Community Dent Oral Epidemiol 2006; 34: 123–9

COMMUNITY DENTISTRY AND ORAL EPIDEMIOLOGY

Application of nonhomogenous Markov models for analyzing longitudinal caries risk

D. T. Kopycka-Kędzierawski and R. J. Billings

Eastman Department of Dentistry, University of Rochester, Rochester, NY, USA

Kopycka-Kędzierawski DT, Billings RJ. Application of nonhomogenous Markov models for analyzing longitudinal caries risk. Community Dent Oral Epidemiol 2006; 34: 123–9. © Blackwell Munksgaard, 2006

Abstract – *Objectives:* Markov modeling is a useful mathematical procedure for calculating probabilities of disease prognosis. Increasingly, Markov models are being applied in medical and health services research and also in social sciences research. The purpose of our study was to use the Markov process to determine time-dependent transition probabilities for caries-free children to convert to a caries-active state and to assess the impact of salivary mutans streptococci (MS) levels on caries status. *Methods:* Our analysis was based on data obtained from a 6-year longitudinal study of risk factors associated with caries onset in children. *Results:* Based on a two-state Markov model, the probability that a caries-free child would convert to a caries-active state during the study ranged between 0.0046 and 0.0471. The highest probability of converting from a caries-free state to a caries-active state was 0.0471 at age 8.5 years. *Conclusions:* In addition to standard statistical methods of analyzing longitudinal caries data, Markov models show promise for use in the analysis of caries risk.

Key words: caries risk; longitudinal data; mutans streptococci; nonhomogenous Markov models

D. T. Kopycka-Kędzierawski, Eastman Department of Dentistry, University of Rochester, 625 Elmwood Avenue, Rochester, NY 14620, USA
Tel: 585-275-1139
Fax: 585-273-1237
e-mail: dorota_kopyckakedzierawski@urmc.rochester.edu

Submitted 26 April 2005; accepted 3 August 2005

The common methods for assessing caries risk are by means of univariate or multivariate models that employ traditional statistical techniques such as odds ratios and regression analysis. These approaches to longitudinal caries risk analysis are useful when risk for dental caries is constant over time, the attrition rate is small and tooth loss occurs infrequently. These methods, however, do not incorporate ongoing risk over time in the resultant risk models.

To analyze ongoing risk over time, we used Markov models to analyze caries onset in initially caries-free children. Markov analysis is a dynamic mathematical modeling technique that is useful for predicting prognoses and probabilities of diseases within populations over time. The finite, discrete Markov process assumes that an individual is always in one of a finite number of states of health referred to as Markov states and the process is described by transition probabilities from one state to another. The time horizon of the analysis is divided into equal increments of time, referred to as Markov cycles. The probability of making a transition from one state to another during a single cycle is referred to as a transition probability. A transition probability is Markovian if it only depends on the present state and not on the previous history of the individual, i.e. the transition probability depends upon the health state persons are in and not how long they have been in that health state or how they arrived there (1). In order for a Markov process to terminate, it must have at least one state that an individual cannot leave. Such states are called absorbing states. In medicine, the absorbing states usually represent death; in our example the absorbing state is caries onset.

Increasingly, Markov models are being applied in medical and health services research and also in social sciences research. For example, a Markov model was successfully used to project tuberculosis incidence in the United States for the years 1980–2010 in disaggregated demographic groups (2). In the social sciences, Markov models were developed to study social capital and violence in the United

States (3). In cancer research, Markov models were constructed to estimate the rate at which dysplasia will progress to cervical cancer (4). Markov modeling has also been used to describe the dynamics of disease progression in sepsis (5).

Markov models offer a unique advantage in that they provide a convenient way of modeling data when decision-making involves a risk that is ongoing over time (6). Thus, they are widely used in cost-effectiveness research as they allow the incremental use of resources and the consequent health benefits to be assessed for any point in time. Despite their applicability, Markov models have not been extensively used in dentistry. A few examples of its application exist, nevertheless. Lu (7) suggested using finite absorbing Markov chains to describe the caries process. Kay and Nuttall (8) attempted to determine transition probabilities for predicting the dental caries status of Scottish children. More recently, Helfenstein et al. (9) applied a Markov chain Monte Carlo method to construct elaborate models in dentistry. Higashi et al. (10) used Markov modeling to evaluate long-term incremental clinical and economic outcomes associated with genetic testing of patients with mild periodontal disease.

Thus, the purpose of this paper is to describe the utility of time-dependent Markov models for the analysis of longitudinal caries data and to assess the impact of mutans streptococcal (MS) levels on caries status.

Materials and methods

Subjects

A cohort of 631 caries-free children, 6–7 years of age, from Rochester, New York and the Finger Lakes region of western New York State who participated in a longitudinal study to identify risk factors associated with caries onset formed the study population for this analysis. Each child was examined at 6-month intervals for up to 6 years. The study was completed in 1996. Cariogenic microbiological profiles were among the risk factors assessed for susceptibility to caries. The complete set of risk factors assessed has been described elsewhere (11, 12).

Two-state Markov model versus three-state Markov model

For the purpose of our approach to the analysis of the data using Markov models, only MS data from the aforementioned risk studies were utilized. Thus, to describe the clinical history of caries onset in initially caries-free children over a 6-year study period, we developed two Markov models: a two-state and a three-state model.

The two-state Markov model, as shown in Fig. 1(a), consists of two health states: (i) a caries-free state and (ii) a caries-active state. The two-state model is described by transition probabilities p_{CF} and p_{CA}. Here, p_{CF} denotes the probability that a child would remain caries-free during a given cycle; p_{CA} denotes the probability that a child would convert from a caries-free to a caries-active state during a 6-month cycle.

A three-state Markov model was constructed to assess the impact of salivary MS levels on remaining caries-free and for transitioning between low and high MS levels over a 6-year study period. The three-state model, as shown in Fig. 1(b), consists of three health states: (i) a caries-free and low MS, (ii) a caries-free and high MS and (iii) a caries-active state. A high level of MS was defined as $\geq 10^6$ colony-forming units (CFU)/ml of saliva, based on the work of Klock and Krasse (13). A low level of MS was defined as $<10^6$ CFU/ml of saliva. There are six possible transitions in the three-state Markov model for each cycle. These are: transitions from a low MS caries-free state to a low MS caries-free state (p_{LL}), from a low MS caries-free state to a high

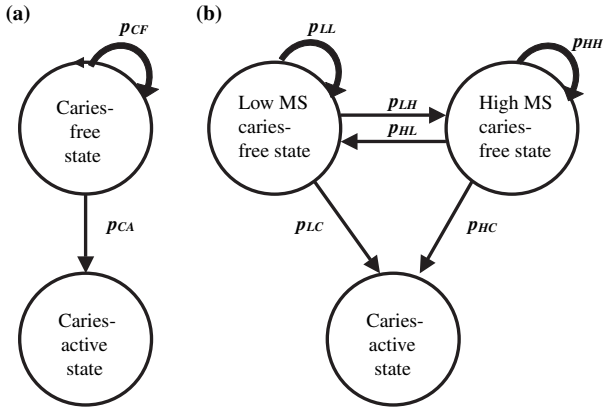

Fig. 1. (a) Two-state Markov model; (b) three-state Markov model. P_{CA}, transition probability from a caries-free to a caries-active state; P_{CF}, transition probability from a caries-free to a caries-free state; P_{LL}, transition probability from a low MS caries-free state to a low MS caries-free state; P_{LH}, transition probability from a low MS caries-free state to a high MS caries-free state; P_{LC}, transition probability from a low MS caries-free state to a caries-active state; P_{HH}, transition probability from a high MS caries-free state to a high MS caries-free state; P_{HL}, transition probability from a high MS caries-free state to a low MS caries-free state; P_{HC}, transition probability from a high MS caries-free state to a caries-active state.

MS caries-free state (p_{LH}), from a low MS caries-free to a caries-active state (p_{LC}), from a high MS caries-free state to a high MS caries-free state (p_{HH}), from a high MS caries-free state to a low MS caries-free state (p_{HL}) and from a high MS caries-free to a caries-active state (p_{HC}). Fig. 1(b) illustrates all possible transitions among the different states. The sum of all possible transitions from a low MS caries-free state equals 1. The sum of all possible transitions from a high MS caries-free state equals 1, as well. The absorbing state in both of our models is caries onset in the permanent dentition (filled and missing teeth were preceded by caries). In a few occurrences, restorations were not preceded by a diagnosis of caries in the prior examination and were excluded from the analysis.

Transition probabilities

To calculate the transition probabilities we used the maximum likelihood method. That is, each transition probability was estimated by the number of observed transitions during a given cycle, divided by the total number of all possible observations during that cycle. By definition, the above formula requires two consecutive clinical examinations and bacterial screenings per individual to calculate a transition probability. For example, to calculate the transition probability p_{LL} from a low MS caries-free state to a low MS caries-free state in any given cycle, we would divide the number of observed transitions from the low caries-free state to the low caries-free state by the number of observations within that cycle, i.e. the number of observations of caries-free children who had a low level of MS at the start of the cycle. To demonstrate trends in the transition probabilities more clearly, we smoothed the empirical data by using the method of moving averages (14). The transition probabilities in our models are time-dependent. Such models are called nonhomogeneous Markov models.

Transition probabilities were calculated for the permanent dentition in both models (28 permanent teeth, excluding third molars). Clinical caries was diagnosed using a modified version of the visual-tactile criteria of Radike (15), i.e. the dental explorer was used only to remove plaque or to detect the presence of translucent sealants. Radiographs were not taken. For the purpose of this analysis, caries onset was defined as a transition from a caries-free state to a caries-active state, i.e. when at least one caries lesion was present in the permanent dentition, an individual was determined to have transitioned to a caries-active state.

Results

The calculated transition probabilities for the two-state model are shown in Table 1(a). The transition probabilities are represented graphically in Fig. 2. Cycle 1, the initial cycle, represents the transitions between ages 6.5 and 7 years; cycle 2 represents the transitions between ages 7 and 7.5 years. Cycle 11, the final cycle, represents the transitions between ages 11.5 and 12 years. The probability that a caries-free child would convert to a caries-active state during one cycle ranged between 0.0046 and 0.0471. The highest probability of converting from a caries-free state to a caries-active state was 0.0471 at age 8.5 years. The lowest probability of developing caries in the next six months was 0.0046 at age 6.5 years.

The transition probabilities calculated for the three-state model are shown in Table 1(b). The transition probabilities are represented graphically in Fig 3(a,b). The transition probability for a caries-free child with low levels of salivary MS to remain in a caries-free state with low levels of salivary MS during the study ranged between 0.841 and 0.882. The transition probabilities from a low level to a high level of MS ranged between 0.105 and 0.147. Based on our data presented in Table 1(b), we calculated that children identified with low levels of MS at baseline, on average (81% of all children remaining caries-free) would be identified with low levels of MS during the course of the study. The probability that a caries-free child with a low level of MS would convert to a caries-active state during the study ranged between 0.005, at age 6.5 years and 0.031 at age 8.5 years. For a caries-free child with a high level of MS, the probability to convert to a caries-active state ranged between 0.0028 at age 6.5 years and 0.092 at age 8.5 years. Usually, children with low levels of salivary MS would have a lower probability of converting to a caries-active state than children with high levels of salivary MS (Fig. 3c). For each cycle, we performed tests for comparison of two proportions (probability of converting to a caries-active state for children with low levels of salivary MS and probability of converting to a caries-active state for children with high levels of salivary MS). Based on these tests (the probabilities of converting to a caries-active state for children with low levels of salivary MS vis a vis children with high levels of salivary MS were statistically significantly different between ages 7.5 and 10 years) they correspond to cycles from 3 to 7 (P-value < 0.05) (Fig. 3c).

Table 1. Transition probabilities by age for the two-state (a) and the three-state (b) Markov models

(a) Cycle	Age (years)	P_{CA}	P_{CF}
1	6.5–7	0.0046	0.9954
2	7–7.5	0.0094	0.9906
3	7.5–8	0.0195	0.9805
4	8–8.5	0.0319	0.9681
5	8.5–9	0.0471	0.9529
6	9–9.5	0.0418	0.9582
7	9.5–10	0.031	0.969
8	10–10.5	0.0178	0.9822
9	10.5–11	0.0176	0.9824
10	11–11.5	0.0129	0.9871
11	11.5–12	0.0173	0.9827

(b) Cycle	Age (years)	P_{LL}	P_{LH}	P_{LC}	P_{HL}	P_{HH}	P_{HC}
1	6.5–7	0.8766	0.1184	0.005	0.3405	0.6567	0.0028
2	7–7.5	0.8466	0.1463	0.007	0.3433	0.6433	0.0134
3	7.5–8	0.8411	0.1465	0.0124	0.3399	0.624	0.0361
4	8–8.5	0.8488	0.1329	0.0183	0.3909	0.5439	0.0652
5	8.5–9	0.8504	0.1182	0.0314	0.4169	0.4914	0.0916
6	9–9.5	0.8659	0.1073	0.0268	0.474	0.4349	0.0911
7	9.5–10	0.8823	0.0943	0.0234	0.4827	0.4581	0.0593
8	10–10.5	0.8802	0.1051	0.0148	0.4745	0.4955	0.03
9	10.5–11	0.8631	0.119	0.0179	0.4375	0.5466	0.0159
10	11–11.5	0.8678	0.12	0.0122	0.3539	0.6302	0.0159
11	11.5–12	0.8693	0.1108	0.0199	0.3688	0.6231	0.0081

P_{CA}, transition probability from a caries-free to a caries-active state; P_{CF}, transition probability from a caries-free to a caries-free state; P_{LL}, transition probability from a low MS caries-free state to a low MS caries-free state; P_{LH}, transition probability from a low MS caries-free state to a high MS caries-free state; P_{LC}, transition probability from a low MS caries-free state to a caries-active state; P_{HH}, transition probability from a high MS caries-free state to a high MS caries-free state; P_{HL}, transition probability from a high MS caries-free state to a low MS caries-free state; P_{HC}, transition probability from a high MS caries-free state to a caries-active state.

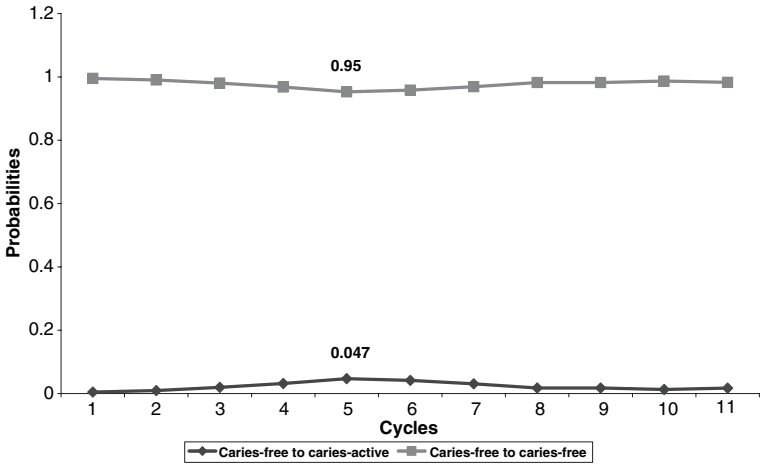

Fig. 2. Transition probabilites for the two-state Markov model.

Discussion

The advantage of Markov models is their simplicity in modeling the natural history of chronic disease. Any given disease can be described by a mathematical model as a series of health states and possible transitions among them when the risk of disease onset or death is ongoing over time. One important limitation of the Markov processes, however, is lack of memory, i.e. often the previous health states of the population are not always known. By using a nonhomogeneous Markov model, an accurate description of the biological process, even when the Markovian assumption is not fully satisfied, can be obtained. We built nonhomogeneous Markov models with time-dependent transition probabilities to

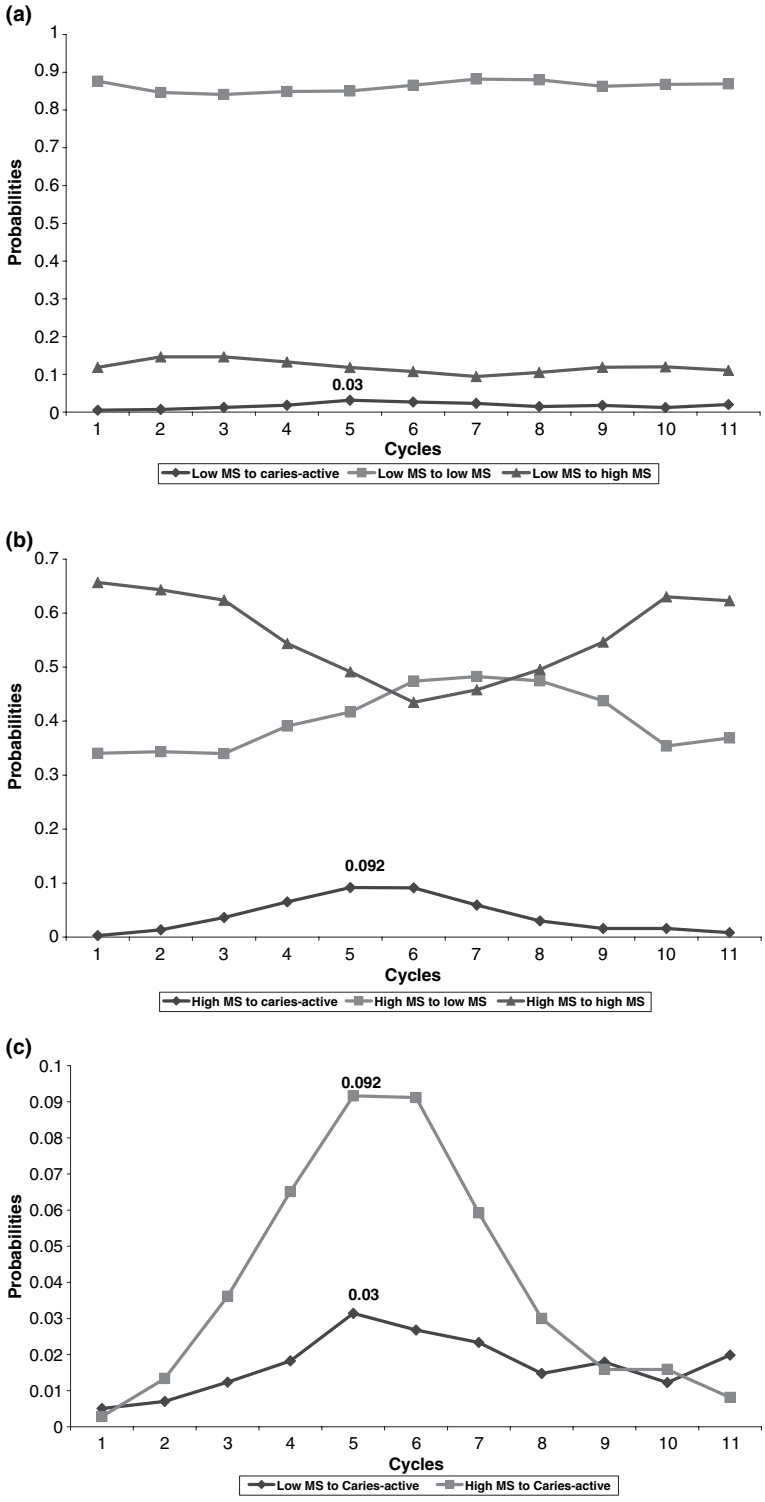

Fig. 3. (a) Transition probabilites for the three-state Markov model (low MS), (b) transition probabilites for the three-state Markov model (high MS), (c) transition probabilites from caries-free to caries-active state (low MS and high MS).

avoid discrepancies between the predictions of our model and observed data. In contrast to standard statistical methods such as regression analysis, Markov models show the dynamics of the disease process over time. Additionally, they incorporate ongoing risk over time.

By using Markov modeling techniques, we calculated transition probabilities for caries onset for different age groups of initially caries-free children. We determined transition probabilities for converting from a caries-free state to a caries-active state; remaining caries-free; and for transitioning between low and high MS levels over a 6-year study period. By incorporating low and high salivary MS states into our model we were able to assess the effect of salivary MS levels on caries

onset in initially caries-free children. Our results suggest that Markov models show promise as an alternative statistical approach to identify children at risk for caries and these methods may be used as additional research tools when presenting dental data. The advantage of Markov models is in their simplicity of describing dynamic disease processes, as in our example, describing caries onset in initially caries-free children. Markov modeling techniques are convenient and simple to implement. They are often used in cost-effectiveness analyses to compare two or more treatment strategies in terms of their costs and effectiveness. The data upon which we based our Markov models were analyzed previously by means of the Kaplan–Meier survival method. Based on that analysis, we concluded that caries-free children who had high salivary MS levels at baseline would have a greater risk of caries onset at any given time than caries-free children who had low salivary MS levels at baseline (16). By using a three-state Markov model, we observed that in general, children with low levels of salivary MS would have a high probability to stay in this state and lower probability of converting to a caries-active state than children with high levels of salivary MS, thus children who were caries-free at baseline and who had high salivary MS levels at baseline would be at greater risk, i.e. more susceptible to caries onset, at any given time than caries-free children who had low salivary MS levels at baseline during the course of the study.

Transition probabilities can also be used to compare the cost-effectiveness of two or more treatment strategies when end points of the analysis other than mortality are studied. For example, Higashi and colleagues used Markov models to evaluate different clinical scenarios for progression from mild to severe periodontal disease (10). We used a Markov model to assess the cost-effectiveness of a hypothetical intervention strategy to reduce the risk of caries onset in initially caries-free children, as compared to a nonintervention strategy (17). In that model, we suggested that the cost-effectiveness of a hypothetical caries preventive strategy, based on a univariate approach or single risk factor model, i.e. salivary MS levels, was clinically beneficial, as 9.5% fewer children in the intervention group would have became caries-active when compared with the control group.

In summary, the model outlined in this paper describes an alternative statistical method that enhances and expands our ability to predict caries risk. At a conceptual level, this approach incorporates ongoing risk over time and, on the statistical level, use of the Markov process may lead to new ways of modeling caries risk. Our aim in this study was to apply the Markov process in the analysis of longitudinal caries data and to present this method as an additional way of analyzing dental data. By estimating the proportions of individuals who would be in the specified health states up to 6 years from baseline, we obtained a comprehensive view of the risk of caries onset. We used caries onset, independent of severity, as a case definition in both models. Future multivariate models will be built to assess the utility of the Markov process to predict caries progression once the transition from a disease-free state to a disease-active state has occurred. For example, we will assess the utility of the Markov process when previously identified variables associated with caries risk are incorporated into the model (11, 12) and we will explore the possibility of using stochastic processes in modeling caries data (18).

Acknowledgements

The authors gratefully acknowledge and thank Drs Dennis H. Leverett (since deceased) and John D.B. Featherstone for their leadership and intellectual contributions to the study upon which this paper is based, Ms Sheila Mundorff-Shrestha for her meticulous oversight of the microbiological assays and Clare Shaffer for her invaluable assistance in the administration and execution of the clinical study. Supported by NIH/NIDCR T32 DE 07328 and R01 DE 08946.

References

1. Drummond MF, O'Brien B, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. 2nd edn. New York: Oxford University Press; 1997.
2. Debanne SM, Bielefeld RA, Cauthen GM, Daniel TM, Rowland DY. Multivariate Markovian modeling of tuberculosis: forecast for the United States. Emerg Infect Dis 2000;6:148–57.
3. Galea S, Karpati A, Kennedy B. Social capital and violence in the United States, 1974–1994. Soc Sci Med 2002;55:1373–83.
4. Sawaya GF, McConnell KJ, Kulasingam SL, Lawson HW, Kerlikowske K, Melnikow J, et al. Risk of cervical cancer associated with extending the interval between cervical-cancer screenings. N Engl J Med 2003;349:1501–9.
5. Rangel-Frausto MS, Pitter D, Hwang T, Woolson RF, Wenzel RP The dynamics of disease progression in sepsis: Markov modeling describing the natural

history and likely impact of effective antisepsis agents. Clin Infect Dis 1998;27:185–90.
6. Sonnenberg FA, Beck R. Markov models in medical decision making: a practical guide. Med Decis Making 1993;13:322–38.
7. Lu KH. An analysis of the caries process by finite absorbing markov chains. J Dent Res 1966;45:998–1015.
8. Kay EJ, Nuttall NM. An example of the application of Markov models in dentistry – a prediction of the fate of molar teeth in 12-year-old Scottish children. Community Dent Health 1993;10:217–33.
9. Helfenstein U, Menghini G, Steiner M, Murati F. An example of complex modeling in dentistry using Markov chain Monte Carlo (MCMC) simulation. Community Dent Health 2002;19:152–60.
10. Higashi MK, Veenstra DL, del Aguila M, Hujoel P. The cost-effectiveness of interleukin-1 genetic testing for periodontal disease. J Periodontol 2002;73:1474–84.
11. Leverett DH, Featherstone JDB, Proskin HM, Adair SM, Eisenberg AD, Mundorff-Shrestha SA et al. Caries risk assessment by a cross-sectional discrimination model. J Dent Res 1993a;72:529–37.
12. Leverett DH, Proskin HM, Featherstone JDB, Adair SM, Eisenberg AD, Mundorff-Shrestha SA et al. Caries risk in a longitudinal discrimination study. J Dent Res 1993b;72:538–43.
13. Klock B, Krasse B. Effect of caries-preventive measures in children with high numbers of *S. mutans* and *lactobacilli*. Scand J Dent Res 1978;86:221–30.
14. Selvin S. Statistical Analysis of Epidemiologic Data. New York: Oxford University Press; 1991.
15. Radike AW. Criteria for diagnosis of dental caries: Proceedings of the Conference on the Clinical Testing of Cariostatic Agents, 1968. Chicago, IL: American Dental Association; 1972. p. 87–8.
16. Kopycka-Kędzierawski DT, Billings RJ. A longitudinal study of caries onset in initially caries-free children and baseline salivary mutans streptococci levels; a Kaplan–Meier survival analysis. Community Dent Oral Epidemiol 2004;32:201–9.
17. Kopycka-Kędzierawski DT. Cost-effectiveness model of caries risk assessment in children, Thesis, Rochester, NY: University of Rochester; 2003.
18. Cox DR, Miller HD. The theory of stochastic process. Bristol: Arrowsmith; 1987.