DALY Calculator

A graphical user interface for stochastic DALY calculation in ***R***

Brecht Devleesschauwer[1,2,3], Arie Havelaar[4,5], Juanita Haagsma[6], Nicolas Praet[3], Niko Speybroeck[2]

[1] Department of Virology, Parasitology and Immunology, Faculty of Veterinary Medicine, Ghent University, Ghent, Belgium
[2] Institute of Health and Society, Faculty of Public Health, Catholic University of Louvain, Louvain-la-Neuve, Belgium
[3] Department of Animal Health, Institute of Tropical Medicine, Antwerp, Belgium
[4] Laboratory for Zoonoses and Environmental Microbiology, National Institute for Public Health and the Environment, Bilthoven, The Netherlands
[5] Institute for Risk Assessment Sciences, University Utrecht, Utrecht, the Netherlands
[6] Department of Public Health, Erasmus Medical Center, Erasmus University, Rotterdam, The Netherlands

ABSTRACT

The disability-adjusted life year or DALY is an increasingly used measure of population health. The **DALY** Package, accessible in the **R** statistical programming environment, provides a Graphical User Interface (GUI) for calculating DALYs and performing uncertainty analysis. The latter is done through Monte Carlo simulations, which are used to compute DALY credible intervals based on the uncertainty in the various input parameters. The GUI of the *DALY Calculator* allows for ease and flexibility, while the **R** environment enables more advanced computations and graphical functions. The underlying calculation methods are designed to promote consistency in the uncertainty analysis of DALYs. These methods are demonstrated through the inclusion of two examples from the foodborne disease burden literature (i.e., the burden of neurocysticercosis in West-Cameroon and the burden of congenital toxoplasmosis in the Netherlands). We believe that the *DALY Calculator* can become a useful tool for students, scientists and health officials involved in the quantification of public health problems.

Corresponding Author:

Brecht Devleesschauwer

Department of Virology, Parasitology and Immunology
Faculty of Veterinary Medicine
Ghent University
Salisburylaan 133
9820 Merelbeke
Belgium
e-mail: Brecht.Devleesschauwer@UGent.be

1. Introduction

The Disability-Adjusted Life Year or DALY is a summary measure of population health widely used in the assessment of disease burden and in the cost-effectiveness analysis of intervention measures (Murray and Lopez, 1996; Lopez et al., 2006). DALYs represent the number of healthy life years lost due to a disease or disability, and do so by incorporating non-fatal and fatal health outcomes, calculated as the years of life lived with disability (YLD) and the years of life lost due to premature death (YLL), respectively. In order to make a direct comparison of YLDs and YLLs possible, disability weights are assigned to various degrees of disability, ranging from zero (full health) to one (worst possible health state). In addition to these disability weights, the DALY calculation requires the choice of two other social values, i.e., age weighting and time discounting. The former gives a higher weight to the healthy life years lived between the age of 9 and 54, as this period of life is considered to be socially more important than the younger and older life spans (Murray, 1994). The latter discounts the years of healthy life lived in the future, at a rate of (usually) 3%. The incorporation of a time discounting rate reflects similar practices in economic assessments, and would prevent policy makers from saving resources for a possible future eradication program, instead of investing in currently available, but less effective, intervention measures (the so-called "disease eradication and research paradox"; Murray, 1994).

The formulas for calculating the YLLs and YLDs are presented in Murray (1994) and Murray and Acharya (1997), and can easily be incorporated in a spreadsheet, such as the "DALY calculation template"[1] prepared by the World Health Organization for its Global Burden of Disease (GBD) project (Mathers et al., 2001). This template allows the computation of YLLs and YLDs for both sexes and various age groups, which are then summarized into a single deterministic DALY measure.

The reliability of the final DALY result depends heavily on the quality of the epidemiological data, which are commonly derived from routine data collection systems, scientific literature, and expert elicitation. The estimates provided by these data sources include an inherent level of uncertainty, mainly due to sampling error, diagnostic uncertainty, and population heterogeneity. To reflect this stochastic nature, epidemiological parameters are often accompanied by a confidence or credible interval, or represented by a probability distribution, rather than being represented by a single point estimate. Monte Carlo simulations have been suggested as the appropriate technique to incorporate this uncertainty in the final DALY result (de Vocht et al., 2010), which can then be presented as a point estimate with a credible interval. This interval allows the assessment of the level of uncertainty in the total DALYs, and facilitates a more reliable comparison of the health impact of different diseases.

To our knowledge, however, there are no standardized tools available for stochastic DALY calculation. Therefore, we designed a Graphical User Interface (GUI) for calculating DALYs that allows the incorporation of the input uncertainty and the computation of a DALY credible interval through Monte Carlo simulations. This program, the *DALY Calculator*, is designed to be used by a variety of users, with different levels of statistical skills, to allow maximum flexibility, and to promote consistency in the DALY uncertainty analysis.

[1] http://www.who.int/entity/healthinfo/bodreferencedalycalculationtemplate.xls

2. Installing and running the *DALY Calculator*

The *DALY Calculator* is developed in **R**, an open-source environment for statistical programming and graphics (R Development Core Team, 2010). The use of the *DALY Calculator* requires the prior installation of **R**, which can be freely downloaded from the Comprehensive R Archive Network (CRAN):

http://cran.r-project.org/

The **DALY** package is available on the CRAN repository, and can be installed by accessing the following menu in **R**:

Packages > Install package(s)…

Next, the required CRAN mirror has to be selected, and the **DALY** package can be selected and installed from the list of available packages.

Finally, every time the *DALY Calculator* has to be used, the **DALY** package has to be loaded in the **R** environment, by typing the following command in the **R** console (Figure 1):

```
library(DALY)
```

This call will load the **DALY** package and initiate the main window of the *DALY Calculator* (Figure 2). To re-initiate this window, the following function has to be called in the **R** console:

```
DALYcalculator()
```

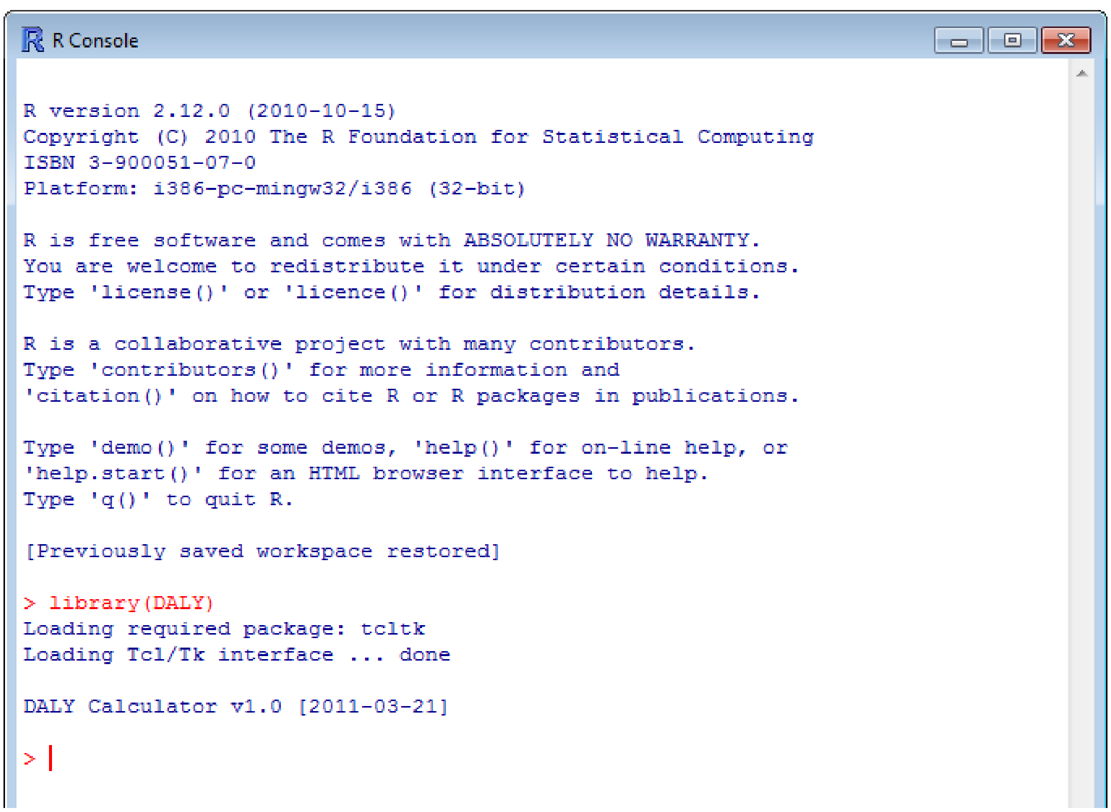

Figure 1. Initiating the DALY Calculator

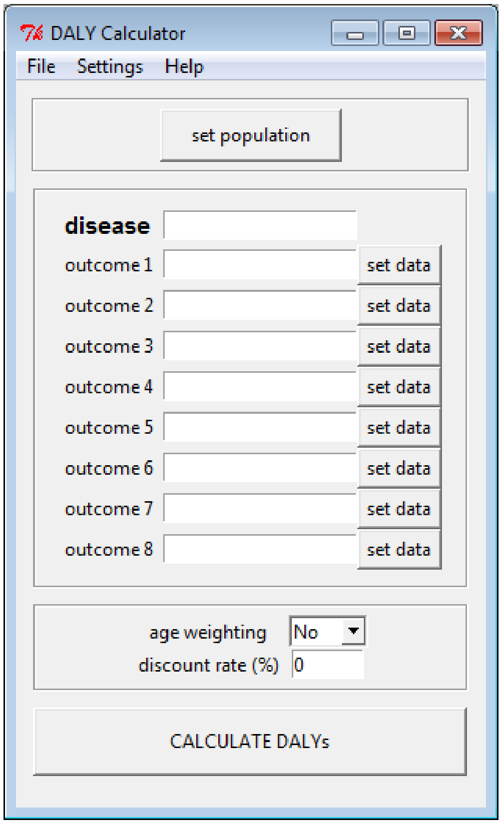

Figure 2. Main window of the DALY Calculator

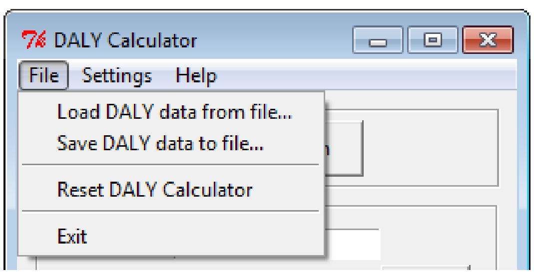

Figure 2.1. DALY Calculator 'File' menu

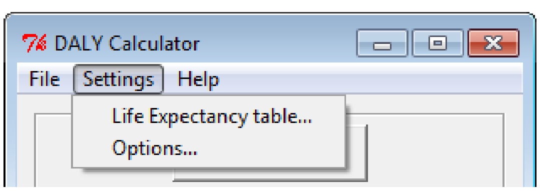

Figure 2.2. DALY Calculator 'Settings' menu

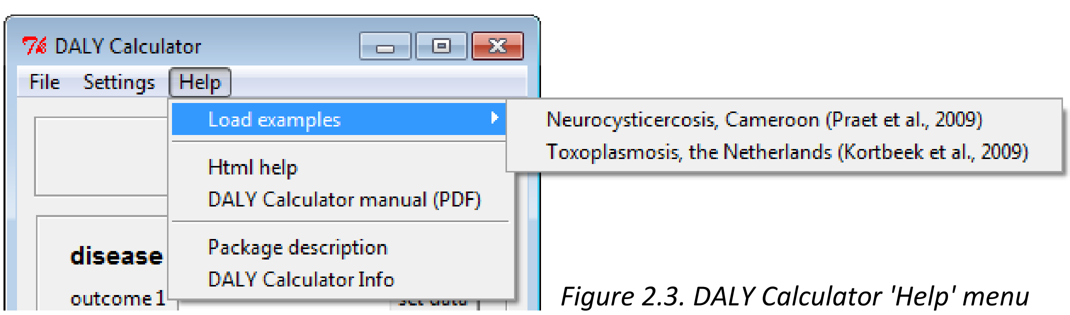

Figure 2.3. DALY Calculator 'Help' menu

3. DALY calculation with the *DALY Calculator*

The *DALY Calculator* is designed to be used in an outcome-based as well as an agent-based approach. Depending on the approach, YLDs, YLLs, DALYs and deaths can be computed for a combination of up to eight different disease categories of one outcome, or of up to eight different outcomes of one agent.

According to the GBD studies, the incidence-based approach was favored over the prevalence-based approach. The current version of the *DALY Calculator* is able to calculate DALYs for a basic incidence-based disease model with well-defined incidence and/or mortality rates per disease category or outcome.

The default age groups used by the *DALY Calculator* are the five age groups used by the GBD 1990 study: 0-4; 5-14; 15-44; 45-59; 60+ (Lopez and Murray, 1996). At least one combination of sex and age group has to be set in order to proceed with the DALY calculation.

To compute YLDs, YLLs, DALYs and deaths, five steps have to be followed (as illustrated in Figure 3). These steps will be explained using one of the two built-in examples (The burden of *Taenia solium* cysticercosis in Cameroon; Praet et al., 2009). Users can load this example by selecting (Figure 2.3):

Help > Load examples > Neurocysticercosis, Cameroon (Praet et al., 2009)

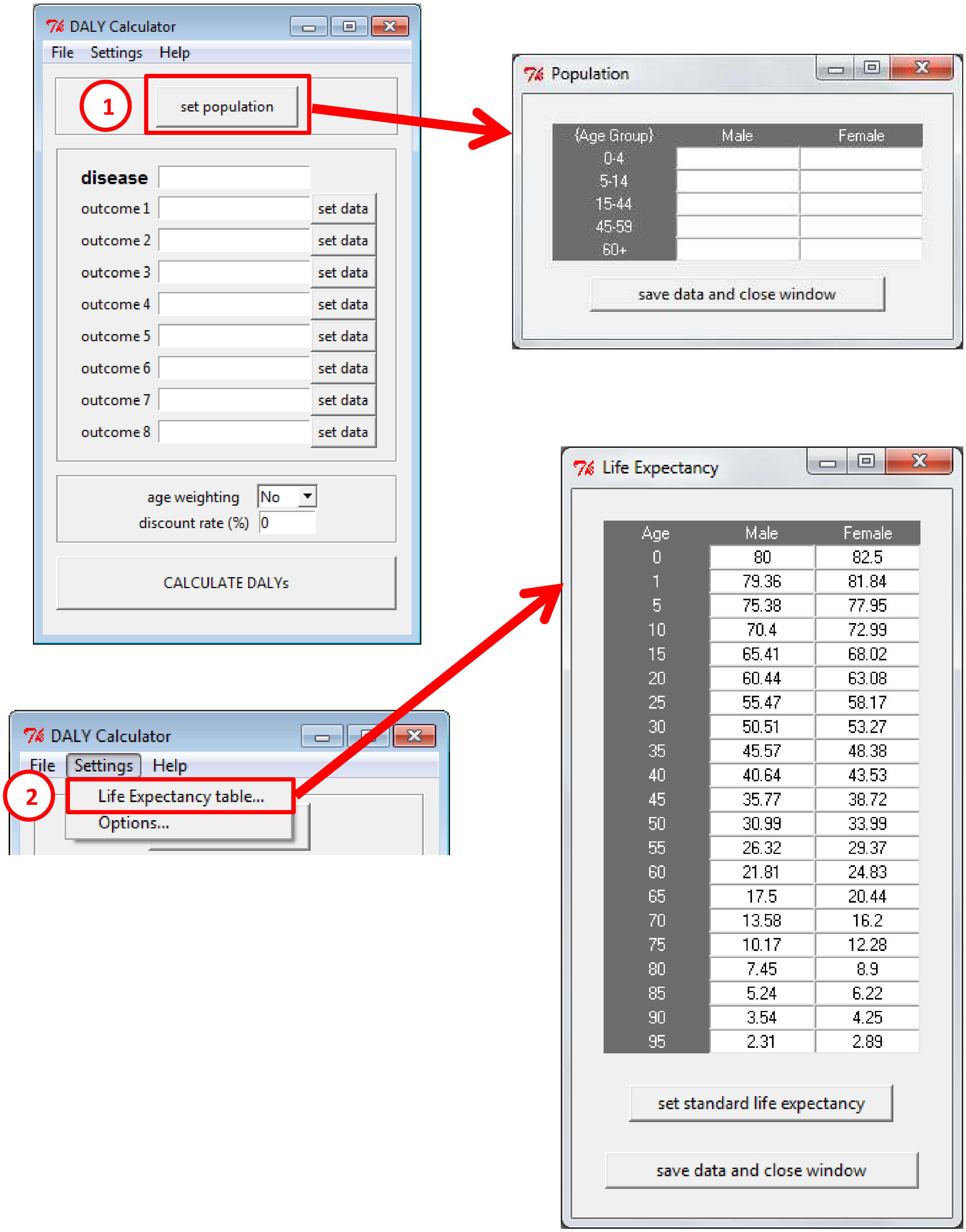

Figure 3. DALY calculation with the DALY Calculator

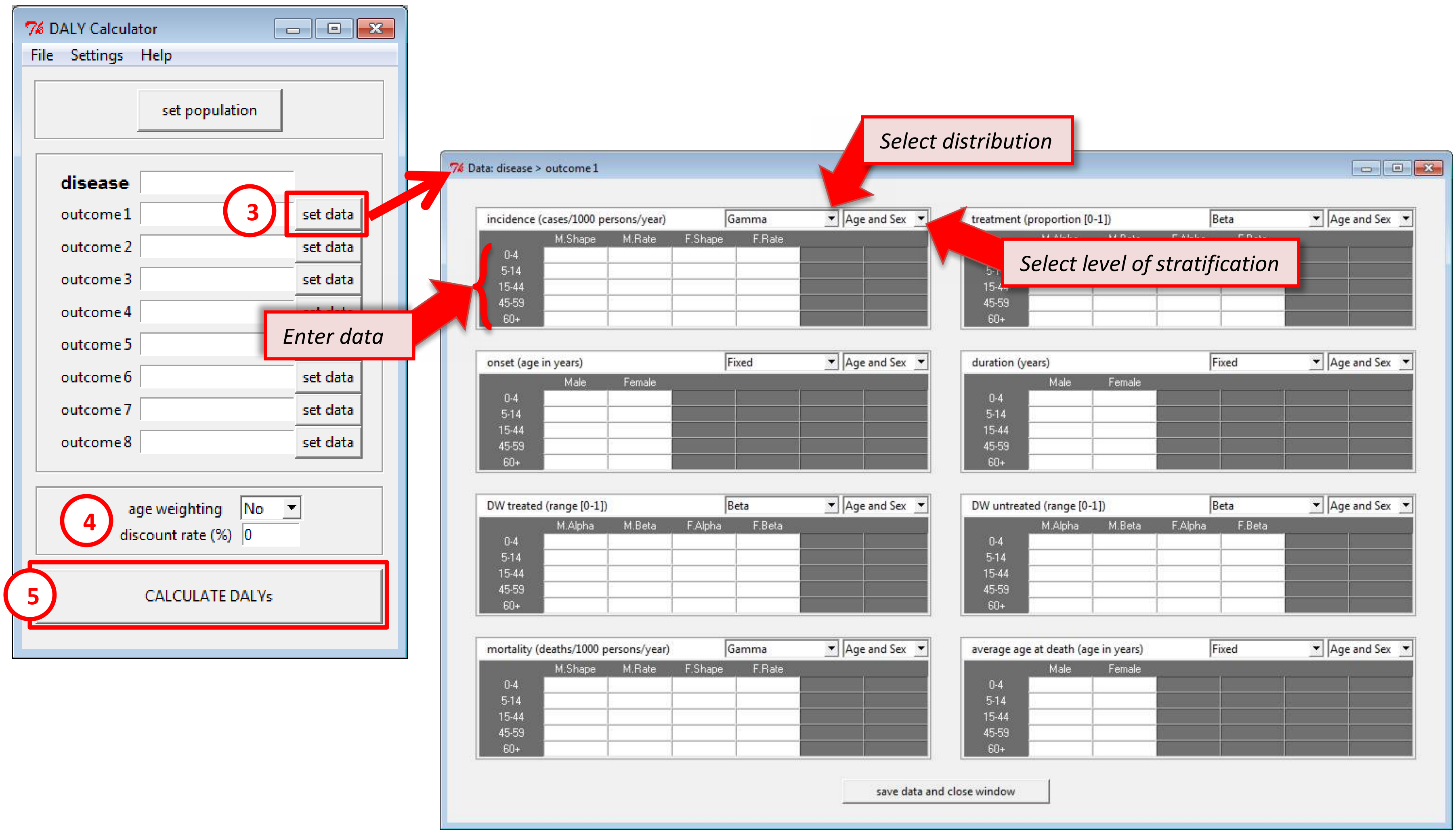

Figure 3. DALY calculation with the DALY Calculator (continued; M = Male, F = Female)

① **Set the population table**

Clicking the "set population" button opens the population window, where the number of males and females, per age group, can be entered. At least one combination of sex and age group has to be set.

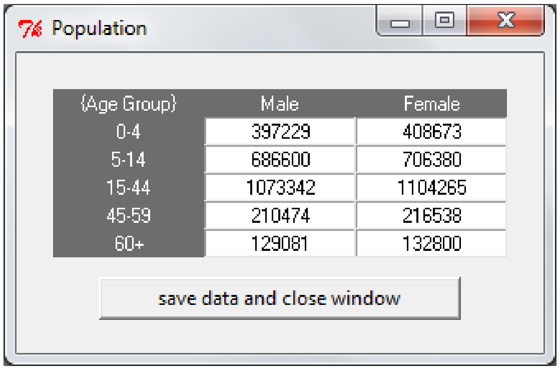

Figure 3.1. Population of West-Cameroon, stratified by sex and age

② **Set the life expectancy table**

The default life expectancy table used by the *DALY Calculator* is the Coale and Demeny model life-table West, level 26 and 25, which has a life expectancy at birth of 80 for males and 82.5 for females (Murray, 1994). However, the user can define his own life expectancy table, by accessing the 'Life Expectancy' window through the 'Settings' menu:

Settings > Life Expectancy

Praet et al. (2009) applied the standard life expectancy table, which is the default life expectancy table of the DALY Calculator. Therefore, no action is required during this step.

③ **Set the input parameters, per disease category or outcome**

After entering the disease name and the name(s) of the disease state(s) or outcome(s), the epidemiological data and disability weights can be entered by clicking the "set data" button. For every parameter, the user can specify the distribution by selecting one of following:

- **Beta-Pert** (mode; min; max)
- **Beta** (alpha; beta)
- **Gamma** (shape; rate)
- **Normal** (mu; sigma)
- **Lognormal-geometric** (logmean; logsigma)
- **Lognormal-arithmetic** (mean; sigma)
- **Uniform** (min; max)
- **Fixed**

Next, the user has to select the specific level of stratification for every parameter. The following four stratification levels have been made available:

- **Age and Sex** (i.e., full stratification)
- **Age** (i.e., data is stratified by age group, but not by sex)
- **Sex** (i.e., data is stratified by sex, but not by age group)
- **None** (i.e., no stratification, data applies to total population)

For calculating the YLDs, the following tables have to be completed:

- **Incidence**: number of new cases per 1.000 persons per year
- **Treatment**: proportion of patients receiving proper treatment [0-1]
- **Onset of the disease**: age of onset in years
- **Duration of the disease**: duration in years
- **Disability Weight for treated cases**: range [0-1]
- **Disability Weight for non-treated cases**: range [0-1]

For calculating the YLLs and deaths, the following tables have to be completed:

- **Mortality**: number of deaths per 1.000 persons per year
- **Average age at death**: age at death in years

Based on the entered average ages at death, the *DALY Calculator* will compute the corresponding life expectancies according to the default or user-defined life table.

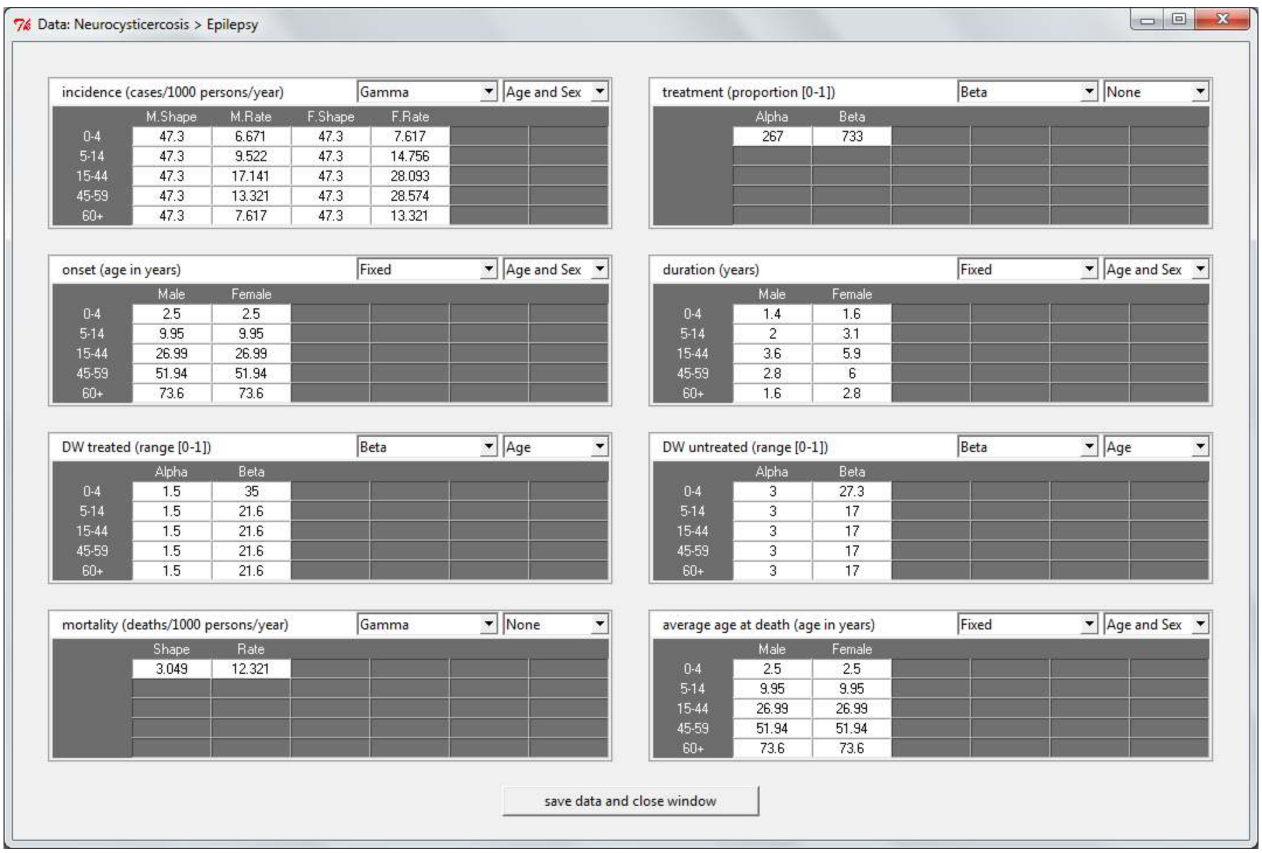

Figure 3.2. Epidemiological parameters for the DALY calculation of Neurocysticercosis in Cameroon, based on Praet et al. (2009)

④ **<u>Set the social values:</u>**

Age weighting: yes or no
Discount rate (%)

The default social values applied by the *DALY calculator* are uniform age weights (i.e., no age weighting) and a zero discount rate. However, the user is given the possibility to define the required set of social values, and to alter these values to assess their influence on the final result.

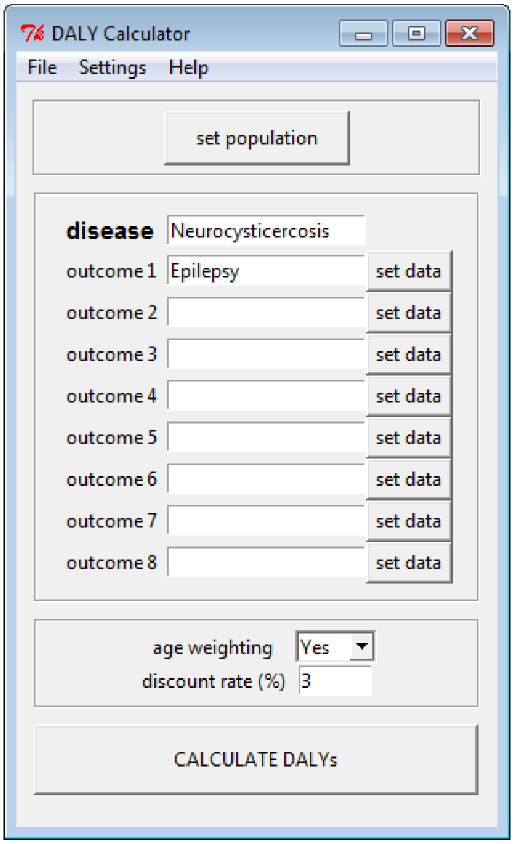

Figure 3.3. DALY Calculator main window, with social values set to full age-weighting and a 3% discount rate, as applied by Praet et al. (2009)

⑤ **<u>CALCULATE DALYs</u>**

(1) Using the GUI

Clicking this button will read the data and compute the YLDs, YLLs, DALYs and deaths per disease category or outcome, and sum these up to total YLDs, YLLs, DALYs and deaths. The mean, median and a 95% credible interval will be printed to the **R** Console, and a histogram of the total DALYs, with accompanying density distribution and credible interval, will be displayed (Figure 5). In order to have a more detailed output of the DALY calculation, with YLDs, YLLs, DALYs and deaths per outcome, the user can select "Advanced output" in the 'Options' window:

Settings > Options

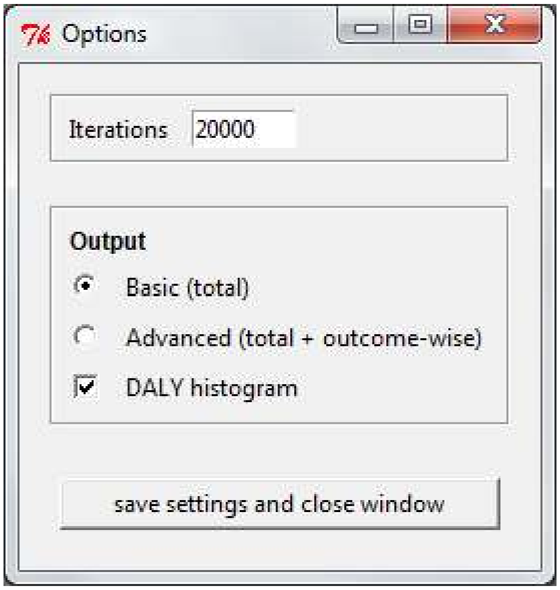

Figure 4. Options window

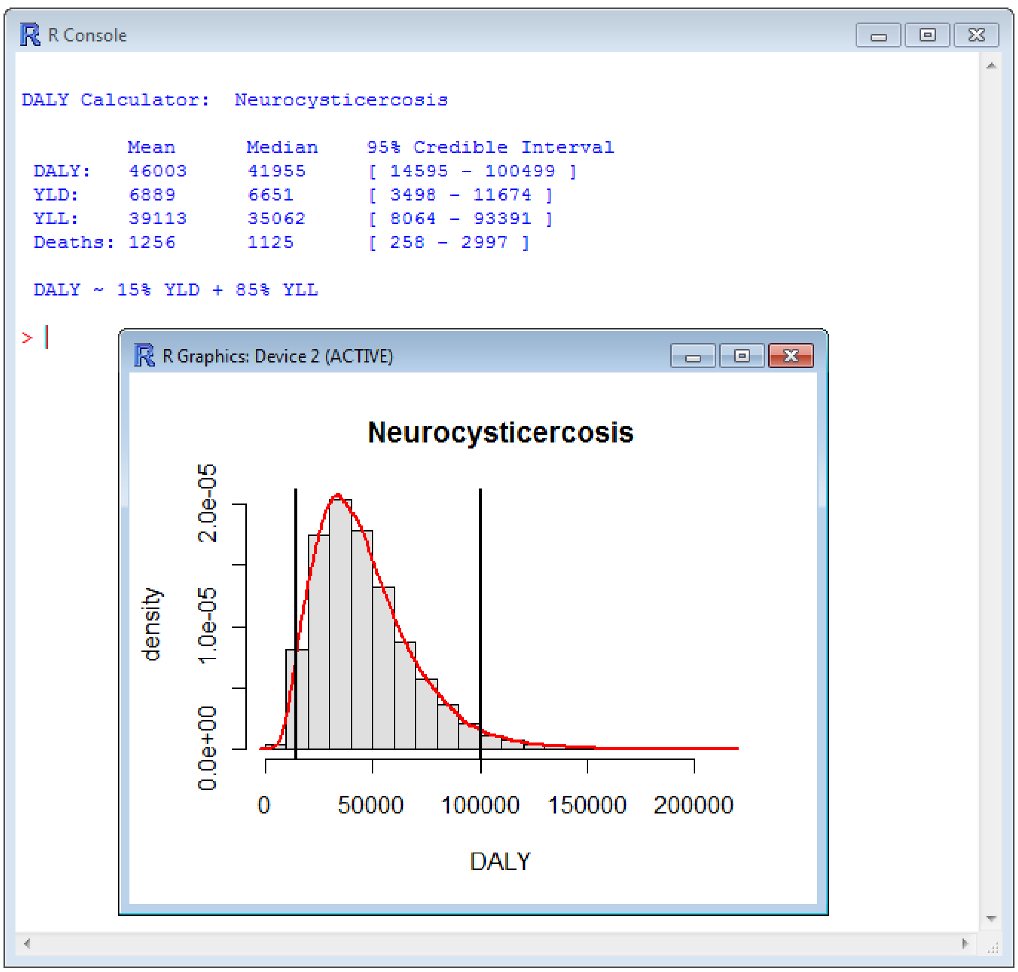

Figure 5. Output of the DALY calculation for Taenia solium Neurocysticercosis in Cameroon (based on Praet et al., 2009)

(2) Using the R Console

Instead of clicking the "CALCULATE DALYs" button in the *DALY Calculator* main window, **R** users can call the `getDALY()` function in the **R** console. This function initiates the Monte Carlo simulation process, and returns the list of total and category- or outcome-wise DALYs, YLDs, YLLs and deaths, computed during each iteration. These results can be stored in an object, which will contain sub-objects for total DALYs, YLDs, YLLs and deaths, and for DALYs, YLDs, YLLs and deaths per category or outcome (available in sub-objects `outcome1` to `outcome8`, depending on the number of outcomes specified by the user). The function `names()` can be used to get an overview of all available sub-objects. Three generic functions are made available for these objects and their sub-objects: `print()`, `summary()` and `hist()`. In addition, users can generate custom output summaries, analyses or graphics through the various statistical and graphical functions available in the **R** environment.

After setting the population and life expectancy table, and entering the epidemiological data and social values, the `getDALY()` function can be called from the **R** console (Figure 6):

```
x <- getDALY()            # save the DALY output in object "x"
names(x)                  # print the objects in the DALY output

hist(x)                   # display the DALY histogram
print(x)                  # print an overview of total DALYs, YLDs
                          # and YLLs
summary(x)                # print a full overview of DALYs, YLDs
                          # and YLLs, in total and per outcome

hist(x$outcome1)          # display the DALY histogram of outcome 1
print(x$outcome1)         # print an overview of DALYs, YLDs and
                          # YLLs due to outcome 1
```

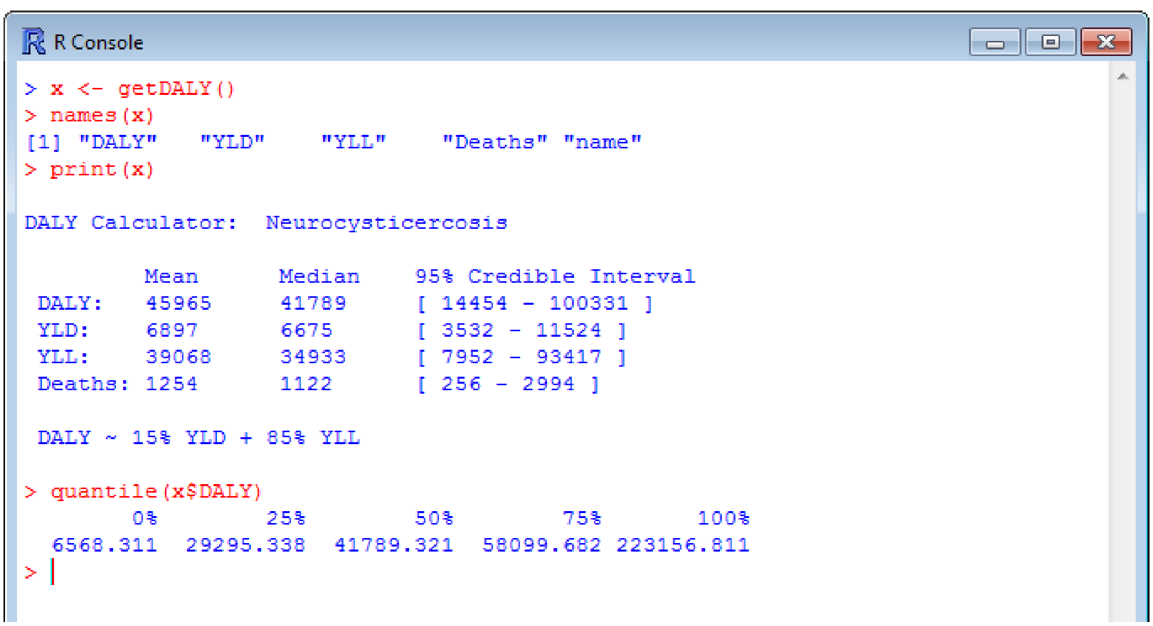

*Figure 6. Output of the DALY calculation for Taenia solium Neurocysticercosis in Cameroon (based on Praet et al., 2009), obtained through accessing the **R** console*

When initiating the GUI of the *DALY Calculator*, certain parameters will be set to their default values. To allow flexibility, however, users are given the possibility to adapt these settings to their needs. The following parameters are set by default:

- The Coale and Demeny model life-table West, level 26 and 25, will be set as default life tables, as proposed by Murray (1994);
- The distributions of incidence and mortality will be set to "Gamma", those of the proportion treated and the disability weights to "Beta", and those of onset, duration and average age at death to "Fixed";
- The level of stratification for all parameters will be set to full stratification (i.e., "Age and Sex");
- The initial social values for the DALY calculation will be no age weighting and a zero discount rate;
- The total number of Monte Carlo iterations will be fixed to 20,000, since this number has been reported to give a sufficiently low difference between subsequent runs (Praet et al., 2009).

4. DALY calculation process

The DALY uncertainty analysis implemented by the *DALY Calculator* is performed through Monte Carlo simulations. This method iteratively simulates random values for each of the epidemiological parameters and disability weights, according to their user-defined distributions. During each iteration, these random values are incorporated in a YLD and YLL calculation, which can then be summed up to a single DALY measure. The calculation of YLDs and YLLs, per outcome o and iteration n, is given by:

$$YLD_{ij} = rbinom(pop_{ij}, inc_{ij}) * (trt_{ij} * DWt_{ij} + \{1 - trt_{ij}\} * DWn_{ij}) * \boldsymbol{Formula}(ons_{ij}, dur_{ij}) \quad (1)$$

$$YLL_{ij} = rbinom(pop_{ij}, mrt_{ij}) * 1 * \boldsymbol{Formula}(aad_{ij}, lxp_{ij}) \quad (2)$$

Where,

i = sex;
j = age group;
YLD_{ij} = YLDs for sex i and age group j;
YLL_{ij} = YLLs for sex i and age group j;
rbinom(n,p) = **R** function which generates a random value based on a binomial distribution with parameters n and p;
pop_{ij} = population belonging to sex i and age group j;
inc_{ij} = randomly generated incidence for sex i and age group j;
trt_{ij} = randomly generated treatment rate for sex i and age group j;
DWt_{ij} = randomly generated Disability Weight–treated for sex i and age group j;
DWn_{ij} = randomly generated Disability Weight–untreated for sex i and age group j;
ons_{ij} = randomly generated age at disease onset for sex i and age group j;
dur_{ij} = randomly generated disease duration for sex i and age group j;
mrt_{ij} = randomly generated mortality for sex i and age group j;
aad_{ij} = randomly generated age at death for sex i and age group j;
lxp_{ij} = calculated life expectancy for sex i and age group j.

The **Formula(a,L)** function used in equations (1) and (2) is the solution of the integral used to calculate the years of life lived with disability or the years of life lost due to premature mortality for one individual. The integral for calculating age-weighted and time-discounted YLLs or YLDs is given by Murray (1994):

$$\int_{x=a}^{x=a+L} Cxe^{-\beta x}e^{-r(x-a)}dx \tag{3}$$

Where,

a = age of onset *or* age at death;
L = disease duration *or* life expectancy at age *a*;
C = age-weighting correction constant (C = 0.1658);
β = parameter from age-weighting function (β = 0.04);
r = discount rate.

In order to allow the calculation of uniformly age-weighted DALYs, an age weighting modulation factor (K) is introduced, which can be either one (non-uniform age weights) or zero (uniform age weights, i.e., no age-weighting) (Murray and Acharya, 1997):

$$\int_{x=a}^{x=a+L} KCxe^{-\beta x}e^{-r(x-a)}dx \tag{4}$$

The solution of integral (4) gives us the general form of the YLD and YLL formula:

$$\frac{KCe^{ra}}{(r+\beta)^2}\left[e^{-(r+\beta)(L+a)}\{-(r+\beta)(L+a)-1\}-e^{-(r+\beta)a}\{-(r+\beta)a-1\}\right] + \frac{1-K}{r}(1-e^{-rL}) \tag{5}$$

When r approaches zero (i.e., no time-discounting), the limit of equation (5) becomes:

$$\frac{KCe^{-\beta a}}{\beta^2}\left[e^{-\beta L}\{-\beta(L+a)-1\}-(-\beta a-1)\right]+(1-K)L \tag{6}$$

For the case where both r and K are zero (i.e., no time-discounting and no age-weighting), equation (6) equals to L, the disease duration (for the calculation of YLDs) or the life expectancy at age a (for the calculation of YLLs).

This process of simulating YLD and YLL values is iteratively repeated for each combination of sex and age group (i,j) within each outcome o. By summarizing, per outcome, the YLDs and YLLs obtained for each combination of sex and age group, a vector of length n is generated, which contains n simulated YLDs and YLLs for outcome o:

$$\boldsymbol{YLD}_o = \langle \textstyle\sum_i \sum_j YLD_{oij} \rangle$$

$$\boldsymbol{YLL}_o = \langle \textstyle\sum_i \sum_j YLL_{oij} \rangle$$

These vectors can then be summarized to a single DALY vector for outcome o:

$$\boldsymbol{DALY}_o = \langle \sum_i \sum_j YLD_{oij} + \sum_i \sum_j YLL_{oij} \rangle = \langle \boldsymbol{YLD}_o + \boldsymbol{YLL}_o \rangle$$

Finally, vectors of total YLDs, YLLs and DALYs are obtained by summarizing the YLD, YLL and DALY vectors obtained for each outcome o:

$$\boldsymbol{YLD}_{total} = \langle \sum_o \boldsymbol{YLD}_o \rangle$$

$$\boldsymbol{YLL}_{total} = \langle \sum_o \boldsymbol{YLL}_o \rangle$$

$$\boldsymbol{DALY}_{total} = \langle \sum_o \boldsymbol{DALY}_o \rangle$$

In a similar way, vectors are generated for deaths per outcome and for total deaths:

$$Deaths_{ij} = rbinom\left(pop_{ij}, mrt_{ij}\right) \qquad (7)$$

$$\boldsymbol{Deaths}_o = \langle \sum_i \sum_j Deaths_{oij} \rangle$$

$$\boldsymbol{Deaths}_{total} = \langle \sum_o \boldsymbol{Deaths}_o \rangle$$

Based upon these vectors, the mean and median YLDs, YLLs, DALYs and deaths can be calculated per outcome and in total, and a credible interval can be derived for these parameters by selecting the required quantiles of their respective vector.

As the epidemiological data and disability weights are often not available per combination of age group and sex, the total uncertainty would be underestimated if the exact level of data stratification would not be taken in account. Therefore, the *DALY Calculator* enables users to specify the level of stratification of the input parameters, and incorporates this information in the DALY calculation process. For the case where the data applies to the total population (i.e., no stratification), a single random value is generated per iteration and applied to each age and sex category. If the data is only stratified by age (sex), a random value is generated during each iteration per age (sex) category, and applied to all sex (age) categories. Finally, in case of full stratification, during each iteration, a random value will be generated for each combination of age group and sex, reflecting the maximum level of detail of the respective input parameter.

The implementation of the above-mentioned features is primarily written in the **R** programming language, with additional support from **Tcl/Tk** (using the **tcltk** package; R Development Core Team, 2010), to design the GUI. In order to speed up the computation process, the *DALY Calculator* interacts with the **C++** programming language, which is used to randomly generate the epidemiological input values (through **R** random number generation functions), and to calculate the YLDs, YLLs and deaths, based on equations (5), (6) and (7). These results are then sent back to the **R** environment, where they are summed up to obtain DALYs and processed into their corresponding sub-objects. Finally, the output of the results and their graphical representation is handled by **R**, which is optimized toward these tasks.

5. Built-in examples

To demonstrate the calculation process of the *DALY Calculator*, two DALY calculation examples are derived from the foodborne disease literature, and are built-in to the *DALY Calculator*:

- "The burden of *Taenia solium* cysticercosis in Cameroon" (Praet et al., 2009); and
- "Congenital toxoplasmosis and DALYs in the Netherlands" (Kortbeek et al., 2009).

These examples can be loaded by accessing the 'Help' menu, selecting 'Load examples', and selecting the required examples (Figure 2.3):

Help > Load examples > Neurocysticercosis, Cameroon (Praet et al., 2009)
Help > Load examples > Toxoplasmosis, the Netherlands (Kortbeek et al., 2009)

5.1. The burden of Taenia solium cysticercosis in Cameroon

Praet and colleagues (2009) estimated the health impact of *T. solium* by calculating the DALYs lost due to neurocysticercosis (NCC)-associated epilepsy, as this is the main health outcome of NCC. Non-uniform age weights and a 3% discount rate were selected as DALY parameters, and YLLs were calculated based on the standard life expectancy. The authors presented the following results:

```
         Mean         95% Credible Interval      % of total DALYs
DALY:    45838.4      [ 14108.1 - 103469.4 ]     100.0
YLD:     6821.4       [ 2765.1 - 12878.4 ]       14.9
YLL:     39017.0      [ 8195.6 - 95512.8 ]       85.1
```

The authors did not directly provide probability distributions for the incidence and mortality of NCC-associated epilepsy, but estimated these values through the following formulas:

$$I = P * N / D \quad (8)$$
$$M = P * N * C \quad (9)$$

Where,

I = Incidence of NCC-associated epilepsy;
P = Prevalence of epilepsy; P ~ Beta(67, 1791);
N = Proportion NCC-associated epilepsy cases; N ~ Uniform(0.236, 0.315);
D = Duration of epilepsy; D ~ Fixed, differs by sex and age group;
M = Mortality due to NCC-associated epilepsy;
C = Case-fatality ratio of epilepsy; C ~ Beta(3.2, 124.8).

Since both incidence and mortality are examples of count data, an underlying gamma distribution was presumed. Monte Carlo simulations were used to generate a compiled distribution of incidence and mortality, based on equations (8) and (9). These compiled distributions were then used to estimate the parameters of the gamma distributions, based on maximum likelihood. This was done using the `fitdistr()` function, available in the **MASS** package (Venables and Ripley, 2002). The results of this process are presented in Figure 3.2.

As the mortality rate presented by Praet et al. (2009) applied to the entire study population, the stratification level was set accordingly (i.e., "None"). Likewise, the "Age" stratification level was selected for the disability weights, as they were only stratified by age category (Figure 3.2).

After entering all required data in the *DALY Calculator*, the following results were obtained (taken from Figure 6):

```
          Mean         Median       95% Credible Interval
DALY:     45965        41789        [ 14454 - 100331 ]
YLD:      6897         6675         [ 3532 - 11524 ]
YLL:      39068        34933        [ 7952 - 93417 ]
Deaths:   1254         1122         [ 256 - 2994 ]

DALY ~ 15% YLD + 85% YLL
```

These results, and their corresponding credible intervals, resemble the values obtained by Praet and colleagues (2009).

5.2. Congenital toxoplasmosis and DALYs in the Netherlands

The second built-in example computes the burden of congenital toxoplasmosis in the Netherlands, as presented by Kortbeek et al. (2009). These authors applied uniform age weighting and no time discounting, and used the Dutch life expectancy table, which has an average life expectancy at birth of 79. DALYs were calculated for the seven most important health outcomes of congenital toxoplasmosis:

- Fetal loss (≥ 24 weeks)
- Clinical symptoms in first year of life:
    - Chorioretinitis
    - Intracranial calcification
    - Hydrocephalus
    - Central nerve system abnormalities
    - Neonatal death
- Asymptomatic at birth:
    - Chorioretinitis later in life

Kortbeek et al. (2009) estimated the most likely burden of congenital toxoplasmosis to be 2,303 DALYs. By adding up the minimum and maximum estimates of the different outcomes, a lower bound of 818 and an upper bound of 6,713 DALYs was obtained. This range represents the possible extremes of the DALYs due to congenital toxoplasmosis, but should not be interpreted as a credible interval.

The only sources of uncertainty included in the authors' calculation model were the incidence and mortality rates of the various outcomes, for which most likely, low and high estimates were presented. In our example, we used a Beta-Pert distribution to represent this uncertainty, and interpreted the "most likely" values as the mean values of the Beta-Pert:

$$\bar{x} = \text{"most likely"}$$

$$m = (6\bar{x} - max - min)/4$$

$$s = \frac{max - min}{6}$$

$$\alpha = \left(\frac{\bar{x} - min}{max - min}\right) * \left\{\frac{(\bar{x} - min) * (max - \bar{x})}{s^2} - 1\right\}$$

$$\beta = \left(\frac{max - \bar{x}}{\bar{x} - min}\right) * \alpha$$

$$Beta\text{-}Pert(m, min, max) = Beta(\alpha, \beta) * (max - min) + min$$

Where,

$\bar{x}$ = mean of the Beta-Pert distribution
m = mode of the Beta-Pert distribution
s = standard deviation of the Beta-Pert distribution
α = shape 1 of the Beta distribution
β = shape 2 of the Beta distribution

The authors did not present different estimates for male and females, therefore the "Age" stratification level was selected for all parameters.

After loading the toxoplasmosis example in the *DALY Calculator*, the life expectancy table has to be adapted, by changing the values for age 0 to 79, for both males and females. The social values will be set automatically to those applied by Kortbeek and colleagues (2009).

The following results were generated by the *DALY Calculator*:

```
        Mean        Median      95% Credible Interval
DALY:   2300        2194        [ 1201 - 3983 ]
YLD:    1247        1151        [ 667 - 2209 ]
YLL:    1053        869         [ 237 - 2528 ]

DALY ~ 54% YLD + 46% YLL
```

Again, the results are similar to those presented by the authors. In addition, the *DALY Calculator* provides a real 95% credible interval, which is, of course, narrower than the range of extremes presented by the authors. Through the **R** environment, it is possible to obtain the extremes of the vector of DALYs generated by the *DALY Calculator*. This can simply be done by applying the `min()` and `max()` function to the object containing the DALY vector:

After setting the data, type the following commands in the **R** Console:

```
x <- getDALY()      # assign DALY Calculator output to object "x"
min(x$DALY)         # get minimum DALY
max(x$DALY)         # get maximum DALY
```

This resulted in a minimum of 646.91 DALYs, and a maximum of 6125.46 DALYs.

6. Saving and loading input data

An additional feature of the *DALY Calculator* is the possibility to save the entered values for the epidemiological parameters to a tab-delimited text file, which can easily be copied into a spreadsheet program for better readability. This function is available through the 'File' menu:

File > Save DALY data to file…

Thus created (and edited) text files can be loaded into the *DALY Calculator* by accessing the 'File' menu and selecting the following:

File > Load DALY data from file…

Finally, a standardized spreadsheet has been made available for setting the epidemiological parameters and their distributions. By converting this file to a tab-delimited text file, the thus set data can easily be loaded into the *DALY Calculator* by calling the above-mentioned function.

7. Current limitations and future work

The current version of the *DALY Calculator* allows basic DALY calculation and uncertainty analysis, which could be useful for a variety of users. However, this first version has some restrictions, which might limit a more advanced performance:

- The current version of the *DALY Calculator* is designed to calculate a basic incidence-based disease model, with well-defined incidence and mortality rates per disease category or outcome. The incorporation of conditional probabilities or the calculation of complex disease models is currently not possible;
- The *DALY Calculator* only allows the simultaneous assessment of up to eight disease categories or outcomes. This is sufficient for most diseases or agents, but could be limiting for certain more complex models;
- Until now, the *DALY Calculator* only allows the simultaneous use of a maximum of five age groups;
- Currently, only two DALY calculation examples are built-in.

Efforts are being made to clear these limitations, and solutions should be provided in the next version of the *DALY Calculator*. In addition, a sensible expansion of the **DALY** package would be the inclusion of sensitivity analysis functions. These will allow assessing the influence of the social values (i.e., age weighting and time discounting) on the final result, and will allow an estimation of the contribution of the input uncertainties to the overall level of uncertainty. Finally, user suggestions will be substantially taken into account during the further development of the *DALY Calculator*.

8. REFERENCES

de Vocht F., Higgerson J., Oliver K., Verma A. (2010). Incorporating uncertainty in aggregate burden of disease measures: an example of DALYs-averted by a smoking cessation campaign in the UK. *Journal of Epidemiology and Community Health*, doi:10.1136/jech.2010.119842.

Kortbeek L.M., Hofhuis A., Nijhuis C.D.M., Havelaar A.H. (2009). Congenital toxoplasmosis and DALYs in the Netherlands. *Memórias de Instituto Oswaldo Cruz 104(2)*, 370-373.

Lopez A.D., Mathers C.D., Ezzati M., Jamison D.T., Murray C.J.L. (2006). *Global burden of disease and risk factors*. Oxford University Press, New York, USA, pp. 475.

Mathers C.D., Vos T., Lopez A.D., Salomon J., Ezzati M. (2001). *National Burden of Disease Studies: a Practical Guide.* Edition 2.0. Global Program on Evidence for Health Policy. World Health Organization, Geneva, Switzerland.

Murray C.J.L. (1994). Quantifying the burden of disease: the technical basis for disability-adjusted life years. *Bulletin of the World Health Organization 72(3)*, 429-445.

Murray C.J.L., Lopez A.D. (1996). *The Global Burden of Disease: a comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020*. Harvard University Press, Cambridge, UK, pp. 990.

Murray C.J.L, Acharya A.K. (1997). Understanding DALYs. *Journal of Health Economics 16(6)*, 703-730.

Praet N., Speybroeck N., Manzanedo R., Berkvens D., Nforninwe D.N., Zoli A., Quet F., Preux P.-M., Carabin H., Geerts S. (2009). The disease burden of *Taenia solium* cysticercosis in Cameroon. *PLoS Neglected Tropical Diseases 3(3)*, e406.

R Development Core Team (2010). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Venables W.N., Ripley B.D. (2002). *Modern Applied Statistics with S*. Fourth Edition. Springer, New York, USA. ISBN 0-387-95457-0.