Calculation of Prevalence Estimates through Differential Equations: Application to Stroke-Related Disability

Javier Mar\textsuperscript{a,b} María Sainz-Ezkerra\textsuperscript{b} Jose Antonio Moler-Cuiral\textsuperscript{c}

\textsuperscript{a}Clinical Management Unit, and \textsuperscript{b}Research Unit, Hospital Alto Deba, Mondragon, and \textsuperscript{c}Public University of Navarra, Pamplona, Spain

Introduction

As Rita Levi-Montalcini, the 1986 recipient of the Nobel Prize for Medicine, comments in a recent report from the World Health Organization, the burden of neurological disorders at the beginning of the third millennium is reaching a significant proportion in countries with a growing percentage of the population more than 65 years old. Thus, in 2005, neurological disorders constituted 6.3% of the global burden of disease [1]. Prevalence represents the frequency of a given disease in the population under study at a certain point in time [2]. Given that neurological disease causes disability, the measurement of its prevalence is a key point in the assessment of its meaning in our society. The states, or levels, of disability caused by neurological diseases vary, but the particular needs for each state of disability require both specific treatment and services. Thus, knowledge of the prevalence becomes essential for planning and management. Furthermore, the ageing of the present population implies that the prevalence of such diseases will increase in the near future, and consequently the burden of those diseases will increase [1, 3]. This shift of resource expenditure towards older people will also be inevitable for less developed countries [4].

Current figures on the prevalence of disabling neurological diseases, such as dementia, come from popula-
tion-based studies. However, experts have suggested that information about dementia is widely scattered and poorly accessible. In part, this is the reason for the inadequate attention given to dementia [5]. For some regions, data about nonfatal health states related to neurological conditions are extremely poor. Without more complete data on such nonfatal states, there may be an absence of prevalence estimates, which, as Murray and Lopez [3] note, often fosters the tacit assumption that there is no problem, especially when compared to the information available about life lost due to premature mortality. The cost and difficulties of carrying out population surveys becomes an insuperable hurdle in many countries. Under these conditions, the role of epidemiological models is crucial in order to obtain estimates of the prevalence and, consequently, the burden of neurological disease [6].

In medical science, models have been used to represent the natural history of diseases and calculate epidemiological outcomes of interest, such as survival, prevalence, and mortality rates. Depending on the goal, different mathematical approaches have been applied [6, 7]. Markov models have been used exhaustively in the economic evaluation of medical interventions [8, 9]. However, the authors of these studies have pointed out the limitation of the models in the management of time. Markov models make a discrete use of time by representing the transition between states in the natural history of the disease of interest alongside a number of cycles (time horizon). Models based on other approaches, such as discrete event simulation or differential equations, make possible a continuous use of time and have been proposed as alternatives to Markov models [10–13].

Dynamic models based on differential equations have been used chiefly in association with infectious diseases [14]. However, they have also been applied to the estimation of epidemiological parameters as in the Global Burden of Disease Study by the World Health Organization and Harvard University. In this study, a generic mathematical disease model called DisMod was developed to exploit the causal relations between the variables (incidence, prevalence, case fatality, and remission) that describe a disease process. It has been extensively used as a tool to obtain epidemiologic parameters, especially in settings where data collection is not an option, as well as to check the internal consistency of existing estimates [15, 16]. Since DisMod allows the modeling of diseases with only a single prevalent state, a natural extension is the development of a model for those diseases with more than one prevalent state. Stroke is such a disease, because it can cause different levels of patient dependence. In the planning of services and management of resources, it is essential to distinguish clearly among the prevalent states. An example of a non-neurological disease with a natural history of more than one prevalent state is hepatitis C, which includes chronic hepatic disease and its complications [17].

In this study, we will focus on stroke. Globally, it represents one of the greatest public health problems, since it is the third leading cause of death and the leading cause of permanent disability in both sexes [18]. In the United States, more than 3 million people have survived a stroke [19], and in 2004, the economic consequences of stroke in terms of health care costs and loss of productivity were estimated to be USD 53.6 billion [20]. The reason for this huge burden is the evolution of patients’ disability. More precisely, 33% of those who have a stroke die in the first year after the acute event, 20–30% become dependent on assistance for everyday activities because of the aftereffects of the stroke, and the remainder (40–50%) fully recover [18, 21].

The objective of our study is the development and resolution of a system of differential equations that represents the natural history of stroke. With the use of data from the male population in the Basque Country in Spain, this approach will make possible the calculation of the prevalence of the autonomous and dependent states of health that are a consequence of stroke.

**Methods**

Differential equations are widely used in such fields as physics and biology to model the evolution of certain processes that vary over a period of time. The unknowns of differential equations are functions that depend on a variable, usually time. These equations express how those functions change with respect to the given variable. This approach appears to be an effective tool for modeling the evolution of stroke because it allows the introduction of the flow of individuals from one state of health to another.

**General Model**

The main goal in this study is the development of a mathematical model to represent the natural history of stroke in the general population. The method lies in creating a system of differential equations that considers the flow of individuals from one state of health to another. The difficulty of this task is the correct representation of the natural history of the disease. Figure 1 shows the corresponding flow diagram and extends the graph associated to a Markov chain that models the evolution of stroke in the general population [22]. Each of the squares in figure 1 represents one of the possible health states that a person who has had a stroke can subsequently experience. The arrows leading from each of the states indicate the possible states a person can reach in a 1-year
period after having left the previous state. The rate at which the transition events occur is indicated above each arrow. The rates from ‘General population’ to ‘Death’ ($\mu_e$) and from General population to ‘Stroke’ ($\alpha_e$) change within each age group ($e$).

With the onset of a stroke, the patient enters the Stroke state and can subsequently reach the following states: ‘Autonomous’, ‘Dependent’, or Death (as a result of the stroke). The time cycles correspond to 1-year periods, so, after 1 year, a patient reaches one of those three states. This means that the Stroke state is transitory because no patient remains in it after the 1-year cycle is complete.

The patients are classified in either the Autonomous or Dependent states, according to the Barthel index: those patients with an index of 95 points or more are considered Autonomous; if the index is less than 95 points, they are classified as Dependent [23]. A patient in either the Autonomous or Dependent state can have a second stroke; this is classified as the ‘Recurrent stroke’ state, which is transitory. After the 1-year cycle, the patient with Recurrent stroke reaches the Autonomous, Dependent, or Death state.

In the model, the deaths from any cause other than stroke are also considered. These deaths are modeled by the transition to Death from any of the states labeled ‘Healthy general population’, ‘Dependent’, or Autonomous. We point out that both the Dependent and the Autonomous patients have a greater probability of dying from causes other than stroke than people from the Healthy general population; the highest death probability is expressed as $1 - e^{-rate}$ [24] makes possible the use of a probability or a rate, depending on the measure that best fits the dimension adjustment of the equations.

From figure 1, the construction of the system of differential equations that reflects the flow of persons entering and leaving each of the states is easy to implement. For each of the five states, one equation has been constructed, as shown in equations 1–5:

$$dN(t)/dt = 1/1,000 \times [\lambda - (\alpha + \mu_e)] \times N(t)$$  \hspace{1cm} (1)

$$dI(t)/dt = \alpha/1,000 \times N(t) - (\beta + \gamma + \delta)/1,000 \times I(t)$$  \hspace{1cm} (2)

$$dDep(t)/dt = \beta/1,000 \times I(t) + \eta/1,000 \times Ir(t) - (\theta + \xi)/1,000 \times Dep(t)$$  \hspace{1cm} (3)

$$dAut(t)/dt = \gamma/1,000 \times I(t) + r/1,000 \times Ir(t) - (\theta + \xi)/1,000 \times Aut(t)$$  \hspace{1cm} (4)

$$dIr(t)/dt = \xi/1,000 \times Dep(t) + \zeta/1,000 \times Aut(t) - (\eta + \nu + \rho)/1,000 \times Ir(t)$$  \hspace{1cm} (5)

Under the initial condition $[N(t), I(t), Dep(t), Aut(t), Ir(t)] = (1,000, 0, 0, 0, 0)$.

The variables and parameters in the equations 1–5 are: $N(t) =$ number of healthy people at time $t$; $Dep(t) =$ number of dependent people at time $t$; $Aut(t) =$ number of autonomous people at time $t$; $I(t) =$ number of people who have a stroke at time $t$; $Ir(t) =$ number of people who have a recurrent stroke at time $t$; $\lambda =$ birth rate; $\alpha =$ rate of stroke; $\beta =$ rate of dependent people; $\gamma =$ rate of autonomous people; $\delta =$ rate of death for healthy people; $\mu_e =$ rate of death for general population; $\alpha_e =$ rate of death for general population; $\beta_e =$ rate of death for general population; $\gamma_e =$ rate of death for general population; $\delta_e =$ rate of death for general population; $\theta =$ rate of death for dependent people; $\xi =$ rate of death for autonomous people; $\rho =$ rate of death for all causes.

Information Sources
The data used in the study have been obtained from the Guipuzcoa Study [21] in the Basque Country, a report by Mar et al. [22] and from the EUSTAT (Basque Institute of Statistics). The data correspond to the male population of the Basque Country on the 1st of January in the year 2000.

Construction of the System of Differential Equations
For the construction of the model, the following assumptions apply regarding the flow rates among states:

Calculation of Disability through Differential Equations
rate; \( \alpha \) = incidence rate of stroke; \( \mu \) = rate of death for any cause other than stroke; \( \beta \) = transition rate from state Stroke to state Dependent; \( \gamma \) = transition rate from state Stroke to state Autonomous; \( \theta \) = transition rate from state Dependent to state Death = transition rate from state Autonomous to state Death; \( \eta \) = transition rate from state Recurrent stroke to state Dependent; \( \rho \) = transition rate from state Recurrent stroke to state Autonomous; \( \xi \) = transition rate from state Dependent to state Recurrent stroke; \( \nu \) = transition rate from state Recurrent stroke to state Death.

Note: The parameter \( \lambda \) is 0 in case of a closed cohort, that is, a situation in which new individuals are not allowed to enter the system; otherwise, an open cohort is represented, and \( \lambda \) is strictly positive (birth rate).

The solution to the system of equations 1–5 is obtained analytically, but we have found that numerical methods provide accurate results.

**Analytical Solution**

It is easy to realize that equations 1 and 2 form an independent system of equations because they can be solved without taking into account the rest of the system – that is, equations 3–5. In particular, the first equation can be solved by itself. The solution reveals that \( N(t) \) varies with time \( t \).

The states Stroke and Recurrent stroke are transitory, meaning that the following relationships hold:

\[
I(t) = a \times N(t) \quad (6)
\]
\[
R(t) = b \times \text{Dep}(t) + b \times \text{Aut}(t) \quad (7)
\]

where \( a = 1 - \exp(-\alpha/1000) \) and \( b = 1 - \exp(-\gamma/1000) \).

The two equations above are interpreted in terms of expectancies.

Equation 6 reflects the number of people in the state of Stroke in time \( t \), \( I(t) \) and can be calculated as the probability of a person from the General population having a stroke. This probability has been calculated from the incidence by the relation \( p = 1 - e^{-\text{rate} \times t} \).

\[
y(0) = (1000, 0, 0)
\]

The solution then relies on solving a linear system: \( y' = \mathcal{A} \times y \), where \( \mathcal{A} \) is the coefficient matrix, \( y = [N(t), \text{Dep}(t), \text{Aut}(t)] \) and \( y(0) = (1000, 0, 0) \).

Mathematica 5.0® and Matlab® software can be used to arrive at the analytical solution to the system of differential equations 8, 9, and 10. The solution is:

\[
N_e = N_{e-1} \times \exp(-\alpha_e + \mu_e)/1000
\]
\[
\text{Dep}_e = \left[ \frac{d}{dt}\left[ \text{Dep}_{e-1} + \text{Aut}_{e-1} \right] \times g_e \times \pi_e - \text{Aut}_{e-1} \times f_e \times \pi_e + N_{e-1} \times \left( \alpha_e \times \beta_e + \alpha_e^2 \times \sigma_e \times \eta_e + h_e \times \sigma_e \times \mu_e \times \phi_e \right) + f_e \times \mu_e \times \phi_e \right]
\]
\[
\text{Aut}_e = \left[ \frac{d}{dt}\left[ \text{Dep}_{e-1} + \text{Aut}_{e-1} \right] \times g_e \times \omega_e - \text{Dep}_{e-1} \times f_e \times \omega_e + N_{e-1} \times \left( \alpha_e \times \nu_e + \alpha_e^2 \times q_e + h_e \times \alpha_e \times \rho_e \right) + f_e \times \mu_e \times \phi_e \right]
\]

where subscript \( e \) makes reference to the value of the variables for the age \( e \), subscript \( e - 1 \) makes reference to the previous age, and

\[
g_e = \exp(-\alpha_e + \mu_e)/1000
\]
\[
f_e = \exp(-\alpha_e + \mu_e)/1000 + (371 + 55 \times \mu_e)/25000
\]
\[
h_e = \exp(255 + 11 \times \mu_e)/50000 + (371 + 55 \times \mu_e)/25000
\]
\[
p_e = \exp(-\alpha_e + \mu_e)/1000 - (255 + 11 \times \mu_e)/50000 - (371 + 55 \times \mu_e)/25000
\]
\[
\pi_e = D - E \times \alpha_e + F \times \alpha_e^2 + P \times \mu_e - Q \times \alpha_e \times \mu_e + V \times \mu_e^2
\]
\[
\beta_e = G \times f_e + B \times g_e - O \times g_e \times \mu_e
\]
\[
\sigma_e = L \times g_e + H \times f_e
\]
\[
\psi_e = N_e \times R \times \alpha_e + \text{Dep}_e \times (S - T \times \alpha_e + W \times \mu_e)
\]
\[
\phi_e = I - J \times \alpha_e + K \times \alpha_e^2
\]
\[
\Omega_e = 3,616,000 \times (94,605 - 8,230 \times \alpha_e + 125 \times \alpha_e^2 + 9,876 \times \mu_e - 300 \times \mu_e \times \alpha_e + 180 \times \mu_e^2)
\]

The given values of the rates, the solution was obtained with the corresponding solver. Because a compact expression of the solution can be quite convoluted, we show the constant values provided by the differential equations solver of the mathematica package (Mathematica 5.0) to help in the understanding of its structure.

Constant values: \( B = 8,643,575,625; C = 169,481,875; D = 17,643,832,500; E = 15,348,950,000; F = 233,125,000; G = 701,354,353; H = 47,261,075; I = 16,565,335,500; J = 14,410,730,000; K = 218,875,000; L = 794,222,127; N = 122,220,800; O = 323,378,250; P = 18,418,740,000; Q = 559,500,000; R = 56,713,290; S = 17,292,876,000; T = 525,300,000; U = 146,664,960; V = 335,700,000; W = 315,180,000.

**Numerical Solution**

Numerical methods for solving a system of differential equations are helpful when an analytical solution is not available or when only a discrete solution is of interest. To select a method, the precision required in the solutions and the complexity of the relations among the variables must be taken into account. For instance, the system including equations 8–10 can be written as \( \dot{y} = \Phi_e(t, Y) \), where \( \Phi_e = A_e y + A_e \) is the matrix of coefficients for the age \( e \). In this study, the Euler method [25] provides accurate approximations to the solution with only 10 iterations. But, when \( \Phi_e \) is more complex, more iterations are needed, or even more accurate methods, such as those of Runge-Kutta, should be applied [25].

In this study, the Euler method has been implemented:

\[
y_{k+1}(e) = y_k(e) + k/r \times A_e y_k(e) \times y_k(e)
\]

where \( e \) represents the age group and \( r \) indicates the number of iterations.
Calculation of Prevalence

Prevalence may be defined as the proportion of a population that has disease at a specific point in time [2, 26], and it relies on all the birth cohorts with living subjects that could include those with the specified attribute at the moment of the calculation [27]. A birth cohort is a group of people born in the same year. In our program, strokes occurring before patients have reached the age of 15 years are negligible. The aim of this study is to calculate the prevalence of patients between the ages of 15 and 100 years in the population of the Basque Country who are Dependent after having had a stroke. To determine the prevalence, we calculate the probability of patients being in each of the states after a given number of years by finding the solution to the system. We assume that initially only one person is in the General population cohort and that zero are in the other two states, therefore, solving the system with the initial condition (1, 0, 0). These probabilities allow us to calculate the evolution of each birth cohort until year 2000. For example, if we consider the number of people born in 1939 who were 60 years old at the beginning of 2000, this figure represents the probability of persons being in that particular state, given that they were alive at the age of 15 in 1954. Specifically, the number of prevalent cases in each of the poststroke states for each age is obtained by multiplying the probability by each cohort size.

The second step is to ascertain the number of individuals in the origin of each annual cohort [27]. With the preceding example, we need to know the number of men and women born in 1939 who were alive in 1954. However, we not only have to take into account those who were born in the Basque Country in 1939 and were still alive in 1954 but also all the individuals who constitute the current population. To estimate that initial cohort with different dates of origin, we begin the calculation with the population who were born in 1939 and were still alive in 2000 (8,708 men and 9,141 women). We proceed retrospectively by dividing that figure by the survival probability, i.e. $P(S)_{2000}$, of the General population at that age. We calculate that number from all-cause mortality rates by sex and year ($r$).

From the total number of men in the cohort born in 1939 (1939 birth cohort) who remained alive in 2000 (8,708), we can calculate that 8,825 [8,708/$P(S)_{2000}$] were alive in 1999, and, finally, an estimated 10,321 men were alive at the age of 15 in 1954. The same process is repeated with all of the 85 birth cohorts between the ages of 15 and 100 years to obtain the number of men for each cohort who were alive at the age of 15, and then we can obtain the number of prevalent cases for each cohort in 2000, as explained before.

Model Validation and Calibration

We carried out an external validation of the model by comparison with the results of a previous study based on Markov models [27] and with the figures described in the literature [28]. Bonita et al. [28] estimated stroke prevalence rates by sex and age group from a population register in Auckland. We applied those figures to the Basque population to obtain the prevalence pool and used that as a second comparator. Calibration refers to the process of determining the relation between the output of the model and the value of the input quantity used for relative risk. The model was calibrated by changing the relative risk of death by age group. The process consisted of repeatedly running the equations system and validating different figures for different ages by comparison with the external results. The final stage was the selection of the relative risk for each age that produced the results of the comparator. To obtain those estimates, we eliminated assumption A and used $\theta_c = r_e \times \mu_e$, where $r_e$ represents the relative risk of death for each age group. We introduced this relation in equations 9 and 10, and then we equated the solution, depending on $r_e$, to the validated values [27], from which an estimate of $r_e$ was obtained. Finally, we solved equations 8, 9, and 10, thereby introducing the new estimates of $\theta_c$. The new model is called a calibrated differential equation model.

Results

We have distinguished between intermediate and final results. The first or intermediate outcomes are the solutions of the equations over a period of 100 years. The evolution of the General population and that of each of the two possible states for an initial open cohort of 1,000 persons are presented in figures 2–4. The evolution of the Dependent and Autonomous states in the case of an open cohort is the same as that shown in figures 3 and 4 because the equations that refer to these states do not change when newborns are included. However, the first equation relates to the changes in the General population, as does the evolution of the General population. More precisely, the graph for the General population in this case shows a straight line because the number of
deaths is compensated by the number of newborns, owing to the fact that the newborn rate is equal to the death rate.

The final results shown in Table 1 are the prevalence estimates by age group and state of disability. This table includes both numbers and the transformation to rates with use of an adjusted and a nonadjusted model. The prevalence figures that were obtained have been compared with those that were estimated where the Markov model was applied. The graph with both curves is shown in Figure 5; it can be observed that the results obtained by the differential equation model overestimate the validated values given previously [27]. This observation is a consequence of assumption A, where the same relative risk,
calculation of disability through differential equations

2.2, for each age group is assumed. This relative risk value is rather small for the age groups less than 70 years old, so more people than would be expected according to the model remained alive after having had a stroke at an early age. Nevertheless, this drift from the validated values allows us to obtain an estimate (table 2) of the relative risk of death for each age group. In figure 5, the results of the Markov model, Auckland population, and those of the calibrated and noncalibrated differential equation model are presented.

Discussion

This paper presents a method for estimating the prevalence of different states in the natural history of stroke according to the level of patient disability. Mathematical models are of interest in epidemiology because they permit integration of facts and values, define the major relations among the components, and link these data to outcomes of interest. They have been widely used for the representation of the natural history of different diseases. The two roles identified for this model are the extrapolation of outcomes in the long term and the translation from intermediate to final outcome. Prevalence estimates of neurological diseases are commonly obtained from population-based epidemiological surveys. However, there are some concerns about those studies [29]. First of all, the figures estimated in those studies are specific for a country or region and are difficult to generalize. Second, surveys are lacking in many countries. Third, demographic structures are changing so rapidly that previous estimates have become useless in many cases. Therefore, models are a technique that can fill the gap that surveys leave. The available parameters, which include incidence, mortality, and evolution of stroke, allow the estimation of another measure of disease, prevalence.

A primary concern is related to the validity of the structure of the model. We have assumed that stroke has two possible disease states, Autonomous and Dependent, in order to calculate the prevalence for both. The assumption of only two permanent states of disability is based on the approach applied in well-known clinical trials assessing treatment for stroke [30].

This approach can be implemented to calculate the prevalence of other chronic diseases, such as diabetes [11] or breast cancer [31] that can result in the development of special patient needs in the natural history of the disease. In our experience, when compared with the results obtained by implementing Markov models, the benefit of the continuous use of time outweighs the mathematical requirements of our model. Nevertheless, when a nonlinear system of differential equations is obtained, an analytical solution may not be available, and only numerical methods can be used to solve the system. In such a case, experts in the use of numerical methods to solve differential equations should be consulted to avoid a misuse

Table 1. Number and rate of stroke prevalence by age and state of disability in men calculated with the differential equation model for the year 2000

<table>
<thead>
<tr>
<th>Age group</th>
<th>Nonadjusted model results</th>
<th>Adjusted model results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Autonomous</td>
<td>Dependent</td>
</tr>
<tr>
<td>0–14</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15–44</td>
<td>254</td>
<td>169</td>
</tr>
<tr>
<td>45–54</td>
<td>562</td>
<td>376</td>
</tr>
<tr>
<td>55–64</td>
<td>1,408</td>
<td>944</td>
</tr>
<tr>
<td>65–74</td>
<td>2,955</td>
<td>2,002</td>
</tr>
<tr>
<td>75–84</td>
<td>2,087</td>
<td>1,420</td>
</tr>
<tr>
<td>85+</td>
<td>454</td>
<td>304</td>
</tr>
<tr>
<td>Total</td>
<td>7,720</td>
<td>5,215</td>
</tr>
</tbody>
</table>

Rates

<table>
<thead>
<tr>
<th>Age group</th>
<th>Number</th>
<th>Rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–14</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15–44</td>
<td>51</td>
<td>51</td>
</tr>
<tr>
<td>45–54</td>
<td>391</td>
<td>391</td>
</tr>
<tr>
<td>55–64</td>
<td>620</td>
<td>620</td>
</tr>
<tr>
<td>65–74</td>
<td>3,008</td>
<td>3,008</td>
</tr>
<tr>
<td>75–84</td>
<td>5,001</td>
<td>5,001</td>
</tr>
<tr>
<td>85+</td>
<td>4,859</td>
<td>4,859</td>
</tr>
<tr>
<td>Total</td>
<td>675</td>
<td>675</td>
</tr>
</tbody>
</table>

Rate = Prevalence rates per 100,000 population; Number = number of men with stroke in the Basque population.

Table 2. Relative risk of death by age obtained by model calibration

<table>
<thead>
<tr>
<th>Age group</th>
<th>Relative risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–14</td>
<td>not applied</td>
</tr>
<tr>
<td>15–44</td>
<td>50</td>
</tr>
<tr>
<td>45–54</td>
<td>18</td>
</tr>
<tr>
<td>55–64</td>
<td>8</td>
</tr>
<tr>
<td>65–74</td>
<td>4</td>
</tr>
<tr>
<td>75–84</td>
<td>2.4</td>
</tr>
<tr>
<td>85+</td>
<td>1.7</td>
</tr>
</tbody>
</table>

Calculation of Disability through Differential Equations

Neuroepidemiology 2008;31:57–66
of the existing solvers or the eventual programming of a specific code. The advantage of ordinary differential equations is that they allow the use of open models. This system’s important consequence is a compact analysis of the total evolution of stroke in a general population. Markov models permit modeling of the evolution in only a homogeneous cohort of patients.

Prevalence has become a major issue in neuroepidemiology as a consequence of the increasing weight of disability-generating chronic conditions in the magnitude of the burden of disease [1]. Developed societies manage the consequences of disability by planning health and social services. In this process, the assessment of needs is the first step, and the prevalence is a key parameter for carrying out planning [32]. Data on the epidemiology of disability as a nonfatal health outcome are scarce in many countries [1–4]. The implementation of surveys to fill the gap is limited by the fact that disability is infrequent, and, therefore, the sample size must be huge to generate reliable results. Especially in developing countries, the difficulties in carrying out population surveys make models the only practical approach to obtaining estimates of outcomes that can be used in the decision-making process for health management. However, to construct a mathematical model for the natural history of any disease is epidemiologically challenging. To achieve this goal, it becomes essential to pay special attention to three criteria. The first is that all the key events must be included; the second is that the flow of individuals between states must be accounted for, and the third is that the dimensions of both sides of the equations must fit at every time. Differential equations are the natural mathematical representations of the evolution of dynamic systems with unknowns that evolve with respect to time. Moreover, systems of differential equations are a popular mathematical tool, and so, most mathematical packages provide effective approaches to solutions. This method of differential equations has been described thoroughly enough to allow its application to other diseases. However, the estimation of the parameters represents an important difficulty and depends on the available figures provided by statistical institutions and the literature. In fact, hypotheses A–C for the case of stroke try to simplify the complicated relations among the flow rates. When this procedure is extended to other diseases, adapting assumptions A–C to the particular disease is advisable. These models are based on incidence, recovery, and mortality. Therefore, this approach can theoretically be applied to any disease whose natural history is known, because it does not restrict the number of states. However, the system expands and becomes more difficult to solve as the number of states increases. It is advisable to establish a balance between the number of states and the number of available parameters.

Differential equations are deterministic methods, so the solution expresses a fixed evolution of the number of individuals in each state. As always in deterministic modeling, a sensitivity analysis can be used to assess the impact of changing the value of some key parameters in the final outcome. Theoretically, a probabilistic approach based on Monte Carlo simulations could be carried out to obtain uncertainty intervals because some parameters can be considered as random distributions. This task involves a major effort in programming calculations because it entails solving the system of differential equations hundreds of times, each time with a different set of values. For the time being, sensitivity analysis seems to be the reasonable approach to managing uncertainty in complex differential equation models.

The prevalence of the various states for stroke was first estimated by means of Markov models [27]. In figure 5, the curve calculated with Markov models and that obtained in the present study have the same pattern. Nevertheless, a higher prevalence was obtained with differential equations, as a consequence of assumption A. The calibration of the model by adjusting the relative risk of death allowed a good fit of the results. When the Auckland prevalence rates were applied to the Basque population, the results showed figures for the Basque Country a third higher than in our study. However, those results fit well when age-adjusted incidence in both populations is taken into account, because the incidence is also one third higher in Auckland. The natural history of stroke has two components that are associated with an excess mortality rate in comparison with that in the general population. The first one relies on the probability of a fatal outcome of the acute event of the first-ever and recurrent strokes. The second is the mortality excess related to other causes. Markov models in the literature use a constant relative risk to represent that excess, which is multiplied by all-cause death rates to obtain the death probabilities by etiology other than stroke. We have not found primary data that allow us to support empirically the conclusion that those relative risks differ by age, but we have validated it with secondary data based on prevalence calculation. If we assess the years of life lost in young or mean-age patients with stroke compared with those lost in young or mean-age persons in the general population who die of causes other than stroke, we think it is a natural conclusion. Stroke is not a destiny consequence,
but rather the result of a combination of pre-existent risk factors that also increase the incidence of other cardiovascular diseases. Given that those rates are very low in young individuals, low in mean-age persons, and high in the elderly, a relative risk that decreases with age seems plausible.

Stroke is usually seen as an acute event that is measured in terms of incidence. However, an accurate approach should also take into account the consequences of the disease that remain over time [33]. Stroke-related disability can be assessed adequately only when a prevalence-based approach is applied. A consequence of our work has been to underscore the growing epidemiologic role of methods to calculate the impact of chronic conditions. Moreover, our society’s ageing population ensures a key role for health care that addresses the special-needs states caused by disability. In this task of enhancing understanding of the social problem of disability-generating neurological diseases, epidemiologic methods play a key role because they entail a change in the relationship between incidence and prevalence. The epidemiologic approach applied to assess the burden of mortality-generating diseases (for instance, cancer or coronary heart disease) is based on the measurement of incidence. This figure is determined by the prevalence of risk factors such as smoking or hypertension and helps to orient public health programs that address primary prevention. Survival methods have been used to deal with the issues raised by incidence, but in neurological diseases, a critical component for the development of policy is the health states of individuals and the health status of populations. The measure of interest then becomes the prevalence of dependent patients by levels of disability in relation to the incidence of disabling diseases, such as stroke, Alzheimer’s, and multiple sclerosis. This approach is clear in patients with stroke because an acute event occurs at the onset of the process. But, it is also valid in instances where incidence is difficult to observe, as in Alzheimer’s disease, because the etiologic process is the same. Notwithstanding the subtle and progressive appearance of symptoms, there is a sequence of pathophysiological changes that leads to the appearance of impaired functional states. When empirical information is scarce, expert opinion can be a source for estimates [29]. The development of methods to assess states of health on the basis of nonfatal outcomes, such as disability, is of high priority, given the predicted increased burden of chronic diseases in the near future [1, 3]. More research is needed to make models based on differential equations a standard technique in the measurement of the consequences of disease.

Acknowledgments

This study was supported by grant 200411010 from the Department of Health, Basque Government and grant PI05/1703 from the ‘Carlos III’ Health Institute, Madrid.

The authors would like to thank Sally Ebeling for editorial assistance in the preparation of this article and Belen Elizalde for the support during the process.


