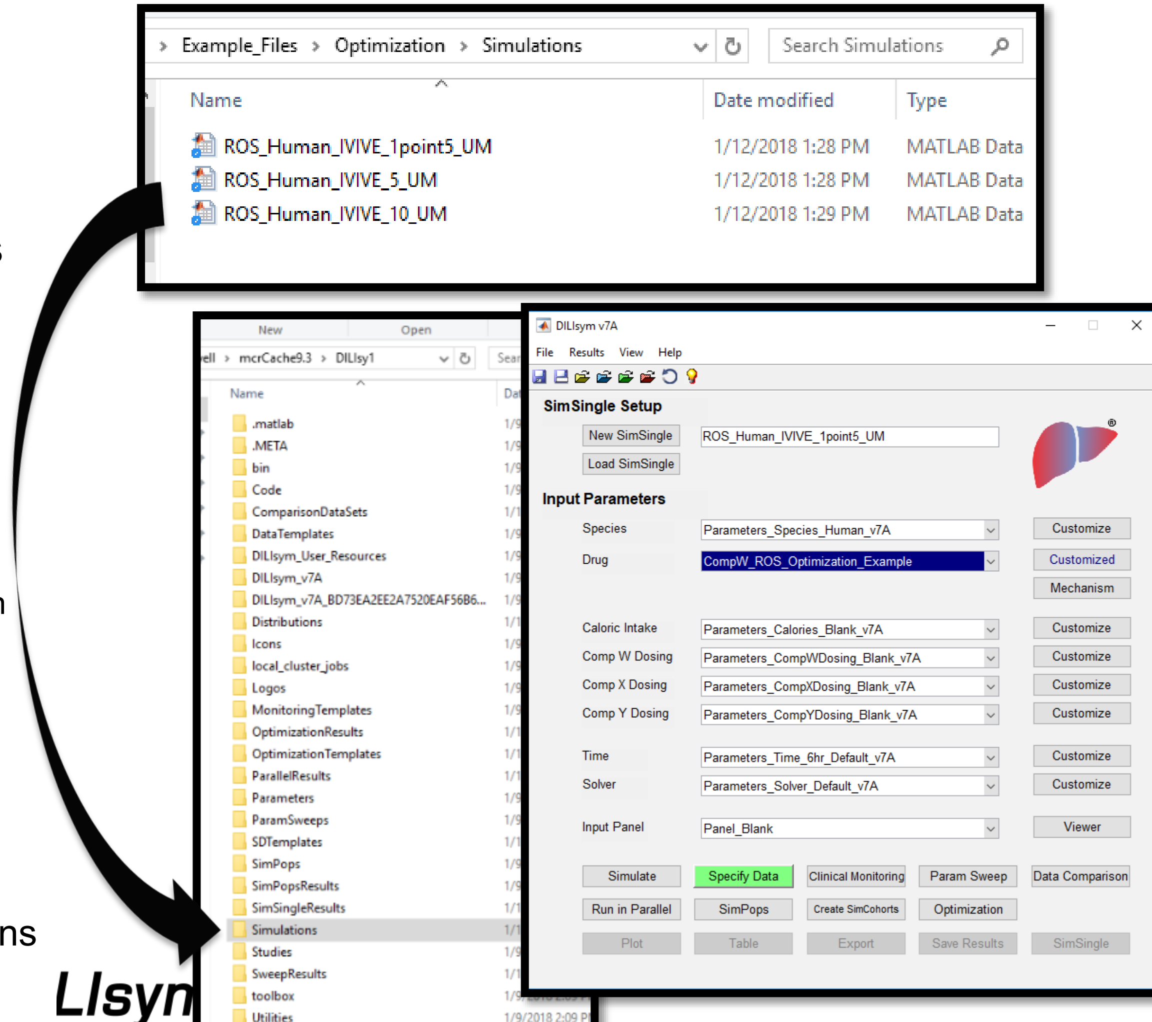




	A	B	C	D	E
1	DO NOT DELETE ROW	DO NOT DELETE ROW	DO NOT DELETE ROW	DO NOT DELETE ROW	
2	DO NOT DELETE ROW	DO NOT DELETE ROW	DO NOT DELETE ROW	DO NOT DELETE ROW	
3	Specified Data - C	Specified Data - C	Specified Data - C	Specified Data - C	
4	Individual 1	cl_liver_compound_w	CL liver Compound W (ug/mL)	1.5	
5	Individual 1	ml_liver_compound_w	ML liver Compound W (ug/mL)	1.5	
6	Individual 1	pp_liver_compound_w	PP liver Compound W (ug/mL)	1.5	
7					
8					

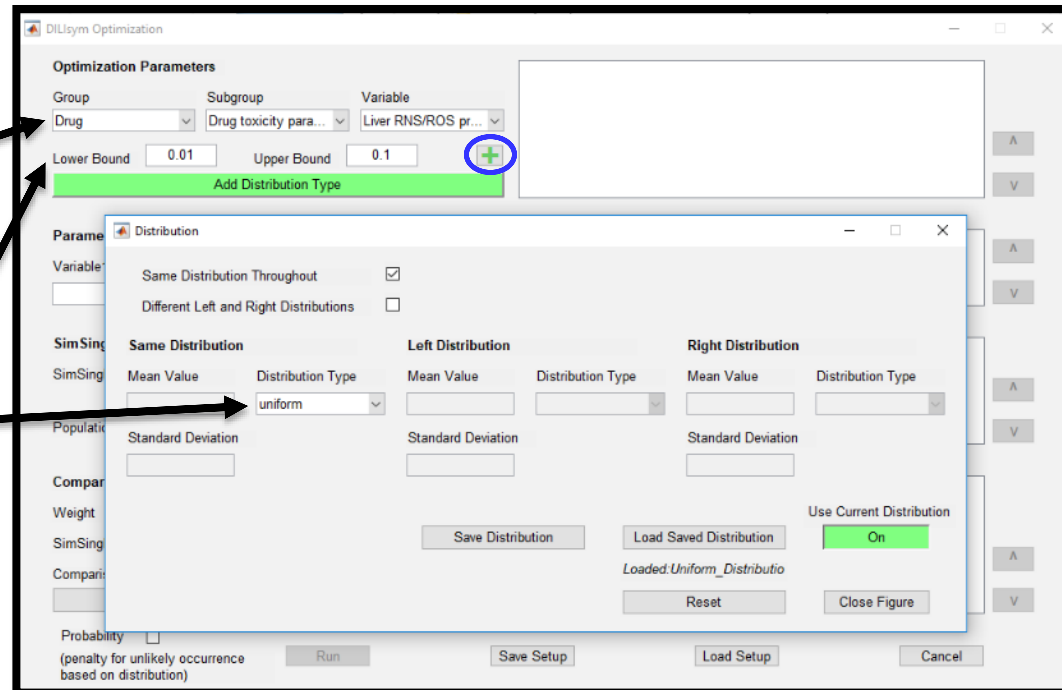
Hands-on Optimization Example – *Step 2 – Place Provided SimSingle Setups in Simulations Folder and Review SimSingles*

- Find Simulations directory by clicking any load option within DILIsym and copying location from Windows Explorer
- Copy three provided SimSingles into your Simulations directory
- Explore SimSingles
  - Human species selected
  - Drug parameter file:
    - molecular weight of 300 g/mol
    - RNS/ROS pathway 1 turned on
  - No meals (to save simulation time)
  - No drug dosing
  - 6 hour simulation time
  - Specified data used to set Compound W liver concentrations



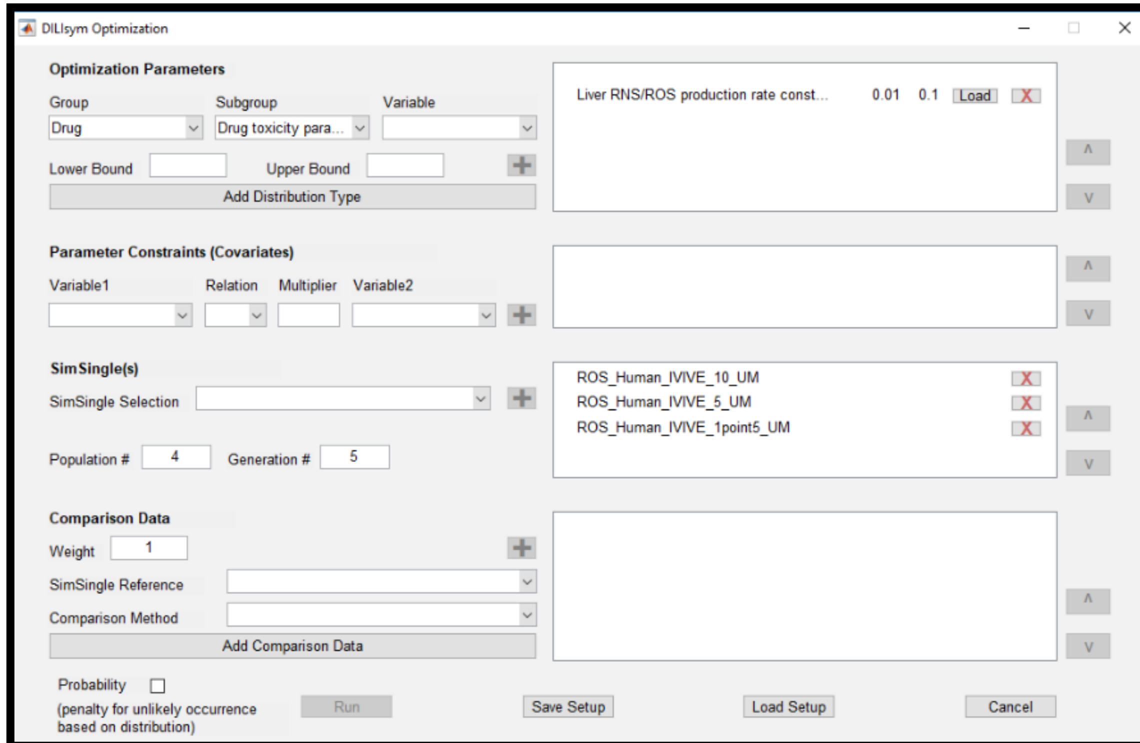
Hands-on Optimization Example – *Step 3 – Add "Liver RNS/ROS Production Rate Constant 1" Parameter as Optimization Parameter*

- Open Optimization interface from DILIsym home screen
  - Select Drug -> Drug toxicity parameters -> ***Liver RNS/ROS production rate constant 1*** from drop down menus
  - Bounds: 0.01 to 0.1
  - Distribution: uniform
    - Same distribution throughout
    - Save distribution
    - Load distribution
    - Turn "Use Current Distribution" button on so it's green
    - Close Figure
  - Add parameter to table with green "+" sign



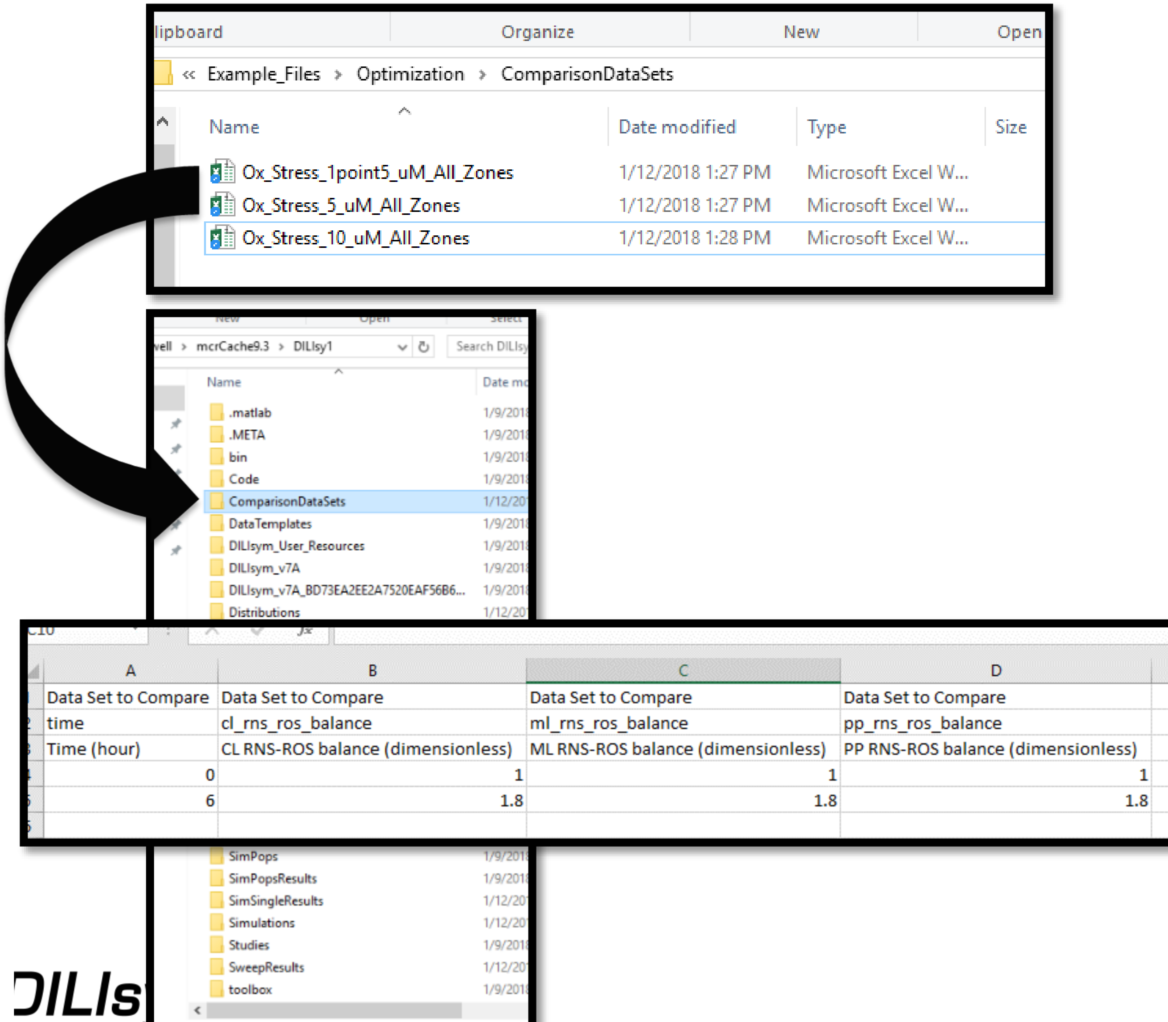
Hands-on Optimization Example – *Step 4 – Add Three Provided SimSingles*

- No parameter constraints (covariates) included in example
- Select each of the SimSingles and add them to table using green “+” sign
  - **ROS_Human_IVIVE_10_UM**
  - **ROS_Human_IVIVE_5_UM**
  - **ROS_Human_IVIVE_1point5_UM**
- Set *Population* size to **4**
- Set *Generation* size to **5**



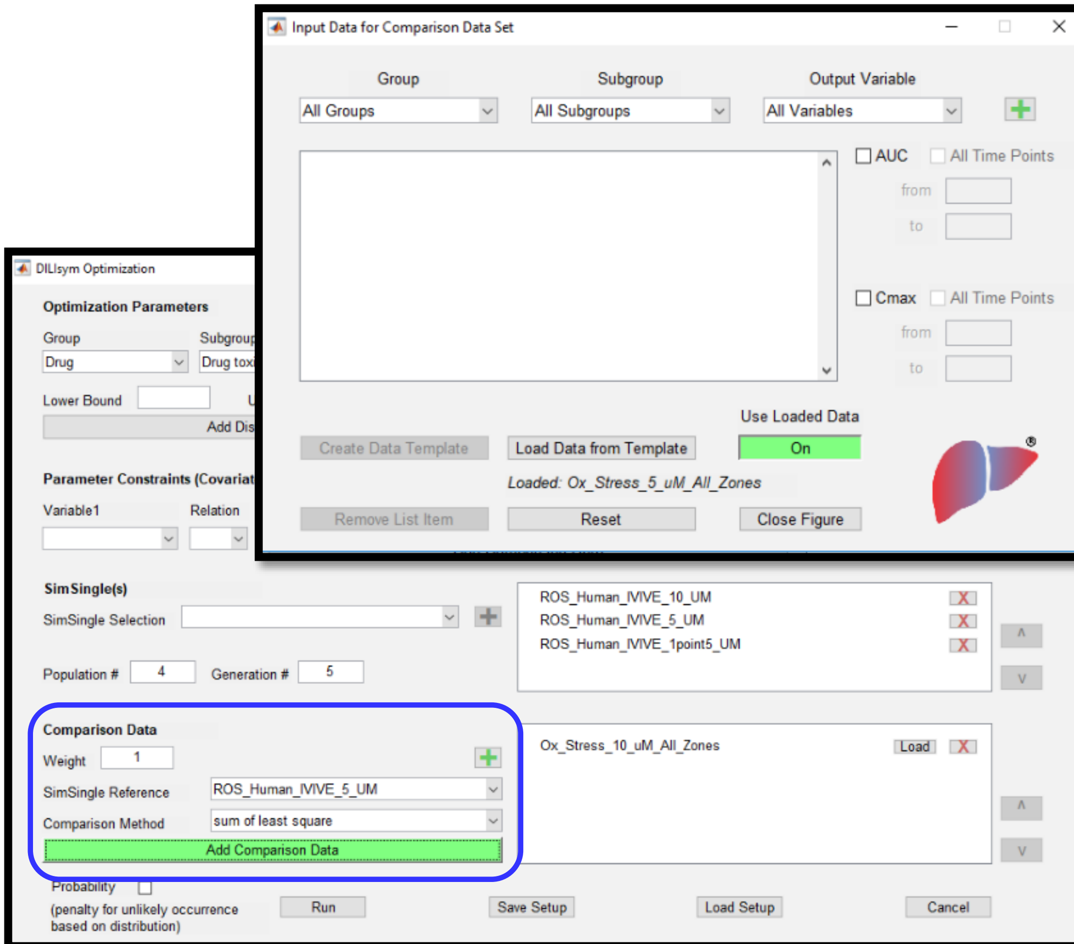
Hands-on Optimization Example – *Step 5 – Place Provided Comparison Data Excel Files in ComparisonDataSets Folder and Review Data*

- Find ComparisonDataSets directory by clicking any load option within DILIsym and copying location from Windows Explorer
- Copy three provided Excel files into your ComparisonDataSets directory
- Explore Excel files
  - Each file has ROS value at 0 and 6 hours for each zone of the liver at the three respective intracellular concentrations
    - 1.5 uM = 1.8 fold change ROS
    - 5 uM = 3 fold change ROS
    - 10 uM = 4.2 fold change ROS



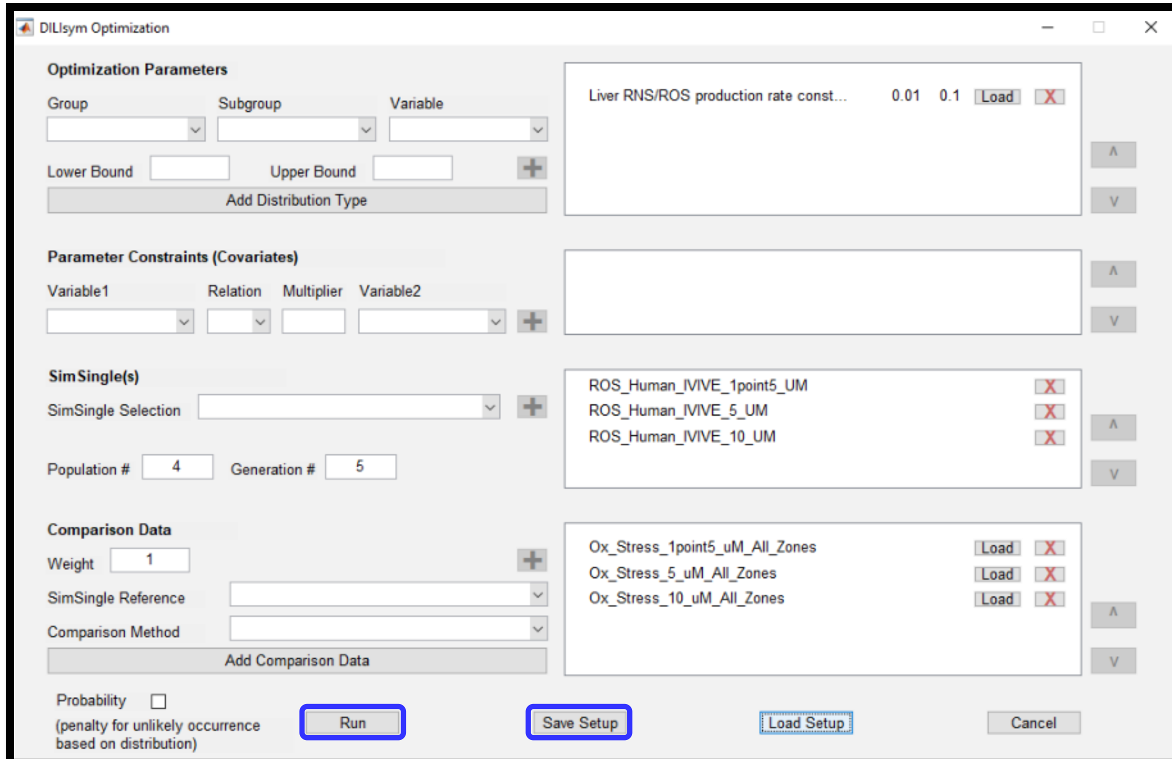
Hands-on Optimization Example – *Step 6 – Add 3 Comparison Data Sets*

- Leave Weight at 1
- Select each SimSingle reference for each exposure level
- Select "sum of least square" method for method
- Click "Add Comparison Data"
  - Load data for the SimSingle selected using the "Load Data from Template" button
  - ***Be sure template selected for loading matches the exposure level in SimSingle selected***
  - Click "Use Loaded Data" button (should turn green)
  - Close window
- Click green "**+**" sign
- Repeat for other two SimSingles
  - When comparison data window opens for the second time, the data from the first run will still be in the window
  - Load in new data set



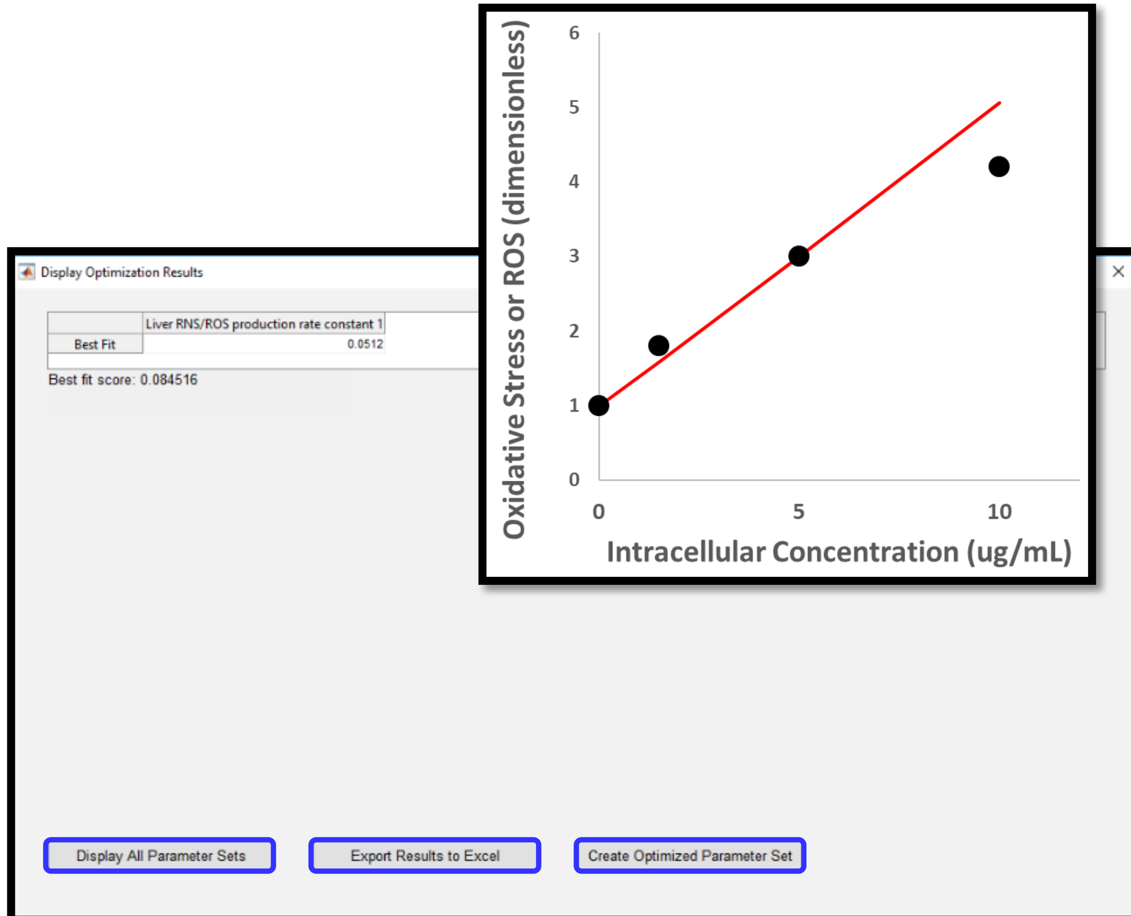
Hands-on Optimization Example –
Step 7 – Start Optimization with Run Button

- Fully defined optimization:
  - Parameter added
  - 3 SimSingles added
  - 3 Comparison Data sets added
- Save Setup before running
- Do NOT check "Probability" initially
- Click "Run"



Hands-on Optimization Example –
Step 8 – Explore Results Once Complete

- The "Display Optimization Results" window will appear
- Explore the buttons available
  - Display all sets
  - Export results to Excel
  - Create Optimized set of parameters from existing set
- Note example fit to data shown at right with solution value of 0.0512 for ***Liver RNS/ROS production rate constant 1***



Hands-on Optimization Example – *Step 9 – Load Setup and Add Probability Component*

- First, remove ***Liver RNS/ROS production rate constant 1*** parameter and add back with same bounds
  - Bounds: 0.01 to 0.1
- Add Distribution Type as shown at right, instead of Uniform
  - Same Throughout
  - Mean = 0.05
  - Stnd Dev = 0.15
  - Type = normal
- Check Probability box and save
- Once complete, Export results to Excel and compare results to those without Probability
- Notice difference in fitness score values
- Did solution value change?