

## How to obtain long term projections for smoking behaviour: A case study in the Dutch population

Maurizio Capannesi<sup>ab\*</sup>, Hendriek C. Boshuizen<sup>b</sup>, Marc C. Willemsen<sup>c</sup> and  
Hans C. van Houwelingen<sup>ab</sup>

<sup>a</sup>Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, University of Leiden, Leiden, The Netherlands; <sup>b</sup>Expertise Centre for Methodology and Information Services, National Institute for Public Health and the Environment, Bilthoven, The Netherlands; <sup>c</sup>STIVORO for a Smoke Free Future, Expertise Centre on Tobacco Control, Den Haag, The Netherlands

A state-transition model is presented to project the prevalence rates of never, current and former smokers within the Dutch population. Changes are determined by the transition rates among these three classes: start rates, quit rates and relapse rates. Both the initial prevalence rates and transition rates are calculated from cross-sectional data using a restricted quadratic multinomial regression spline and a restricted quadratic logistic regression spline, respectively. Through a Monte Carlo experiment an uncertainty analysis was performed to assess the level of reliability of the results, while a sensitivity analysis was carried out in order to detect the input variables that mostly contribute to the output variability. The results obtained through this model show a considerable but rather slow decrease of smokers up to year 2050.

**Keywords:** state transition models; restricted quadratic regression spline model; smoking prevalence rates; transition rates; uncertainty analysis; sensitivity analysis

**AMS Subject Classification:** 49Q12; 92D30; 65D10

### 1. Introduction

Tobacco is the second major cause of death in the world and it is currently responsible for the death of one in ten adults worldwide (about five million deaths each year) [1]. The burden of tobacco addiction on health in high income countries has been estimated as 6–15% of the total health care costs [2]. Due to the health consequences and the large amount of costs involved, public health policy makers are interested in tools to describe future trends of smoking prevalence in the population and of the related diseases [3–5]. Several instruments to support decision makers in planning interventions, in terms of costs and human resources, have been developed [3,4,6,7]. All these tools are based on projections of smoking prevalence rates.

Our main purpose is to obtain projections of the prevalence rates for never smokers/smokers/former smokers within the Dutch population using a state transition model that describes the population in terms of smoking behaviour. The projections shown in this research are coherent with the well-known model of the cigarette epidemic thought in the 1990s [8], in which The Netherlands is classified as a country in the last stage of its smoking epidemic. But the smoking behaviour in this country seems to move more slowly towards lower values than is required to reach the national targets as set by the Dutch government, that is 20% of smokers in 2010 [9].

---

\*Corresponding author. Email: Maurizio.Capannesi@rivm.nl

To get important information about the robustness of our projections we performed an uncertainty and sensitivity analysis of the model output. The uncertainty analysis gives the uncertainty of the model output based on the uncertainty of the model input. The aim of a sensitivity analysis is to determine the input variables that mostly contribute to the output variability and pinpoint those variables that require further study if we would want to make better projections. Both kinds of analyses are strongly related. For both of them a Monte Carlo analysis was carried out, based on performing multiple evaluations with randomly selected model input, and then using the results of these evaluations to determine both uncertainty in model projections, and the contribution of the input variables to this uncertainty [10,11].

## 2. Methods

### 2.1 Data sources and modelling input rates

Three different data sources have been employed. The first is the mortality probabilities. They were obtained from Statistics Netherlands (Vorborg, The Netherlands) specified by gender and within gender by one-year age class (15–85) for 2003 and the Statistics Netherlands demographic projections specified by gender and within gender by one-year age class (15–85) from 2006 up to 2050 [12]. We decided to use 2003 mortality probabilities instead of those from 2004 because the data from which we calculated transition rates and prevalence rates refer to 2003; however, we re-evaluated the model with the 2004 mortality probabilities, and this did not produce any relevant difference. We assume that both the mortality rates and the demographic projections are error free.

The second data source is the relative risks with respect to never smokers for all cause mortality specified by gender and within gender by smoking class (smokers and former smokers) and within smoking class by 5-years age class. These relative risks were obtained from the Dutch National Institute for Public Health and the Environment, Bilthoven, The Netherlands [13].

The third data source used to estimate the input prevalence rates and the transition rates for smoking behaviour are data from research commissioned by STIVORO, Den Haag, The Netherlands, the Dutch expert centre on tobacco control and carried out by TNS NIPO (The Dutch Institute for Public Opinion and Market Research, Amsterdam, The Netherlands) in 2003 in 19,029 persons. TNS NIPO uses a database with persons who have agreed to participate in research. From this a representative sample of the Dutch population aged 15 and older was selected and asked to fill in a questionnaire on their own PC, through the internet or by sending the program on diskette (at 2003 the 80% of households in The Netherlands had a PC [14]). Every week, 200 households were selected randomly, and within these households two persons older than 15 were appointed to answer the questionnaire. We decided to use the 2003 database because this was a year in which no special tobacco control interventions have been introduced by the Dutch Government.

Subjects were required to answer the following questions: ‘Do you smoke now (yes/no)?’. If yes: ‘Did you smoke exactly one year ago (yes/no)?’. If the question was answered affirmatively respondents were classified as *continuous current smokers*. If answered with ‘no’ they were asked: ‘Did you start smoking for the first time (yes/no)?’. If answered affirmatively, the respondents were classified as *starters*, if answered with ‘no’ they were classified as *relapsers*. If respondents answered ‘no’ to the first question ‘Do you smoke now?’ they were asked: ‘Did you ever smoke (yes/no)?’. If answered with ‘no’ they were classified as *continuous never smokers*. If answered with ‘yes’ they were asked when they had quit smoking. All respondents who quit smoking sooner than one year ago were

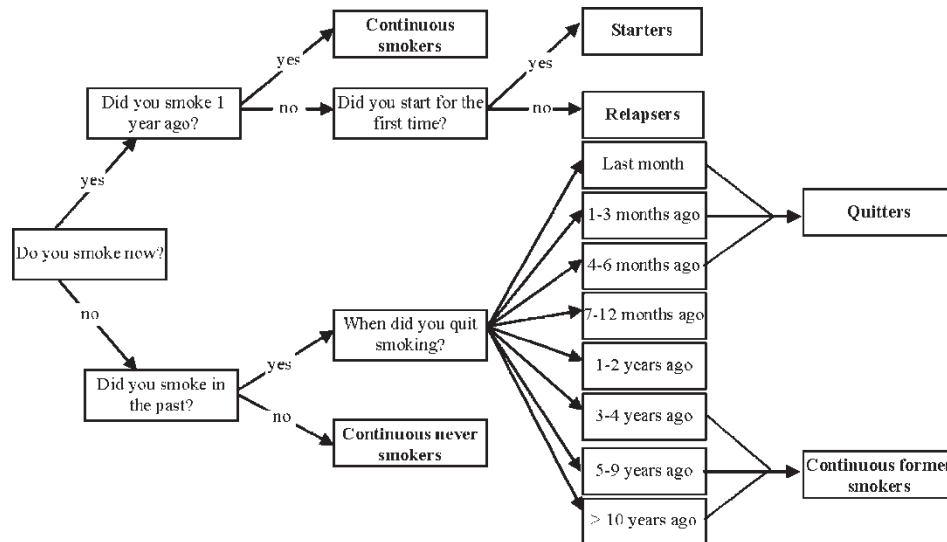


Figure 1. STIVORO algorithm to classify the respondents.

classified as *quitters*. All respondents who quit smoking longer than one year ago were classified as *continuous former smokers*. Figure 1 shows schematically this classification algorithm. Moreover gender and age at the moment of the interview were recorded.

The data we are dealing with are dichotomous concerning the transition rates, while the prevalence rates are a multinomial variable (never smokers, smokers and former smokers). The prevalence rates specified by gender and within gender by age were, therefore, calculated by fitting a restricted quadratic multinomial logistic regression spline while a restricted quadratic logistic regression spline has been chosen to model the transition rates [15]. The two knots yielding the best Akaike Information Criterion were selected in the whole age range of 14–86. The values chosen for the knots are 17 and 33, 25 and 35, 15 and 19 for start rates, quit rates and relapse rates in males, respectively. For females the knots were 23 and 25, 59 and 63 for start and quit rates, respectively. For the relapse rates a logistic regression model without knots was chosen as the best model. For the prevalence rates the knots were 43 and 75, 45 and 75 for males and females, respectively. The choice of the restricted quadratic logistic regression spline model to generate the age specific transition rates and prevalence rates was suggested by the necessity to find a parametric model that fits the data as well as a non parametric one. The method we adopted can be seen as a middle way between a logistic (or multinomial) regression and a smoother. With this model we still have a parametrically formulated model, but can fit the data much better. As we can assume that the model parameters are multinormally distributed, in this way we were able to generate samples of input variables as needed to perform the uncertainty and sensitivity analysis by repeatedly drawing a sample of the model parameters.

Further details and a mathematical description of the restricted quadratic logistic regression spline model are given in Section 2.2.

The start rates at younger ages derived by fitting a restricted quadratic logistic regression spline show a high instability due to the small numbers. Also the fitted rates are strongly inconsistent with the difference of the prevalence rates between neighbouring ages. Such an inconsistency is unrealistic, as available data show that the prevalence of smoking in

teenagers has been rather stable in recent years. Therefore, we decided to calculate the start rates in the age range 15–19 from the difference in prevalence rates of never smokers between neighbouring ages.

## 2.2 The restricted quadratic logistic regression spline model

The initial prevalence rates and the transition rates specified by gender and within gender by age were calculated by fitting a restricted quadratic logistic regression spline model.

As an example, we describe the case of starting rates. Let  $Y$  be the binary outcome variable (1 = starters and 0 = continuous never smokers) and the  $x$  represents the age measured in years. The restricted quadratic regression spline model is:

$$\log[\pi(x)/(1 - \pi(x))] = \beta_0 + \beta_1 \ln(x) + \beta_2(s_2^2 - s_3^2),$$

where

$$s_2 = \begin{cases} \ln(x) - \ln(x_1) & x > x_1 \\ 0 & \text{otherwise} \end{cases}, \quad s_3 = \begin{cases} \ln(x) - \ln(x_2) & x > x_2 \\ 0 & \text{otherwise} \end{cases}$$

and  $x_1$  and  $x_2$  are the knots chosen for the spline model and  $\pi(x)$  is the probability of starting to smoke at age  $x$ . For the quit rate the outcome variable will be  $Y$  (1 = quitters and 0 = continuous current smokers), while for the relapse rate:  $Y$  (1 = relapsers and 0 = continuous former smokers).

Concerning the prevalence rates, the model is a restricted quadratic multinomial regression spline. Let be  $Y$  the multinomial outcome variable (0 = never smokers, 1 = smokers and 2 = former smokers). The model is:

$$\log[\pi_1(x)/\pi_0(x)] = \beta_{10} + \beta_{11} \ln(x) + \beta_{12}(s_2^2 - s_3^2),$$

$$\log[\pi_2(x)/\pi_0(x)] = \beta_{20} + \beta_{21} \ln(x) + \beta_{22}(s_2^2 - s_3^2),$$

where  $s_2$  and  $s_3$  are defined by the previous formulas and  $\pi_0(x)$ ,  $\pi_1(x)$  and  $\pi_2(x)$  are the probabilities at age  $x$  of being a never smoker, a smoker and a former smoker, respectively.

## 2.3 State-transition model

The structure of the state-transition model is defined by the three states: *never smokers*, *smokers* and *former smokers*. The possible state transitions are from never smokers to smokers, from smokers to former smokers and from former smokers to smokers. These changes are governed by transition rates, which we will call *start rate*, *quit rate* and *relapse rate*, respectively.

Starting from the six categories obtained from the STIVORO survey (*continuous never smokers*, *starters*, *continuous current smokers*, *quitters*, *continuous former smokers* and *relapsers*) we can estimate directly the prevalences for the three states at the beginning of the simulation and also the transition rates between them. Figure 2 shows the possible transitions among the states and the relations between the state-transition model and the STIVORO survey categorization.

The state *Never smokers* consists of ‘continuous never smokers’ and the state *Smokers* comprises ‘starters’, ‘relapsers’ and ‘continuous current smokers’. The *Former smokers* state consists of ‘quitters’ and ‘continuous former smokers’.

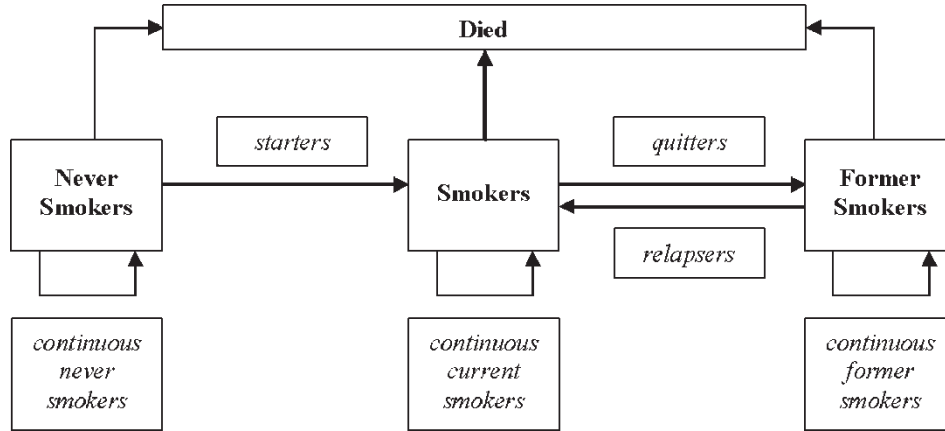


Figure 2. Smoking states and the possible transitions among them.

*Start rates* (specified by sex and age) can be estimated by dividing the number of ‘starters’ by the sum of the number of ‘starters’ and the number of ‘continuous never smokers’, *quit rates* by dividing the number of ‘quitters’ by the sum of the number of ‘quitters’ and the number of ‘continuous current smokers’. Similarly, *relapse rates*, specified by age and sex, can be obtained by dividing the number of ‘relapsers’ by the sum of the number of ‘relapsers’ and the number of ‘continuous former smokers’.

The state-transition model satisfies the Markov property: conditional on the current model state the probability distribution of the model states one time-step ahead is independent on the past model states. In this model, we assume transition rates are constant over time. Due to data limitations, our model simulations start at age 15. This implies that the prevalence rate for people aged 15 is constant over time. The projections of the prevalence rates for the years ahead are obtained through the following equation:

$$p_{a+1,t+1,s}^T = \{1 - m_{a,s}\}^{-1} p_{a,t,s}^T (\mathbf{I} - \mathbf{M}_{a,s}) \mathbf{T}_{a,s},$$

where  $\mathbf{p}_{a,t,s}$  and  $\mathbf{p}_{a+1,t+1,s}$  are the vectors of the prevalence rates for age  $a$  at time  $t$  and for age  $a + 1$  at time  $t + 1$ , respectively, for gender  $s$ .  $\mathbf{T}_{a,s}$  is the transition matrix,  $\mathbf{M}_{a,s}$  is the matrix in which the diagonal elements are the mortality probabilities for each smoking risk factor class, for age  $a$  and gender  $s$  and all the other elements null,  $\mathbf{I}$  is a  $3 \times 3$  identity matrix,  $m_{a,s}$  is the mortality probability for age  $a$  and gender  $s$ . It can be rewritten in the following full form

$$\begin{bmatrix} p_{a+1,t+1,s}^1 \\ p_{a+1,t+1,s}^2 \\ p_{a+1,t+1,s}^3 \end{bmatrix}^T = \frac{1}{1 - m_{a,s}} \begin{bmatrix} p_{a,t,s}^1 \\ p_{a,t,s}^2 \\ p_{a,t,s}^3 \end{bmatrix}^T \begin{bmatrix} 1 - m_{a,s}^1 & 0 & 0 \\ 0 & 1 - m_{a,s}^2 & 0 \\ 0 & 0 & 1 - m_{a,s}^3 \end{bmatrix} \\ \times \begin{bmatrix} t_{a,s}^{1 \rightarrow 1} & t_{a,s}^{1 \rightarrow 2} & 0 \\ 0 & t_{a,s}^{2 \rightarrow 2} & t_{a,s}^{2 \rightarrow 3} \\ 0 & t_{a,s}^{3 \rightarrow 2} & t_{a,s}^{3 \rightarrow 3} \end{bmatrix},$$

where  $t_{a,s}^{ri \rightarrow ri'}$  is the transition rate from state  $ri$  to state  $ri'$ , and  $m_{a,s}^{ri}$  is the mortality rates specified for the risk factor class  $ri$  ( $ri = 1$  Never smokers,  $ri = 2$  Smokers and  $ri = 3$  Former smokers) for age  $a$  and gender  $s$ , calculated through the formula

$$m_{a,s}^{ri} = (RR_{a,s}^{ri} m_{a,s}) / \left( \sum_{ri} RR_{a,s}^{ri} p_{a,s}^{ri} \right),$$

where  $RR_{a,s}^{ri}$  is the relative risk with respect to never smokers for all cause mortality specified by gender and within gender by smoking class  $ri$ .

#### 2.4 Uncertainty and sensitivity analysis

In order to assess the uncertainties of the prevalence projections and to study how the variation of the prevalence rates projections depends on the uncertainties in the input variables (i.e. transition rates and input prevalence rates) a Monte Carlo analysis was performed. We generate a 5000 sample of input variables in the following way: we first generate a sample of 5000 sets of the parameters of the restricted quadratic spline logistic regression model and the restricted quadratic spline multinomial regression model for the transition rates and for the prevalence rates, respectively, by taking random draws from the multinormal distributions of the parameters. From these, we obtained both transition rates and prevalence rates (input), and used these to project the prevalence rates (output).

The main uncertainty and sensitivity analyses consist of the 95% pointwise confidence intervals (95% CI) for the projected prevalence rates in the years from 2006 to 2050 and of the correlation coefficients among the input rates and the output prevalence rates. Also the standardized prevalence projections were calculated, which are the projections obtained by fixing the population age structure to be equal to the age structure of 2006. Both the Monte Carlo experiment and the state-transition model were implemented using S-PLUS<sup>®</sup> [16].

### 3. Results

The results of interest are the prevalence rates of never smokers, smokers and former smokers within the Dutch population aged 15–85. Our projections show a considerable decrease of smokers, albeit slow, up to year 2050. The projected prevalence rate of smokers aged 15–85 decreases from 27.1% (95% CI: 26.3–27.9) in 2006, to 17.8% (95% CI: 16.1–19.6) in 2050, which means more than one million smokers less in four decades, as shown in Figure 3. At the same time we can see an increase in the number of never smokers from 41.8% (95% CI: 41.1–42.4) to 57.6% (95% CI: 54.6–59.7). The projected percentage of former smokers is increasing from 31.1% (95% CI: 30.4–31.9) to 32.4% (95% CI: 31.1–33.9) in 2015, then decreasing to 24.6% (95% CI: 23.1–26.9) in 2050 (Figure 3). In detail, the percentage of smoking males decreases from 29.6% (95% CI: 28.4–30.8) in 2006, to 19.9% (95% CI: 17.2–22.6) in 2050, while for females it decreases from 24.7% (95% CI: 23.6–25.8) to 15.8% (95% CI: 13.7–18.2). We also see an increase in the percentage of never smokers: 36.5% (95% CI: 35.5–37.4) to 54.3% (95% CI: 49.7–57.9) for males and 47% (95% CI: 45.9–47.9) to 60.7% (95% CI: 57.3–63.4) for females. The trends in the percentages of former smokers differ by sex: essentially decreasing for males, from 33.9% (95% CI: 32.8–35) to 25.8% (95% CI: 23.4–29.2), while for females they increase from 28.4% (95% CI: 27.3–29.5) to 30.8% (95% CI: 28.7–32.9) in 2018, and then decrease to 23.5% (95% CI: 21.2–26.2) in 2050 (Figure 4).



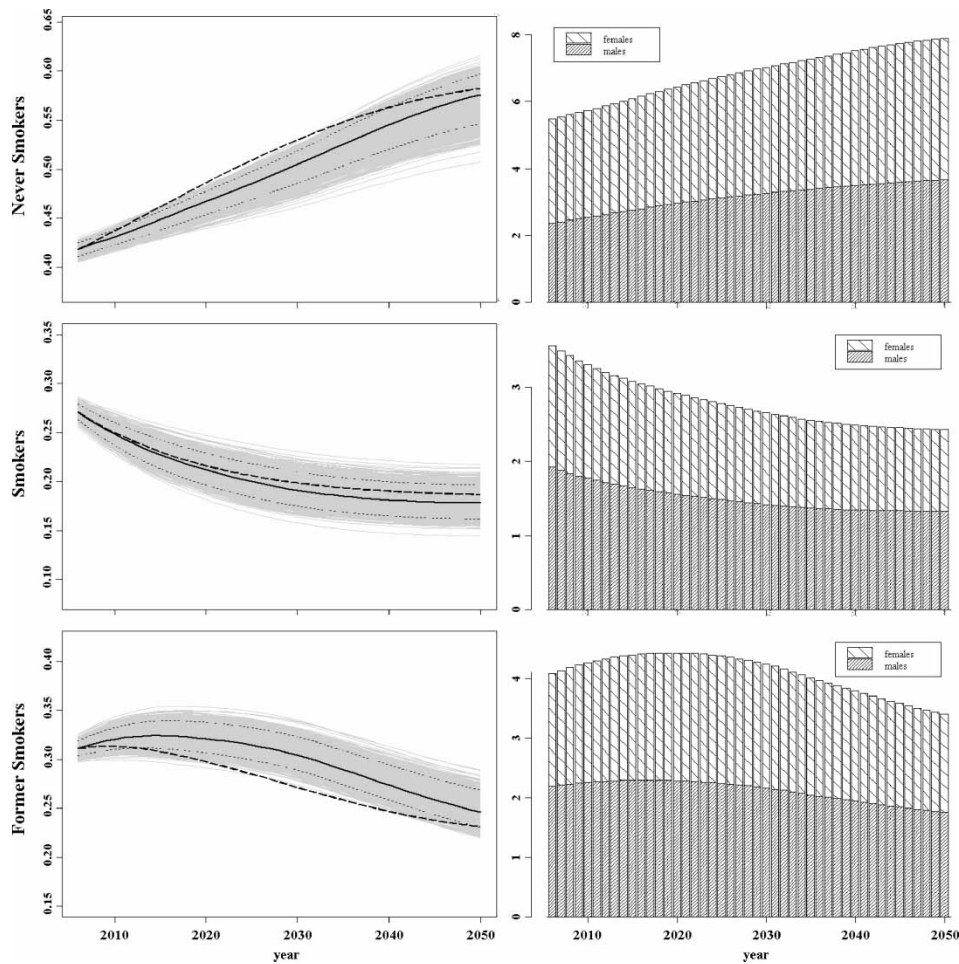


Figure 3. On the left, prevalence rates projections in the Dutch population aged 15–85. Black solid curve, prevalence projections; short dashes, standardized prevalence projections; black dotted lines, 95% pointwise confidence interval; and grey solid curves simulated prevalence projections. On the right, prevalence projections in millions.

For smokers, the projected standardized prevalence rates do not differ greatly from the unstandardized prevalence rates; which means that if the demographic structure of the Dutch population would remain the same, essentially the same results would be obtained for the prevalence rates of smokers. This is due to the fact that the demographic structure of the part of the population that mostly contributes to the proportion of smokers in the population remains the same in the next decades. In contrast, the projected standardized prevalence rates for never smokers are higher than the unstandardized rates and the projected standardized prevalence rates for former smokers are lower than the unstandardized rates.

Results of the sensitivity analysis obtained for males and females are similar. If a model input has a large effect on a particular model output in males, in general the same input–output relationship will be found in females, although with different magnitude.

In general, the input prevalence rates mostly influence the projected prevalence rates during the early years, but their influence decreases in importance with ongoing time. Only

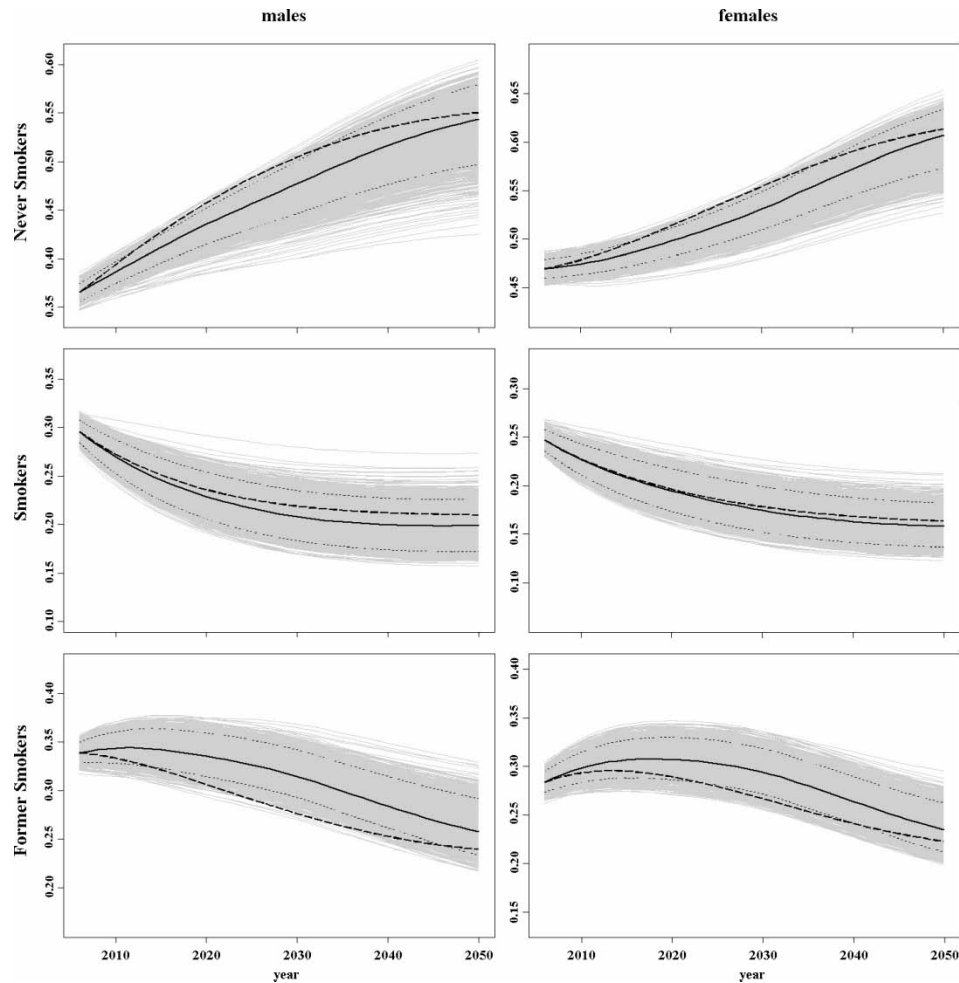


Figure 4. Prevalence rates projections in the Dutch population aged 15–85 specified by gender. Black and curve, prevalence projections; short dashes, standardized prevalence projections; black dotted line, 95% pointwise confidence interval; and grey solid curves, simulated prevalence projections.

the input prevalence rates for the ages 15–35 remain influential, while input transition rates increase their relevance with ongoing time.

The projected never smokers prevalence rate is positively correlated with the input never smokers prevalence rates, but negatively with the input prevalence rates of the other two smoking factor classes. Analogously, smokers and former smokers prevalence projections are positive correlated with the smokers input prevalence rates and former smokers input prevalence rates, respectively. Output prevalence rates are positive correlated with the input transition rates that determine the flow into this class (for example, the output smokers prevalence rate is positively correlated with the input relapse rates), and analogously there is a negative correlation with the flow out of the class (for example, the output never smokers prevalence rates and the input start rates). In the situations in which a transition rate does not cause a flow in or out of a class, the correlation between this type of transition rate and the output prevalence rate of that class will be null, at least during the first



years of the projection period. For example, the correlation coefficients between the projected prevalence rates of never smokers and the quit rates and the relapse rates are null. The projected prevalence rate of former smokers is not correlated with the input start rates at the beginning of the projection period, but the correlation increases with time; this is reasonable since the former smokers prevalence rates are related to the smokers prevalence rates and through this to the start rates.

The start rates for ages 15–19 were obtained in an alternative way compared to the other ages. They were not independently estimated from the data, but estimated from difference in prevalence rates between adjoining ages. Due to this a discontinuity is seen in the trend of the correlation coefficient for the start rates around the input age of 20. We also performed the Monte Carlo simulation experiment with start rates estimated directly from the data. In that case, the discontinuity in the correlation coefficient disappears.

#### 4. Discussion

State-transition models are intuitive methods to trace future smoking behaviour within a population, but the reliability of their results is based on the validity of the assumptions made and on the accuracy of the input data they require. An important assumption is that future transition rates are equal to that observed in 2003. Another assumption is that transition rates and especially the probability of relapse do not depend on the length of time someone is already in a state. This is probably not true for relapse rates. However, our projection probably partly takes this into account because the age dependency in relapse rate partly reflects different stopping durations in different ages.

As our uncertainty and sensitivity analysis shows, the importance of the input variables (transition rates and prevalence rates) in determining the variability of the state-transition model output varies greatly with the year the projections refer to, as well as with the kind of output. Hence, the level of the concern about the reliance of the input data will depend also on the type of projections needed: short or long period projection. For example, if the prevalence rate of never smokers in 2050 is the output of interest, great care should be put in the estimation of the start rates. The database we used for this research comes from a cross-sectional survey investigating the smoking behaviour within the Dutch population. Therefore, we were able to work with data which allowed us to obtain prevalence rates and transition rates directly from the same source.

Concerning the way we modelled the input rates we would like to underline once more the modification introduced in the start rates from age 15 to 19. This modification was needed because the start rates at younger ages obtained through the STIVORO survey are highly unstable. Recent time trends of smoking prevalence (data not shown) show that prevalence rates at young ages have been stable over the last years, making our approach a reasonable one.

Our sensitivity analysis shows that prevalence rates at young ages are important for future trends, but that transition rates are more important. Therefore, data collection should focus on these transition rates, also in order to monitor developments that will influence future smoking trends.

With the state-transition model developed in this paper we obtained projections that show a noteworthy but rather slow decline of the proportion of smokers in the Dutch population within the next decades, both for males and females, and at the same time an increase of people who never smoked. Trends for the percentage of former smokers are different by gender, but nevertheless will get to the same level of around 25% in 2050 (in particular 26.1% for males and 23.5% for females). Prevalence rate projections obtained

for the whole Dutch population aged 15–85 show that a prevalence of smokers in 2010 of 24.8% (95% CI: 23.7–26). This suggests that reaching the national target of 20% of smokers [9] in 2010 would be difficult without adopting any further strategy since a difference of almost the 5% consists of more than 600,000 smokers who should not start and/or should quit smoking.

### Acknowledgements

The authors are grateful to Dr Rudolf Hoogenveen, Dr Talitha Feenstra and Dr Pieter van Baal for their help and their suggestions. This research has been sponsored by the Dutch Ministry of Health, Welfare and Sport, Den Haag, The Netherlands.

### References

- [1] WHO, *Tobacco: Deadly in Any Form or Disguise. World No Tobacco Day*, World Health Organization, Geneva, Switzerland, 2006.
- [2] P. Jha and F.J. Chaloupka, *Curbing the Epidemic: Governments and the Economics of Tobacco Control*, The World Bank, Washington DC, (Development in Practice) 1999.
- [3] R.T. Hoogenveen, A.E.M. de Hollander, and M.L.L. van Genugten, *The Chronic Disease Modelling Approach*, RIVM, Bilthoven, 1998.
- [4] D.T. Levy, J.E. Bauer, and H. Lee, *Simulation modeling and tobacco control: Creating more robust public health policies*, Am. J. Public Health 96(3) (2006), pp. 494–498.
- [5] L. Gunning-Schepers, *The health benefits of prevention: A simulation approach*, Health Policy 12(1/2) (1989), pp. 1–255.
- [6] T.L. Feenstra, P.H.M. van Baal, R.T. Hoogenveen, S.M.C. Vijgen, E. Stolk, and W.J.E. Bemelmanns, *Cost-Effectiveness of Interventions to Reduce Tobacco Smoking in the Netherlands. An Application of the RIVM Chronic Disease Model*, RIVM, Bilthoven, 2006.
- [7] P.H.M. van Baal, R.T. Hoogenveen, G.A. de Wit, and H.C. Boshuizen, *Estimating health-adjusted life expectancy conditional on risk factors: Results for smoking and obesity*, Popul. Health Metr. 4(1) (2006), p. 14.
- [8] A.D. Lopes, N.E. Collishaw, and T. Piha, *A descriptive model of the cigarette epidemic in developed countries*, Tob. Control 3 (1994), pp. 242–247.
- [9] Ministerie van Volksgezondheid, Welzijn en Sport Nationaal Programma Tabaksontmoediging 2006–2010. Den Haag. (Ministry of Health, Welfare and Sport. National Program for Tobaccodiscouragement 2006–2010) (2006).
- [10] A. Saltelli, K. Chan, and E.M. Scott, *Sensitivity Analysis*, Wiley, New York, 2000.
- [11] A. Saltelli, S. Tarantola, F. Campolongo, and M. Ratto, *Sensitivity Analysis in Practice. A Guide to Assessing Scientific Models*, Wiley, New York, 2004.
- [12] Statistics Netherlands, Statline. [<http://www.cbs.nl>] (Accessed: June 2006).
- [13] C.F. van Kreijl, A.G.A.C. Knaap, and J.M.A. van Raaij, *Our Food, Our Health – Healthy Diet and Safe Food in The Netherlands*, RIVM, Bilthoven, 2006.
- [14] Statistics Netherlands, Statline. [<http://www.cbs.nl>] (Accessed: April 2007).
- [15] S. Greenland, *Dose–response and trend analysis in epidemiology: Alternatives to categorical analysis*, Epidemiology 6(4) (1995), pp. 356–365.
- [16] W.N. Venables and B.D. Ripley, *Modern Applied Statistics with S-PLUS*, 4th ed. Springer, New York, 2002.