

**Process for calculating ZIMS Life Table values in report ZR-047: Fecundity and Mortality**  
*Version 1.0*  
October 28, 2005

See Excel file “Life Table Zims.xls” for examples of methods.

With the development of ZIMS, we have the opportunity to revise the methods and approaches we currently use to calculate life-tables. This proposal outlines the methods ZIMS will use to calculate life-tables, highlighting change that will be made to the current methodology.

The life-table should capture the following kinds of events that occur to animals in a population:

- births
- deaths
- migration into population
- migration out of population
- lost-to-follow-up
- sex changes
- use of germ-plasm during life or after death of animal
- animals coming in and out of reproductive opportunity
- confidence intervals around parameters related to survival

**A. Definitions:**

The analysis window is the filter used to define the time/geographic period during which data are included in the analysis. It must be defined as a time period, but additional filters may be used as well, such as geographic restrictions (all animals in this set of zoos, or this region) as well as other restrictions (e.g., animals that were hand-reared). Need to have adequate filters to define this risk window. One of the filters needs to be user entered individual numbers (select specific individuals to analyse – or import a file of those individuals)

Population is defined as the set of animals that during any period of their life have entered the analysis window.

Reproductive Opportunity is defined as those periods of time when an individual is in an environment where reproduction can occur (e.g., opposite sex animal in same enclosure without contraceptives). User determines if reproductive rates will be calculated based only on periods of reproductive opportunity, or at any time.

**A. Tally the Data**

Create two separate arrays, one for males, one for females; create an additional one for sex “other” if User selected preference MethodUnknownSex is to use other life table. Each table with the following column headers for each of the following (see table 1).

- DAY
- BIRTHS
- START (animals in population at start of analysis window)

- IMPORT (entered population after birth; includes wild-caught animals)
- DEATHS: Non-management related deaths
- MGMT DEATHS: Management death (basically euthanasia)
- EXPORT (left population)
- ALIVE
- LOST-TO-FOLLOWUP (LTF; missing, no data after date DD/MM/YY)
- OFFN (number of birth produced by natural means)
- OFFA (number of births produced by artificial means)

The following two are only needed if user requested fecundity calculation based on reproductive risk only:

- InReproRisk (animals entering environment with possible reproduction)
- OutReproRisk (animals leaving reproductive environment)
- 

Create the following variables to track counts:

- UnknownAge = number of animals in population with unknown birthdates
- UnkSire = number of sires that are not registered in ZIMS or are listed as Unknown
- UnkAgeSire = number of sires who have unknown birth dates
- UnkDam – number of dams that are not registered in ZIMS or are listed as Unknown
- UnkAgeDam = number of dams who have unknown birth dates
- SexchangeMF = tally of sex changes
- UnkSex = number of non-male or non-females animals
- NParKnownAge = number births to known aged parents
- NNeonateBorn = number of animals entering population within 30 days of birth
- NNeonateDied = number of deaths within 30 days of birth

Record all ages as age in days from day of birth

For each animal in the population:

- If animal unknown age, increment UnknownAge by 1 and skip to next individual

**\* Calculating survival rates**

- if it's a female, use female array
- if it's a male, use male array
- if it is sex indeterminate:
  - Increment UnkSex by 1.0.
  - Depending on User selected option MethodUnknownSex:
    - use other array,
    - otherwise if animal is neither male nor female, increment event counts by 0.5 in male and female table rather than by 1.0 as described below (see below for variation on this)

- [note – in all cases if a record already exists with the DAY = age for that case, do not add another record, simply increment the appropriate column with 1]
- [note – whenever a row is added, it should be inserted in DAY order. So an event that occurs at age 75 should be inserted as a new row between the next younger age and the next older age]
- if it was born into the population, add a record with DAY = 0, and set BIRTHS = 1 or increment BIRTHS by 1 if a record with DAY = 0 already exists; Increment NNeonateBorn;
- if alive and in the population when analysis window starts, add record for DAY = age and set or increment START by 1; if <= 30 days of age, increment NNeonateBorn by 1.
- if it entered the population after birth after the start date, insert a record with DAY = the age it entered the population and set or increment IMPORTS by 1; if age <=30 days, increment NNeonateBorn by 1.
- if it died, insert a row with DAY = age in days it died and:
  - if it died a non-management death (not euthanized) set or increment DEATHS by 1;
  - if it was a management related death (euthanasia, others?) set or increment MGMT DEATH by 1;
  - if it died age <= 30 days, increment NNeonateDied by 1.
- if it was alive at the end of the analysis window and still in the population, insert a row with DAY = age at the end of the analysis window and set or increment ALIVE by 1;
- if it left the geographic area defined by the geographic dimension in the analysis window during the temporal period defined in the analysis window, insert record with DAY = age left then set or increment EXPORT with 1; if no geographic window was defined when defining the analysis window, then ignore this case;
- if it disappeared from the population, or was lost-to-followup during the analysis window, insert record with DAY = day last know to exist and set or increment LTF with 1;
- Note: an animal may exit and then re-enter the population multiple times. Each time it enters or leaves counts as a separate event that must be registered.

**\* Reproductive opportunity**

**\*If user selected to calculate fecundity based on reproductive opportunity, process the following section. Otherwise skip**

Insert record for any age animal entered reproductive opportunity, set or increment InReproRisk by 1

Insert record for any age animal left reproductive opportunity; set or increment OutReproRisk by 1

If animal alive at beginning of risk window and in reproductive opportunity, insert record with age for when entered, increment or set InReproRisk by 1

If animal alive at end of risk window, or exported from geographic risk window, insert record for that age and increment OutReproRisk by 1

If animal dies and is in reproductive risk window, increment OutReproRisk by 1 in same record recording DEATH

LTF and EXPORTs when animal in reproductive opportunity: increment OutReproRisk for that record as well.

If animal imported while in reproductive opportunity, increment InReproRisk for that import record as well.

\* Fecundity

- proceed if animal is born in analysis window. If not, loop to next animal
  - If dam and/or sire unknown, tally this by incrementing variable UnkDam and/or UnkSire by 0.5.
  - Increment NParKnownAge by 0.5 for each parent that is known (e.g., both parents known, increment by 1.0; only one parent known increment by 0.5; both parents unknown, don't increment)
  - if age of dam and/or sire unknown, tally this by incrementing a variable UnkAgeDam by 0.5 and/or UnkAgeSire by 0.5.
  - If both parents unknown or both parents age unknown, loop to next individual. Follow the instructions below for this parents who are both known and have know age when offspring born.
  - calculate age of dam when birth occurred, insert record with DAY = dam age or, if birth was through use of frozen gamete from female, the age of the dam when the gamete was collected;
  - calculate age of sire when birth occurred, insert record with DAY = sire age or, if birth was through use of frozen gamete from male, the age of the sire when the gamete was collected;
  - if user selected BirthSexRatio to be 50/50:
    - set or increment OFFN by 0.5 in both the male and female tables if the male and/or female parent provided its gamete naturally (natural mating, for female if AI's with sperm)
    - set or increment OFFA by 0.5 both the male and female tables if the male and/or female provided its gamete via ex vitro methods (for male if female AI'd with fresh sperm; for female and male if gamete fertilized externally and replaced). Or do we tally this for ANY kind of repro intervention??
  - if user selected BirthSexRatio to be assigned to real sex:
    - set or increment OFFN by 1 in the table of the sex of the individual if the male and/or female parent provided its gamete naturally (natural mating, for female if AI's with sperm)
    - set or increment OFFA by 1.0 in the table of the sex of the individual if either the male or the female contributed gamete via via ex vitro methods
    - if the sex of the individual is unknown, then increment OFFN or OFFA by 0.5 in both the male and female table

Next animal

## B. Producing the ZIMS Product-Moment Life Table (see Table 2 in “Life Table ZIMS.xls”)

Create two tables for males and females; if user selected MethodUnknownSex as separate life table, also create table for “other” sex. Each with the following headers:

DAY:	day of event
RISKQ <sub>x</sub> :	number of animals in population on this day potentially contributing information on events
<b>Only need this if user requested base fecundity on reproductive risk:</b>	
RISKM <sub>x</sub>	number of animals in reproductive opportunities on this day
Q <sub>x</sub> :	Proportion of animals alive day before that die on this day of age
MQ <sub>x</sub> :	Proportion of animals alive day before that die due to management reasons on this day of age
EXPORT <sub>x</sub> :	Proportion of animals alive day before that are exported from population on this day of age
LTF <sub>x</sub> :	Proportion of animals alive day before that are lost-to-followup on this age
IN <sub>x</sub> :	Proportion of animals alive day before that are imported on this age
Q' <sub>x</sub> :	Q <sub>x</sub> for desired combination of events selected by the user under TrackEvents. Can be set to be any of the following when running report: = Just Q <sub>x</sub> = [Q <sub>x</sub> + MQ <sub>x</sub> ] to reflect all deaths (Default) = [Q <sub>x</sub> + MQ <sub>x</sub> + EXPORT <sub>x</sub> ] to reflect removals from the pop (no longer death rate) = [Q <sub>x</sub> + MQ <sub>x</sub> + EXPORT <sub>x</sub> + LTF <sub>x</sub> ] to reflect all removals from pop = [Q <sub>x</sub> + MQ <sub>x</sub> + EXPORT <sub>x</sub> + LTF <sub>x</sub> – IN <sub>x</sub> ] to reflect all change events to the population = Any alone or any combination of: [ +Q <sub>x</sub> , +EXPORT <sub>x</sub> , +LTF <sub>x</sub> , -IN <sub>x</sub> , -MQ <sub>x</sub> ]. For example, user should be able to just combine (Q' <sub>x</sub> = EXPORT <sub>x</sub> – IN <sub>x</sub> ) to calculate rates for just migration events in the population.
P' <sub>x</sub> :	= 1 – Q' <sub>x</sub> = measure of surviving or remaining in population from day x-1 to day x
L' <sub>x</sub> :	= Survival/exist in population from day 0 to day x
M <sub>x</sub> :	= Average number of born to individuals contribution gametes on age x . This should be filterable to allow both Artificial and/or Naturally created offspring.
SUM_M <sub>x</sub> :	Cumulative sum of M <sub>x</sub> values

These are the instructions for completing the life-table – it perhaps may be easier to trace the formulas in the excel life-table shown in Table 2 of the attached:

For the first row (age = x = 0, ROW = 1)

$$\text{DAY} = 0$$

$$\text{RISKQ}_{x_0} = \text{Number of births from Table 1}$$

$$\text{RISKM}_{x_0} = 0$$

$$Q_0 = \text{DEATHS}_0 / \text{RISK}_0$$

$MQ_0$ ,  $EXPORT_0$ ,  $IN_0$  and  $LTF_0$  are all the number of those events that occurred on DAY = 0 divided by  $RISK_0$  (e.g.,  $EXPORT_0 = \# EXPORTS_0 / RISK_0$ )  
 $L'_x = 1$  by definition for Row 1,  $x = 0$

For all other rows ( $x$  is the age, row is the row number in the table):

DAY = next age that any event happened

$RISKQ_x = RISKQ_x(\text{previous row}) + START(\text{previous row}) + IMPORT(\text{previous row}) -$   
 $DEATHS(\text{previous row}) - MGMT DEATHS (\text{previous row}) -$   
 $EXPORTS(\text{previous row}) - LTF(\text{previous row}) - ALIVE(\text{previous row})$

$RISKM_x = RISKM_x(\text{previous row}) - OutReproRisk(\text{previous row}) +$   
 $InReproRisk(\text{previous row})$

$Q_x$ ,  $MQ_x$ ,  $EXPORT_x$ ,  $IN_x$ ,  $LTF_x$  are all calculated as the number of each of those events divided by  $RiskQ_x$

$Q'_x$  and  $P'_x$  are as defined above for each row

$L(\text{row}) = L'_x(\text{previous row}) * P'(\text{previous row})$

$E'_x(\text{row}) =$  see formulas and Life Table ZIMS.xls

$V'_x =$  see formulas and Life Table ZIMS.xls – but note that need to calculate  $r$  before can calculate this. So this is calculated only after all  $L_x$  and  $M_x$  are calculated.

$M_x = OFFA/RISKQ_x$ ,  $OFFN/RISKQ_x$ , or  $(OFFA + OFFN)/RISKQ_x$  depending on user selection option OffspringInclude. If user selected to base fecundity on reproductive opportunity only, the denominator is replaced by  $RISKM_x$  in each case above.

### **C. Adding Reproduction From Unknown Aged Or Unknown Parents (Calculating $M'_x$ )**

After the above calculations are made, there will be information on the number of unknown or unknown aged parents in the population. If there are NO unknown parents, or parents with unknown ages (i.e.,  $UnkSire + UnkDam + UnkAgeSire + UnkAgeDam = 0$ ), then this section can be skipped. If ANY of these are  $> 0$ , then proceed:

The user should be provided with the following information and prompted:

NParKnown “Births are attributed to know age parents”

UnkSire + UnkDam “births to unknown parents”

UnkAgeSire + UnkAgeDam “births to unknown aged parents”

“This is”  $(UnkSire + UnkDam + UnkAgeSire + UnkAgeDam) / NParKnown$  “% of the known data”

“Do you wish to include these unknown birth events in proportion to the known events?” = DistriburtMx

If DistributeMx = yes, proceed with this section.

This section is only processed if the user selects the DistributeMx option to “Yes”

Births to unknown aged or unknown parents need to be included in the life-table or reproduction is underestimated. These births are tallied in the UNKSIRE and UNKDAM variables above.

Our practice has been to add these births to the distribution of Mx values in proportion to the Mx values that have been observed:

Let  $M'_x$  = sex specific Mx rates with unassigned birth added

For the male life table:

Let  $N_{Bx}$  = number of births to age x males with unassigned births proportionally added.  
Then

$$NB_x = (OFFN_x + OFFA_x) \left[ 1 + \frac{UnkAgeSire}{N_m} \right]$$

Where  $N_m$  = sum of all  $(OFFN_x + OFFA_x)$  for males.

Then let  $M'_x = NB_{x} / RISKQ_x$ .

The denominator is replaced with  $RISKM_x$  if user specified to base fecundity calculations on reproductive opportunity.

Also,  $OFFN_x$  and  $OFFA_x$  should be replaced by simply  $OFFN_x$  or  $OFFA_x$  depending on user specifications.

For the female life table:

See above formulas for males, but replace *UnkAgeSire* with *UnkAgeDam*

(see Table 2 in “Life Table ZIMS.xls” using only OFFN in denominator)

Sum $M'_x$  = cumulative sum of the  $M'_x$  values at each x.

#### **D. Calculating Confidence Intervals for Survival Data**

This section deals with how to calculate confidence intervals for the  $L'_x$  data. The method uses the log-log survivorship function defined by Kalbfleish and Prentice (1980) and presented by

Hosmer, D. W. Jr. and Lemeshow, S. (1999). Applied Survival Analysis. John Wiley & Sons. Page 43. It is a common method for calculating confidence intervals on product-limit life tables.

It involves three steps:

1. Calculate the Standard Error of every estimate of the log-log survival estimates:  $SE_{L'_x}$
2. Calculate the confidence interval for the log-log  $L'_x$
3. Calculate the confidence interval for the  $L'_x$

Step 1.

The standard error for each  $L'_x$  value is calculated as:

$$SE_{L'_x} = \left( \frac{\sum_{t=1}^x \frac{DEATHS_t}{RISKQ_t (RISKDQ_t - DEATHS_t)}}{(\ln(L'_x))^2} \right)^{1/2}$$

The numerator should include all the items that are factored into the  $Q'_x$  – which could include all deaths, exports, imports and LTF.

$\ln(L'_x)$  is the natural log of  $L'_x$ .

Step 2.

The user has specified the level of confidence for this interval: 90%, 95% or 99%. The corresponding Z in formulas are 1.645, 1.96 and 2.576 respectively.

Lower X% confidence interval for the log-log function:

$$C_L = \ln[-\ln(L'_x)] - (Z)(SE_{L'_x})\{\ln[-\ln(L'_x)]\}$$

Upper X% confidence interval for the log-log function:

$$C_U = \ln[-\ln(L'_x)] + (Z)(SE_{L'_x})\{\ln[-\ln(L'_x)]\}$$

Step 3.

Convert this log-log survivorship function back into the survivorship function:

The Lower confidence interval for  $L'_x$  is:

$$LCI_{L'_x} = \exp[-\exp(C_U)]$$

The Upper confidence interval for  $L'_x$  is:

$$UCI_{L,x} = \exp[-\exp(C_L)]$$

where the  $C_u$  and  $C_L$  are from step 2. Note that the upper CI in step two is used for the lower CI in step 3 and visa-versa.

Thus, each  $L_x$  value will have its corresponding lower confidence interval estimate ( $LCI_{L,x}$ ) and an upper confidence interval estimate ( $UCI_{L,x}$ ).

The Excel file “Life Table ZIMS.xls” shows an example of how to calculate these.

These lower and upper figures will be used later to calculate confidence intervals around population summary statistics.

#### **D. Creating a User-Defined Age-Class Life-table**

The life table in Table 2 records probability events on only those days that ANY event happened. It may not coincide with any time interval we might be interested in. For example, we need to export data to PM2000 in annual age classes. But we want to build in more flexibility than that since annual age classes are not always the best things. For example, for small animals, weekly or monthly age classes are better; for gorillas, elephants, 5-year age classes might be better. Or we might want the first age class to be age 0 to 30, the second age 31 to 182, the third age 183 to 365, etc.

Thus, we want a report and life-table generator to work with any user-define age classes. I think by definition, the first age class needs to be 0. The others can be anything as long as each is older than the previous. The  $L'_x$  and  $M'_x$  values then for that class are then the survivorship/change from birth (age 0) to that age and the sum reproductive contribution for all ages cover the span of that age class.

The  $L'_x$  value for the user specified  $x$  is simply the  $L'_x$  for that  $x$  if  $x$  is a age in the life-table, or the  $L'_x$  for the next youngest  $x$ .

The  $E'_x$  value for the user specified  $x$  is simply the  $E'_x$  for that  $x$  if  $x$  is a age in the life-table, or the  $E'_x$  for the next youngest  $x$ .

The  $V'_x$  value for the user specified  $x$  is simply the  $V'_x$  for that  $x$  if  $x$  is a age in the life-table, or the  $V'_x$  for the next youngest  $x$ .

The  $Q'_x$ ,  $P'_x$  can then be back calculated from the  $L'_x$ :

$$P'_x = L'(x+1)/L'_x$$

$$Q'_x = 1 - P'_x$$

The  $M'x$  value for the user specified  $x$  is the sum of  $M'x$  values between the previous user-specified  $x$  up to and including the current user-specified  $x$ . Table 2 uses a function on the Sum $M'x$  column to do this.

For both these, see Table 3 in “Life Table ZIMS.xls”. Change the Age class definitions in the table and see corresponding changes to the other. Note that the report in Table 3 is for display purposes only. All summary calculations are done using data from Table 2.

## **D. Dealing with dirty data**

### **Date estimates:**

For any dates that are estimated, the life-table calculations should use the date that is the mid-point of a specified range, or the estimated date if it's a date with a date field entered with a +/- field. [Note: Future versions of ZIMS will export raw(er) data and actual dates will be randomly drawn from these ranges]

### **Unknown sexes:**

The user should have the ability to specify if they want to tally unknown sex in a separate life table, or treat unknown sexes as 50% male, 50% female. In the first case, the unknown sex life-table is generated as defined above. In the second, for each unknown sexed animal that appears in the defined population, the event counts are incremented by 0.5 in **both** the female and male life-table. For example, if an unknown sex dies, the DEATHS is incremented by 0.5 in both the female and male life-table. Every time a sex change occurs, SexChange is incremented by 1.

### **Unknown Parents:**

When parents are unknown, the  $Mx$  values can not be calculated since the age of the parents is indeterminable. The variables UnkDam and UnkSire are incremented, and used as described above to calculate  $M'x$ .

### **Multiple parents:**

Use the most-likely parent, or average of most-likely parent. User Option AgeMultiParents: If multiple parents all equally likely, use average age of all ages of parents when birth occurred or assume unknown age of parent and increment UnkDam and/or UnkSire.

### **Selfing:**

use age of self

### **Changing sex:**

The date the animal changed sex, it is treated like an export from its previous sex life-table and is added like an import in its changed sex life-table (i.e., removed from risk from previous, added to risk in current). Its life events (repro, death, etc) are applied against the sex it was at the time of those events occurred.

## E. Smoothing

See document for Process for SPARKS Life-table. The same process should be used but applied ONLY to the report of life-table by age class. The values for the age classes are smoothed as described in the SPARKS process.

## F. Summary Statistics from Life-Tables

We would like to be able to calculate the following sex-specific values parameters from the life-table for the male and female life tables. Since there will be no  $M_x$  data for the “other” table, these can not be calculated. All these calculations use table 2 data:

- life-table  $r$  :

The sum is over all  $x$  ages in the table,  $r$  is the value that makes this sum = 1.0000. Should solve  $r$  to the nearest 3 sig. digits. Usually  $r$  is solved using an iterative process.

$$\sum L'_x M'_x e^{-rx} = 1.0$$

Use this value of  $r$  in the formulas below.

An upper and lower confidence interval for  $r$  can be calculated by replacing the  $L'_x$  with its lower confidence limit estimates ( $LCI_{L'_x}$ ) and upper confidence limit estimates ( $UCI_{L'_x}$ ) from section D above.

- life-table lambda ( $\lambda$ ):

$$\lambda = e^r$$

An upper and lower confidence interval for lambda can be calculated by replacing the  $r$  with its lower confidence limit estimates and upper confidence limit estimates from above.

- Net reproductive rate ( $R_0$ ):

$$R_0 = \sum L'_x M'_x$$

An upper and lower confidence interval for  $R_0$  can be calculated by replacing the  $L'_x$  with its lower confidence limit estimates ( $LCI_{L'_x}$ ) and upper confidence limit estimates ( $UCI_{L'_x}$ ) from section D above.

- generation length ( $T$ ):

$$T = \frac{\sum xL'_x M'_x e^{-r}}{\sum L'_x M'_x e^{-r}}$$

An upper and lower confidence interval for  $T$  can be calculated by replacing the  $L'_x$  and  $r$  with their lower confidence limit estimates and upper confidence limit estimates from above.

- life-expectancy ( $E_x$ ):

$$E_x = \frac{\sum_{y=x}^{\text{max age}} L'_y}{L'_x}$$

An upper and lower confidence interval for  $E_x$  can be calculated by replacing the  $L'_x$  with its lower confidence limit estimates ( $LCI_{L'_x}$ ) and upper confidence limit estimates ( $UCI_{L'_x}$ ) from section D above.

- Reproductive value ( $V_x$ ):

$$V_x = \left( \frac{e^{rx}}{L'_x} \right) \sum_{y=x}^{\text{max age}} L'_y M'_y e^{-ry}$$

An upper and lower confidence interval for  $V_x$  can be calculated by replacing the  $L'_x$  and  $r$  with their lower and upper confidence limit estimates.

- 30 day survival rate ( $L_{30}$ ):

The  $L'_x$  for 30 days if 30 is an age in the life-table, or the  $L'_x$  for the next youngest  $x$  in the table.

## G. Life-Stages and Life-Tables

For species with life-stages for individuals:

For species with life-stages where individuals are tracked (i.e., non-groups), very much the same method is used. For the day a particular event happens (e.g., death, gave birth, shipped out), the life-stage is known. Thus, when the tables are presented, the life-stage associated with that age is shown. However, for the table with age classes, here the

age classes are already predefined according to the life-stages – so the user does not need to define age classes. Thus the procedures are the same as for individuals with non-age classes – but age classes are predefined and stage should be added as an additional column to the right of age. Examples of each of the tables with life-stages are shown in Table 097-5.

## **H. Groups**

We will need to determine what demography to do on groups.

## **I. Business rules**

- There may be appropriate warning flags if sample sizes are small.