

Document downloaded from:

<http://hdl.handle.net/10251/35672>

This paper must be cited as:

Cortés López, JC.; Villanueva Micó, RJ.; Sánchez-Sánchez, A.; Santonja, F. (2011). epiModel: A system to build automatically systems of differential equations of compartmental type-epidemiological models. *Computers in Biology and Medicine*. 41(11):999-1005. doi:10.1016/j.combiomed.2011.06.018.



The final publication is available at

<http://dx.doi.org/10.1016/j.combiomed.2011.06.018>

Copyright Elsevier

epiModel: A system to build automatically systems of differential equations of compartmental type-epidemiological models

J.-C. Cortés^a, Almudena Sánchez-Sánchez^{a,*}, Francisco-J. Santonja^b,
Rafael-J. Villanueva^a

^a*Instituto Universitario de Matemática Multidisciplinar, Edificio 8G piso 2, Universitat Politècnica de València, 46022 Valencia (Spain)*

^b*Departamento de Estadística e Investigación Operativa, Universidad de Valencia, Valencia (Spain)*

Abstract

In this paper we describe *epiModel*, a code developed in *Mathematica* that facilitates the building of systems of differential equations corresponding to type-epidemiological models whose characteristics are defined in text files following an easy syntax. It includes the possibility to obtain the equations of models involving age and/or sex groups.

Keywords: Modelling, Type-epidemiological model, Automatic building of systems of differential equations, *Mathematica*

1. Introduction

From the celebrated paper of Kermack and McKendrick in 1927 [1] epidemiologists and mathematicians have been used mathematical models to understand the transmission dynamics of diseases. The advances in this area led to more complex models and, therefore, larger systems of differential equations. For instance, the model developed in [2] for the study of the spread

*Corresponding Author.

Email addresses: jccortes@imm.upv.es (J.-C. Cortés),
alsncsnc@posgrado.upv.com (Almudena Sánchez-Sánchez),
francisco.santonja@uv.es (Francisco-J. Santonja), rjvillan@imm.upv.es (Rafael-J. Villanueva)

of Human Papillomavirus (HPV) consists of more than 7,000 equations, or the model described in [3] for the dynamics of meningococcal disease with around nine hundred equations.

The compartmental diagrams have helped the developing of the epidemiological models and their translation into differential equations. However, when the models include a lot of subpopulations and consider age and/or sex groups, the building of the system of differential equations turns difficult when handling a large number of functions and parameters.

It has been developed software to help researchers to model, study, simulate and analyse the models, for instance MathModelica [4] mainly addressed for Engineering purposes. Other as Berkeley Madonna [5], *Mathematica* [6], MATLAB [7], MLAB [8] or even Grapher [9] among others, are powerful tools once the user provides them the system of differential equations to be studied.

Thus, with the aim to facilitate researchers in the epidemiology area to build model equations, in this paper we present *epiModel*, a code developed in *Mathematica* capable to generate automatically the system of differential equations and its parameters from a short and easy description of the model contained in a text file.

epiModel consists of three files (see scheme in Figure 1):

- "ModelDefinition". This file is where the user describes the characteristics of the model using a simple syntax explained in Section 2.
- "epiModel". This file contains the code that enable the transformation of the data model in "ModelDefinition" into a system of differential equations.
- "ModelBuilder.nb". This is a *Mathematica* file that loads the files "ModelDefinition" and "epiModel" and execute them in order to generate two new files:
 - "Model.data" with the system of differential equations and
 - "parameters.data" with a list of all the model parameters.

The paper is organised as follows. In Section 2 we describe how to build the file "ModelDefinition". Once the "ModelDefinition" has been built, in Section 3 we explain how to generate the files "Model.data" and "parameters.data". In Section 4, three examples are presented: in the first, we

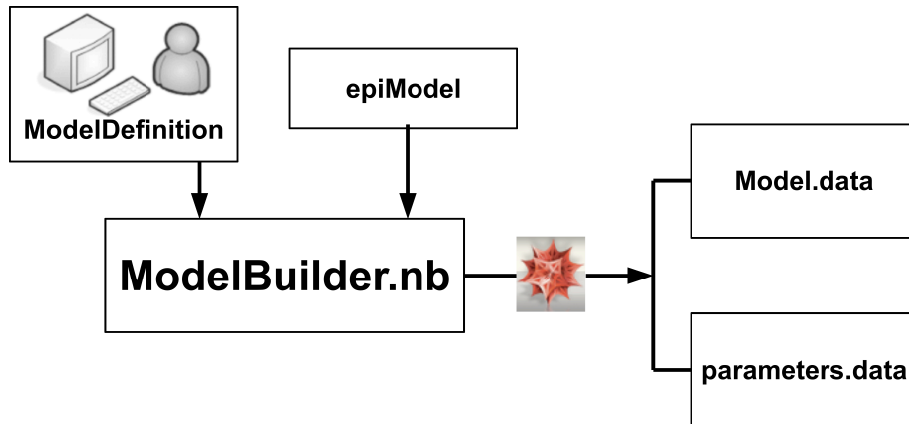


Figure 1: Process of how *epiModel* works. "ModelBuilder.nb" loads data from "ModelDefinition" and "epiModel" creates "Model.data" and "parameters.data".

generate the system of differential equations corresponding to a typical SIRS model; in the second, the same is done for a SIR model with two age groups; finally, the equations for a SIR model with two age and two sex groups are generated in the last example. In Section 5, conclusions are given.

It is not the objective of this paper to explain the code line by line, but we must say that it has been developed following a slight improvement of the idea of Capasso [10, 11] about how to represent an epidemiological model in matrix form. In fact, when a model is generated, the file "Model.data" contains the system of equations and the matrices corresponding to the matrix form of the model.

Then, if the vector $z(t) = (z_1(t), \dots, z_n(t))$ contains as entries the model subpopulation functions and we denote by

$$diag(z(t)) = \begin{pmatrix} z_1(t) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & z_n(t) \end{pmatrix},$$

any compartmental model (even including age and/or sex groups) can be written as

$$\frac{dz(t)}{dt} = c + Lz(t) + A diag(z(t)) Bz(t), \quad (1)$$

where c is a vector of size n and L , A and B are matrices of size $n \times n$. c contains parameters corresponding to the model independent terms, L to the linear terms and A and B to the model non-linear terms, placed in the appropriate coordinates.

The difference with the Capasso's idea is the inclusion of matrix A that allows us to take several subpopulations (usually the same in different age/sex groups) as a part of the same transmission (non-linear) term. In Section 4, in the first and second example, we will provide the obtained matrices c , L , A and B .

epiModel is available at <http://www.imm.upv.es>.

2. How to build the file "ModelDefinition"

This is a text file and consists of three parts: a general variable; the definition of subpopulations; and the definition of the parameters. Note that the syntax of this file should fit the *Mathematica* syntax.

It is important to preserve the names of the variables defined below (\backslash [NTilde], **SP**, **TI**, **LIN**, **NOLIN**, x) because they will be called by "epi-Model".

2.1. General variable

This variable indicates the structural characteristics of the model, i.e., the age groups. The variable is

	Name	Value	Description
1.	\backslash [NTilde]	number	Number of age groups.

2.2. Definition of the subpopulations

Data corresponding to subpopulations are stored in a list named **SP**. Each row of the list consists of the following fields

	Name	Value	Description
1.	Number	Number	Subpopulation ID number.
2.	Description	String	Name of the subpopulation.

In the following example, three subpopulations of a model are defined.

```

SP = {
{1, "Susceptible"},
{2, "Infected" },
{3, "Recovered" }
};

```

2.3. Defining Parameters

The parameters have two ways to be classified. The first depends on the part of the model to which they contribute:

- Those that are included in the independent term of the model and they will be in the list **TI**.
- Those that are linear terms of the model and they will be in the list **LIN**.
- Those that are part of the non-linear term of the model and will be stored in the list **NOLIN**.

The second classification depends on the type of parameter. To explain this, we should note that compartmental models are illustrated by diagrams where the boxes represent the subpopulations and the arrows represent the terms involving the model parameters. However, not all the arrows are equal: some only enter in a box; some of them only exist from a box; others exist from a box and enter in another; some of the latter are especial because connect the same box for different age groups (see Figure 2). These four possibilities lead to the second parameters classification:

- **Type 1:** also called *birth type*, because this parameter comes from arrows that only enters in a subpopulation, as the newborns do.
- **Type 2:** also called *death type*, because it is related to an arrow that only exists from a subpopulation, like dead people leaving the system.
- **Type 3:** also called *input-output type*, because this parameter measures the flow from one box or subpopulation to another, for instance, the disease transmission or the recovering illness average time.
- **Type 4:** also called *between ages type*, because this parameter is related to the population growth and connects the same box in two consecutive age groups.

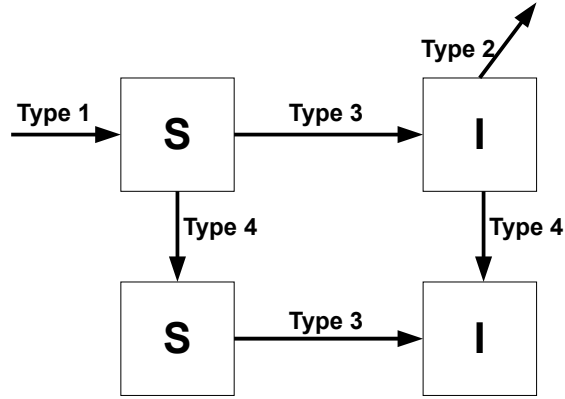


Figure 2: Parameter types depending on where the arrows enter and exit in compartmental models.

2.3.1. Parameters of independent term and linear term

Each independent term or linear term parameter is encoded with a list containing the following fields:

	Name	Value	Description
1.	Name		Parameter name.
2.	Type	1, 2, 3, 4	Parameter type.
3.	The arrow exits from the subpopulations	$\{n1, n2, \dots\}$	Subpopulations from which the arrow related to the current parameter, exits. The list $\{\}$ means that the arrow does not exit from any box (e.g., a birth type parameter) .
4.	The arrow enters into the subpopulations	$\{n1, n2, \dots\}$	Subpopulations from which the arrow related to the current parameter, enters. The list $\{\}$ means that the arrow does not enter into any box (e.g., a death type parameter).
5.	Depending on the age group?	True/False	True means that this parameter can be different depending on the age group.
6.	Description	String	Description of the parameter.

It should be taken into account the following considerations concerning with the definition of parameters:

- Any parameter should not be named x because this variable defines the subpopulations.
- A parameter of **Type 3** can not have common elements in the fields 3 and 4 of the above table. These situations can be avoided by defining various parameters properly.
- If a parameter does not depend on age groups, i.e., the field 5 in the table is False, it will appear only in the first age group between the boxes included in the lists of the fields 3 and/or 4.
- **Type 4** parameters cannot appear in the variable **TI**.
- The **Type 4** parameters have always to be dependent on the age group.

In order to avoid confusions, it is not convenient to use the same or similar variable names for different parameters. Now, let us show some examples of how to encode parameters corresponding to independent and linear model terms:

- **Type 1:** Suppose that the newborns enters directly to subpopulation 1 at a rate $\backslash[\text{Mu}]$. The model does not consider age groups. Then, this term is encoded as

```
{\[Mu], 1, {}, {1}, False, "Birth rate"}
```

- **Type 2:** Now, we consider, an age group model with three subpopulations where the death rate depends on the age group and all people of any subpopulation is susceptible to die. This is encoded as:

```
{d, 2, {1,2,3}, {}, True, "Death rate"}
```

- **Type 3:** Let us suppose that, after recovering from a disease, the individuals have an average temporary immunity $\backslash[\text{Gamma}]$ in a typical SIRS model with age groups. This parameter will be encoded as:

```
{\[Gamma], 3, {3}, {1}, True, "Average immunity time"}
```


- **Type 4:** In an age group model with three subpopulations, the growth rate c is encoded as follows:

`{c, 4, {1,2,3}, {}, True, "Growth rate"}`

2.3.2. Parameters of non-linear term

Each non-linear term parameter is encoded using the following fields:

	Name	Value	Description
1.	Name		Parameter name.
2.	Transmission is affected by the subpopulations	$\{n1, n2, \dots\}$	Those subpopulations related to any infectious state (infectious, latent, etc.).
3.	The arrow exits from the subpopulation	$\{n1\}$	Subpopulation whose individuals are susceptible to be infected.
4.	The arrow enters into the subpopulation	$\{n2\}$	Subpopulation where an infected individual enters (latent, infectious, etc.).
5.	Depending on the age groups?	True/False	True means that this parameter can be different depending on the age group.
6.	Description	String	Description of the parameter.

The same advices done for independent and linear parameters can be applied to non-linear ones. In the following example, the list

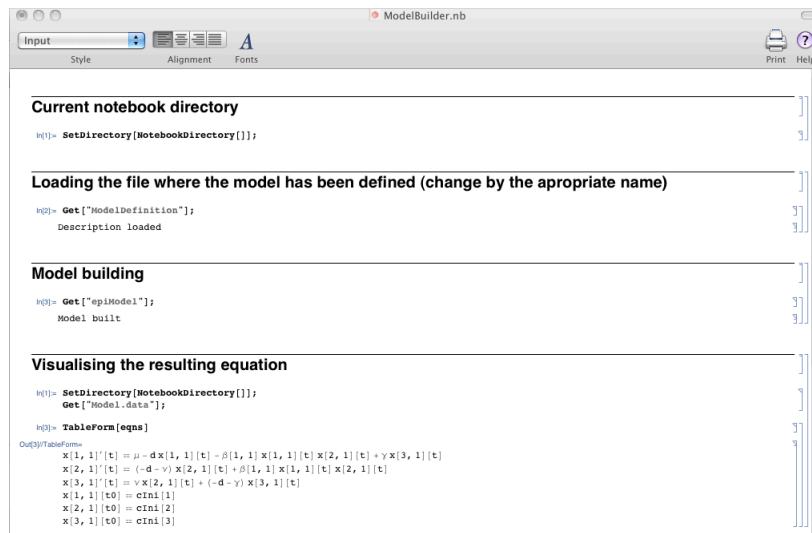
`{\[Beta], {3,4}, {1}, {2}, True, Transmission rate}`

indicates that if an individual A in subpopulation 1 of any age group has a successful contact (the disease is transmitted) with another individual belonging to subpopulations 3 or 4 of any age group, individual A moves to subpopulation 2 in the same age group as A was previously.

3. Steps to build the system of differential equations

As we said before, the code was developed using *Mathematica* [6]. Then, in order to build the system of differential equations, we need to have *Mathematica* installed in the computer. Moreover, the files "ModelDefinition", "epiModel" and "ModelBuilder.nb" have to be in the same directory. Then, open the notebook "ModelBuilder.nb" using *Mathematica*. This notebook consists of 4 cells.

- 1st. This cell allows to set the current directory (the directory containing the files to build the model) as the working directory.
- 2nd. The second cell loads the text file "ModelDefinition" where we have defined the model, following the rules described in Section 2.
- 3rd. The third cell loads "epiModel" and executes it. Then, the files "Model.data" and "parameter.data" appear in the working directory.
- 4th. Once the system of equations is in the "Model.data" file, this cell loads this file and displays the resulting system of equations. This is useful to verify that no errors have occurred and check the correctness of the equations.



```
Input
Style Alignment Fonts Print Help

Current notebook directory
H01) SetDirectory[NotebookDirectory[]];

Loading the file where the model has been defined (change by the appropriate name)
H02) Get["ModelDefinition"];
Description loaded

Model building
H03) Get["epiModel"];
Model built

Visualising the resulting equation
H04) SetDirectory[NotebookDirectory[]];
Get["Model.data"];
H05) TableForm[eqns]

Out[3]TableForm:
x[1, 1][t] = μ - d x[1, 1][t] - β[1, 1] x[1, 1][t] x[2, 1][t] + γ x[3, 1][t]
x[2, 1][t] = (-d - γ) x[2, 1][t] + β[1, 1] x[1, 1][t] x[2, 1][t]
x[3, 1][t] = ν x[2, 1][t] + (-d - γ) x[3, 1][t]
x[1, 1][t0] = cIni[1]
x[2, 1][t0] = cIni[2]
x[3, 1][t0] = cIni[3]
```

Figure 3: Screenshot of "ModelBuilder.nb" in *Mathematica*.

In the case that no errors appeared, two files will be generated: "Model.data" and "parameters.data".

3.1. The file "Model.data"

This file contains the variables ti , mc , $mcNoL1$, $mcNoL2$, $Fvars$, $eqns$, $vars$.

- ti corresponds to the vector independent term of the model, that is, vector c in expression (1).
- mc corresponds to the coefficient matrix of the linear model, i.e, matrix L in (1).
- $mcNoL1$ and $mcNoL2$ are the coefficient matrices which allow the construction of the non-linear part of the model, that is, matrices A and B in (1), respectively.
- $Fvars$ is a vector function where each entry corresponds to each sub-population in the model.
- $eqns$ is a list with the system of differential equations. It is built computing the expression (1) using all the above matrices.
- $vars$ is the same as $Fvars$ but removing t in the functions.

Thus, if we want to solve numerically in *Mathematica* a model (system of differential equations) we execute

```
sol = NDSolve[ eqns, vars, {t, t0, tEnd} ]
```

and in order to evaluate and draw the solutions we can execute

```
Plot[ Evaluate[ Fvars /. sol ], {t, t0, tEnd} ]
```

3.2. The file "parameters.data"

This file has a complete list of the parameters appearing in the model. This is useful because some of them can be replaced by known values and the ones that do not can be included in a procedure to be estimated.

4. Examples

In this section let us show three examples with different options to build type-epidemiological models.

4.1. SIRS model

The first example is the classical SIRS model, where we have three subpopulations: Susceptible (S), Infectious (I) and Recovered (R). The transmission is carried out with effective contacts between a susceptible individual and an infectious individual. Once an individual has been infected he/she recovers and acquires a temporal immunity. When this finishes, the individual becomes again susceptible. This description has been depicted in Figure 4.

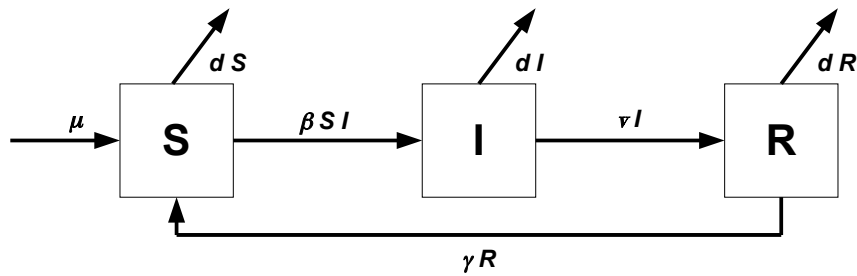


Figure 4: Diagram of a Susceptible-Infectious-Recovered-Susceptible model.

To build the "ModelDefinition" file, it should be taken into account that:

- It is a model without age groups.
- μ is a Type 1 parameter belonging to the independent term model, because it represents the newborns that enter directly into the susceptible subpopulation.
- Parameter d is the death rate (Type 2) and exits from all the subpopulations.
- β is the transmission parameter belonging to the non-linear term.
- The Type 3 parameters γ and ν belong to linear terms. They are the average time of infection and immunity.

Note that the number of age groups indicated is 1 and there are not parameters depending on the age groups. In this way, the "ModelDefinition" file will be as follows:

```

(* Number of age groups *)
\[NTilde] = 1;

(* Subpopulations *)
SP = {
{1, "Susceptible"},
{2, "Infected" },
{3, "Recovered" }
};

(* INDEPENDENT TERM *)
TI = {
{\[Mu], 1, {}, {1}, False, "Birth rate"}
};

(* LINEAR TERM *)
LIN = {
{ \[Nu], 3, {2}, {3}, False, "Average time of infection"},
{\[Gamma], 3, {3}, {1}, False, "Average time of immunity"},
{ d, 2, {1,2,3}, {}, False, "Death rate"}
};

(* NON LINEAR TERM *)
NOLIN = {
{\[Beta], {2}, {1}, {2}, False, "Transmision rate"}
};

```

After running "epiModel" and building the model from the above data in the "ModelDefinition" file, *Mathematica* returns the following system of differential equations ($x[1, 1]$ is the susceptible, $x[2, 1]$ the infectious and $x[3, 1]$ the recovered subpopulations):

$$\begin{aligned}
x[1, 1]'[t] &== \mu - dx[1, 1][t] - \beta x[1, 1][t]x[2, 1][t] + \gamma x[3, 1][t] \\
x[2, 1]'[t] &== (-d - \nu)x[2, 1][t] + \beta[1, 1]x[1, 1][t]x[2, 1][t] \\
x[3, 1]'[t] &== \nu x[2, 1][t] + (-d - \gamma)x[3, 1][t] \\
x[1, 1][t_0] &== cIni[1] \\
x[2, 1][t_0] &== cIni[2] \\
x[3, 1][t_0] &== cIni[3]
\end{aligned}$$

and the matrices that, computing the expression (1), enable the construction of the above system are:

$$c = \begin{pmatrix} \mu \\ 0 \\ 0 \end{pmatrix}, \quad L = \begin{pmatrix} -d & 0 & \gamma \\ 0 & -d - \nu & 0 \\ 0 & \nu & -d - \gamma \end{pmatrix},$$

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & \beta & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

4.2. SIR model with two age groups

This is a typical SIR (Susceptible-Infectious-Recovered) model with two age groups. We have two susceptible groups, S_1 and S_2 , one for each age group, the same for infectious, I_1 and I_2 and for recovered R_1 and R_2 . An individual in S_1 , I_1 or R_1 grows and can enter in the corresponding box of the 2nd age group, S_2 , I_2 or R_2 , respectively. The transmission is carried out with effective contacts between a susceptible individual and an infectious individual of any age group. Once an individual has been infected, after some time, recovers. This description has been depicted in Figure 5.

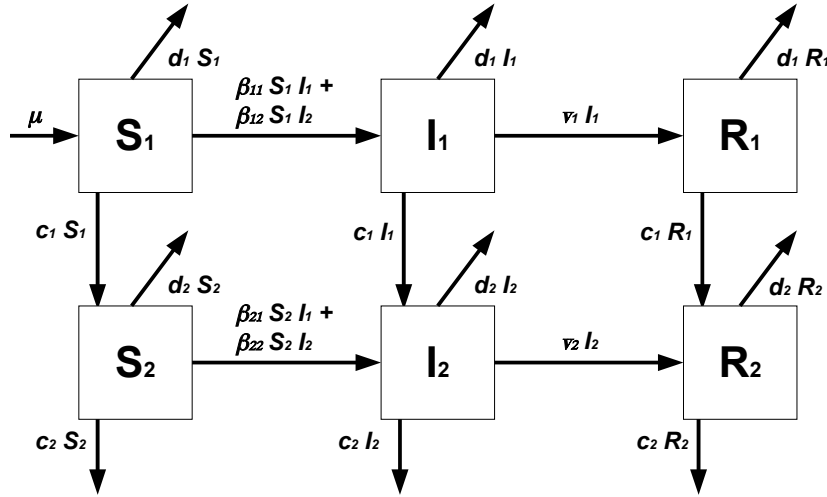


Figure 5: Diagram of a Susceptible-Infectious-Recovered model with two age groups.

To build the "ModelDefinition" file, we take into account that:

- It is a model with 2 age groups.
- μ is a Type 1 parameter belonging to the independent term model that enters directly into the susceptible subpopulation of the first age group, i.e., it does not depend on the age group.
- β is the transmission parameter belonging to the non-linear term. "epi-Model" will generate four different parameters ($\beta[1, 1]$, $\beta[1, 2]$, $\beta[2, 1]$ and $\beta[2, 2]$) depending on the crossed products between susceptible and infectious subpopulations.
- The Type 2 death parameter d depends on the age group. "epiModel" will generate $d[1]$ parameter for the first age group and $d[2]$ for the second one.
- The Type 3 parameter ν belongs to linear term. It is the average recovery time and also depends on the age group.
- Parameter c is Type 4 and is the population growth rate between these age groups.

Thus, the "ModelDefinition" file will be as follows:

```
(* Number of age groups *)
\NTilde = 2;

(* Subpopulations *)
SP = {
{1, "Susceptible"},
{2, "Infected" },
{3, "Recovered" }
};

(* INDEPENDENT TERM *)
TI = {
{\Mu}, 1, {}, {1}, False, "Birth rate"
};

(* LINEAR TERM *)
```

```

LIN = {
{ \[Nu], 3, {2}, {3}, True, "Average time of infection"},
{ d, 2, {1,2,3}, {}, True, "Death rate"},
{ c, 4, {1,2,3}, {}, True, "Growth rate"}
};

```

(* NON LINEAR TERM *)

```

NOLIN = {
{\[Beta], {2}, {1}, {2}, False, "Transmision rate"}
};

```

Then, running "epiModel" and building the model from the above data in the "ModelDefinition" file, *Mathematica* returns the following system of differential equations ($x[1, 1]$, $x[2, 1]$ and $x[3, 1]$ are the susceptible, infectious and recovered subpopulations, respectively, of the first age group, and $x[1, 2]$, $x[2, 2]$, $x[3, 2]$ for the second one):

$$\begin{aligned}
x[1, 1]'[t] &== \mu + (-c[1] - d[1])x[1, 1][t] \\
&\quad -\beta[1, 1]x[1, 1][t]x[2, 1][t] - \beta[1, 2]x[1, 1][t]x[2, 2][t] \\
x[2, 1]'[t] &== (-c[1] - d[1] - \nu[1])x[2, 1][t] \\
&\quad +\beta[1, 1]x[1, 1][t]x[2, 1][t] + \beta[1, 2]x[1, 1][t]x[2, 2][t] \\
x[3, 1]'[t] &== \nu[1]x[2, 1][t] + (-c[1] - d[1])x[3, 1][t] \\
x[1, 2]'[t] &== c[1]x[1, 1][t] + (-c[2] - d[2])x[1, 2][t] \\
x[2, 2]'[t] &== c[1]x[2, 1][t] + (-c[2] - d[2] - \nu[2])x[2, 2][t] \\
x[3, 2]'[t] &== \nu[2]x[2, 2][t] + c[1]x[3, 1][t] + (-c[2] - d[2])x[3, 2][t] \\
x[1, 1][t_0] &== cIni[1] \\
x[2, 1][t_0] &== cIni[2] \\
x[3, 1][t_0] &== cIni[3] \\
x[1, 2][t_0] &== cIni[4] \\
x[2, 2][t_0] &== cIni[5] \\
x[3, 2][t_0] &== cIni[6]
\end{aligned}$$

and the matrices that permit the construction of the above system, computing the expression (1), are:

$$L = \begin{pmatrix} -c[1] - d[1] & 0 & 0 & 0 & 0 & 0 \\ 0 & -c[1] - d[1] - \nu[1] & 0 & 0 & 0 & 0 \\ 0 & \nu[1] & -c[1] - d[1] & 0 & 0 & 0 \\ c[1] & 0 & 0 & -c[2] - d[2] & 0 & 0 \\ 0 & c[1] & 0 & 0 & -c[2] - d[2] - \nu[2] & 0 \\ 0 & 0 & c[1] & 0 & \nu[2] & -c[2] - d[2] \end{pmatrix},$$

$$c = \begin{pmatrix} \mu \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad A = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & \beta[1,1] & 0 & 0 & \beta[1,2] & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \beta[2,1] & 0 & 0 & \beta[2,2] & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Note that the model includes parameter $c[2]$. This parameter has sense if people in the second age group leave the system in a form different from death, otherwise, $c[2]$ would be redundant because $d[2]$ plays the same role and then, $c[2]$ should be zero.

4.3. SIR model with two age groups and two sexes

”epiModel” is not designed to build systems of differential equations from gender models, however, considering some tricks, we can transform a typical age group model into an age group and gender model. These tricks are:

- Consider first the age groups for females and then for males. Then, if there are n age groups for each sex, we should consider a model with $2n$ age groups ($\backslash[\text{NTilde}]=2n$).
- The growth parameter $c[n]$ connecting the last female age group (group n) with the first male age group (group $n + 1$) is zero.
- Birth rate should be considered age depending in order to take into account different birth rates for females and males. Then, Birth rate, say $\mu[i]$ is a Type 1 parameter depending on age group. However, newborns enter in the first age group, group 1 for females ($\mu[1]$) and group $n + 1$ for males ($\mu[n + 1]$). The remainder $\mu[i]$ are zero.

This example consists on a typical SIR (Susceptible-Infectious-Recovered) model with two age groups and two sexes for each age group. We have two susceptible female groups, S_1 and S_2 , and two susceptible male groups, S_3 and S_4 , one for each age group. The same for infectious, I_1, I_2, I_3 and I_4 , and for recovered, R_1, R_2, R_3 and R_4 . An individual in $S_1, I_1, R_1, S_3, I_3, R_3$ grows and can enter into the box $S_2, I_2, R_2, S_4, I_4, R_4$ of the 2nd age group, respectively. People leave the system by death.

For this example, let us suppose that the disease considered is heterosexually transmitted. The transmission is carried out with effective contacts between a susceptible male or female and an infectious individual of any age group of the other sex. Once an individual has been infected, after some time, recovers. To build the "ModelDefinition" file, we take into account that:

- It is a model with 4 age groups, two for females and two for males.
- μ is a Type 1 parameter belonging to the independent term model that enters directly into the susceptible subpopulations of the first age group for males and females. This requires that μ depends on the age group and $\mu[2] = \mu[4] = 0$.
- β is the transmission parameter belonging to the non-linear term. "epiModel" will generate $\beta[i, j]$ for $i, j = 1, 2, 3, 4$, one for each type of contact. Taking into account that this disease is heterosexually transmitted, parameters $\beta[1, 1]$, $\beta[1, 2]$, $\beta[2, 1]$, $\beta[2, 2]$, $\beta[3, 3]$, $\beta[3, 4]$, $\beta[4, 3]$ and $\beta[4, 4]$ are zero.
- The Type 2 death parameter d depends on the age group. "epiModel" will generate $d[i]$ parameters for $i = 1, 2, 3, 4$, the two first for the female age groups and the remainder for male age groups.
- The Type 3 parameter ν belongs to linear term. It is the average recovery time and also depends on the age group and gender.
- Parameter c is of Type 4 and it denotes the population growth rate between these age groups. As we mentioned before, $c[2] = 0$. Moreover, as people leave the system by death, $c[4] = 0$ because plays the same role as the death parameter $d[4]$.

Thus, the "ModelDefinition" file will be as follows:

```
(* Number of age groups *)
\NTilde = 4;

(* Subpopulations *)
SP = {
{1, "Susceptible"},
```

```

{2, "Infected" },
{3, "Recovered" }
};

(* INDEPENDENT TERM *)
TI = {
{\[Mu], 1, {}, {1}, True, "Birth rate"}
};

(* LINEAR TERM *)
LIN = {
{ \[Nu], 3, {2}, {3}, True, "Average time of infection"},
{ d, 2, {1,2,3}, {}, True, "Death rate"},
{ c, 4, {1,2,3}, {}, True, "Growth rate"}
};

(* NON LINEAR TERM *)
NOLIN = {
{\[Beta], {2}, {1}, {2}, False, "Transmission rate"}
};

```

Then, running "epiModel", files "Model.data" and "parameters.data" appear. Now, in order to obtain the desired system of differential equations, we have to assign the following values to parameters: $c[2] = c[4] = 0$, $\mu[2] = \mu[4] = 0$, $\beta[1, 1] = \beta[1, 2] = 0$, $\beta[2, 1] = \beta[2, 2] = 0$, $\beta[3, 3] = \beta[3, 4] = 0$ and $\beta[4, 3] = \beta[4, 4] = 0$. Thus, we obtain the system:

$$\begin{aligned}
x[1, 1]'[t] &== \mu[1] + (-c[1] - d[1])x[1, 1][t] \\
&\quad -\beta[1, 3]x[1, 1][t]x[2, 3][t] - \beta[1, 4]x[1, 1][t]x[2, 4][t] \\
x[2, 1]'[t] &== (-c[1] - d[1] - \nu[1])x[2, 1][t] \\
&\quad +\beta[1, 3]x[1, 1][t]x[2, 3][t] + \beta[1, 4]x[1, 1][t]x[2, 4][t] \\
x[3, 1]'[t] &== \nu[1]x[2, 1][t] + (-c[1] - d[1])x[3, 1][t] \\
x[1, 2]'[t] &== c[1]x[1, 1][t] - d[2]x[1, 2][t] \\
x[2, 2]'[t] &== c[1]x[2, 1][t] + (-d[2] - \nu[2])x[2, 2][t] \\
x[3, 2]'[t] &== \nu[2]x[2, 2][t] + c[1]x[3, 1][t] - d[2]x[3, 2][t] \\
x[1, 3]'[t] &== \mu[3] + (-c[3] - d[3])x[1, 3][t] \\
x[2, 3]'[t] &== (-c[3] - d[3] - \nu[3])x[2, 3][t] \\
x[3, 3]'[t] &== \nu[3]x[2, 3][t] + (-c[3] - d[3])x[3, 3][t] \\
x[1, 4]'[t] &== c[3]x[1, 3][t] - d[4]x[1, 4][t] \\
x[2, 4]'[t] &== c[3]x[2, 3][t] + (-d[4] - \nu[4])x[2, 4][t] \\
x[3, 4]'[t] &== \nu[4]x[2, 4][t] + c[3]x[3, 3][t] - d[4]x[3, 4][t] \\
x[1, 1][t_0] &== cIni[1] \\
x[2, 1][t_0] &== cIni[2] \\
x[3, 1][t_0] &== cIni[3] \\
x[1, 2][t_0] &== cIni[4] \\
x[2, 2][t_0] &== cIni[5] \\
x[3, 2][t_0] &== cIni[6] \\
x[1, 3][t_0] &== cIni[7] \\
x[2, 3][t_0] &== cIni[8] \\
x[3, 3][t_0] &== cIni[9] \\
x[1, 4][t_0] &== cIni[10] \\
x[2, 4][t_0] &== cIni[11] \\
x[3, 4][t_0] &== cIni[12]
\end{aligned}$$

where $x[i, j][t]$ is the subpopulation susceptible for $i = 1$, infectious for $i = 2$ and recovered for $i = 3$, and age group 1 females for $j = 1$, age group 2 females for $j = 2$, age group 1 males for $j = 3$ and age group 2 males for $j = 4$.

5. Conclusion

In this paper, we present a *Mathematica* code that translates the description of a type-epidemiological compartmental model in a simple syntax into

a system of differential equations. The obtained system can be used to estimate parameters, simulate different scenarios or predict short and long-term behaviour.

This code is easy to use, save time building the systems and avoid errors. Moreover, it can be applied to models involving age groups and/or gender. It is particularly interesting when we have to handle a large number of groups. You can test it changing in any of the examples in Section 4, the variable `\[NTilde]` by 100 (one hundred one-year age groups), appearing 300 equations.

epiModel is available at <http://www.imm.upv.es>.

Acknowledgement This work has been partially supported by the FIS PI-10/01433, the Spanish Government MICINN and FEDER grant MTM2009-08587 and the Universidad Polit3cnica de Valencia grant PAID-06-09 (ref. 2588).

- [1] W.O. Kermack, A.G. McKendrick, Contributions to the mathematical theory of epidemics-I, Proc. Roy. Soc. 115A (1927) 700-721. Reprinted in Bull. Math. Biol. 53(1-2) (1991) 33-55.
- [2] E.H. Elbasha, E.J. Dasbach, R.P. Insinga, Model for Assessing Human Papillomavirus Vaccination Strategies, Emerg Infect Dis. 13(1) (2007) 2841.
- [3] C.L. Trotter, N.J. Gay, W.J. Edmunds, The natural history of meningococcal carriage and disease, Epidemiol Infect. 134(3) (2006) 556-66.
- [4] <http://www.mathcore.com/products/mathmodelica/>
- [5] <http://www.berkeleymadonna.com/>
- [6] <http://www.wolfram.com/products/mathematica/index.html>
- [7] <http://www.mathworks.com/>
- [8] <http://www.civilized.com/>
- [9] <http://en.wikipedia.org/wiki/Grapher>
- [10] E. Beretta, V. Capasso, On the general structure of epidemic systems. Global asymptotic stability, Comp. Math. Appl. 12A (1986) 677-694.

- [11] V. Capasso, *Mathematical Structures of Epidemic Systems*, Lecture Notes in Biomathematics, Springer, 1993.