BIOSTATISTICS

ebph

Comparison of multi-state Markov models for cancer progression with different procedures for parameters estimation. An application to breast cancer

Leonardo Ventura[1], Giulia Carreras[1], Donatella Puliti[1], Eugenio Paci[1], Marco Zappa[1], Guido Miccinesi[1]

ABSTRACT

BACKGROUND: the knowledge of sojourn time (the duration of the preclinical screen-detectable period) and screening test sensitivity is crucial for understanding the disease progression and the effectiveness of screening programmes. For this purpose a model of the natural history of the disease is needed. The aim of this work is to provide an illustration of the application of multistate Markov models for breast cancer progression to the data of the Florentine screening programme, in order to estimate the sojourn time and sensitivity for breast cancer screening.

METHODS: three different multi-state Markov models of increasing complexity were used with three different estimation procedures based on non-linear least squares, maximum likelihood, and on a Bayesian approach. All the models produced estimates for screening sensitivity and mean sojourn time. The data used in our application seem to lead to a non-identifiability problem, since the estimation procedures for both the Maximum Likelihood and Non-Linear Least Squares gave estimates that changed with the parameters' initial values or difficultly converged. In order to take this problem into account we used the Bayesian Approach by incorporating prior information on all the parameters.

RESULTS: the mean sojourn time varied between 2-7 years and 3-5 years for women aged 50-59 and 60-69, respectively. When the model complexity was increased a higher variability in estimates was observed among the estimation procedures. The results of the screening sensitivity estimates were highly variable, both among estimation techniques and models - varying between 63% and 100%, and between 77% and 100% for women aged 50-59 and 60-69, respectively.

CONCLUSIONS: results are in accord with the literature; those obtained through the Bayesian Approach seem to be more reliable.

Key words: Markov model; Mean sojourn time; Sensitivity; Non-linear least squares; Maximum likelihood; Gibbs sampling

(1) Cancer Prevention and Research Institute (ISPO), Florence, Italy

CORRESPONDING AUTHOR: Leonardo Ventura, Cancer Prevention and Research Institute (ISPO), via delle Oblate 2, 50141 Florence, Italy. Tel: +39 055 7972502; Fax: +39 055 7972535. e-mail: l.ventura@ispo.toscana.it

INTRODUCTION

Early detection by mass screening programmes is of major public health importance for control of at least some kinds of progressive chronic disease. By advancing the diagnosis of disease, the hope is to reduce the risk of serious consequences, especially death. For cancer, the definitive demonstration that early detection is of benefit is a reduction in disease-specific mortality in a screened group.

To properly define the more appropriate screening intervals and understand the effect of screening activities on the incidence of cancer, a model of the natural history of the disease is needed.

Following Day et al [1], we assume that the disease progresses as shown in Figure 1. If no intervention were to take place, an individual would enter the preclinical phase of the disease detectable by screening at time T_0 and would begin to manifest symptoms, i.e. the disease would become clinically apparent at time T_1. For the individual, the sojourn time is defined as $T_1 - T_0$. Suppose that a person is screened at time T_2 ($T_0 < T_2 < T_1$) and the preclinical stage is diagnosed. For the individual, the lead time - the interval by which diagnosis is brought forward - is defined as $T_1 - T_2$. The probability that the screening test correctly identifies an individual as being in the preclinical phase is termed the "sensitivity" (Se) of the test.

The time that a disease remains in the preclinical detectable phase (PCDP), the sojourn time, is important in determining the potential effectiveness of a screening programme. It provides an absolute upper limit to the lead time obtainable and indicates what interval between screens is likely to achieve a useful lead time. If the sojourn time is long, the maximum attainable lead time is correspondingly long. If the sojourn time is short, however, the potential gains from screening are smaller and screening should take place more frequently in order to increase the probability that preclinical disease is found before it surfaces to the clinical stage. Furthermore, lead time estimates are needed to properly interpret changes in cancer incidence after the start of a screening program and, particularly, to evaluate the presence of overdiagnosis.

Different models have been developed to estimate mean sojourn time (MST) and Se, for breast as well as for other cancer screening programmes.

Day and Walter [1] proposed a simplified model to estimate both the MST and the test Se for breast cancer screening and showed that the sojourn time distribution is well approximated by an exponential distribution. The exponential distribution implies that the average advance in diagnosis due to screening is equal to the MST and that the standard deviation of the sojourn time is equal to its mean.

Using generalized linear models, Paci and Duffy [2] used information on interval cancers between screens by applying an exponential form for the sojourn time distribution, to estimate MST, Se and positive predictive value based on the Florence breast cancer screening programme.

SCHEME FOR THE PROGRESSION OF THE DISEASE
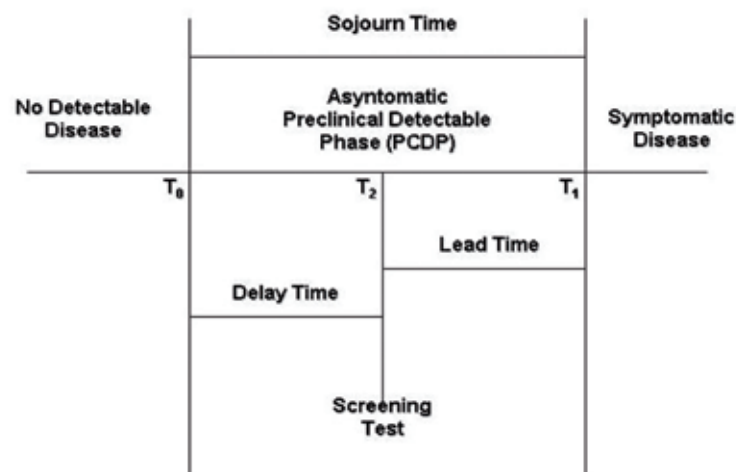
This model was modified by Prevost et al [3] and applied to colorectal cancer screening to take into account the possibility that the Se is substantially less than 100%, using a two states Markov model. More complex multi-state Markov models were successively proposed to assess the natural history of the disease [4, 5]. Multi-state Markov models - which consider more states in modeling the natural history of the disease - allow for the specification of the transition rates from preclinical to clinical states through different pathways (dimension, nodes involvement, non-progressive disease).

Uhry et al [6] applied multi-state Markov models to estimate natural history parameters that were applied to data from breast screening programmes in three districts in France. Using the node status as a prognostic factor, they applied a multi-state Markov model in order to estimate the Se of the screening procedure and the MST, and to predict the mortality reduction associated with screening, To explore the effect of overdiagnosis on parameter estimates, they added a non-progressive preclinical state.

Using data from Finnish screening program, Wu et al [7] presented a refinement of the multi-state Markov model to estimate the progression rate of breast cancer, making allowance for measurement errors and for different methods to detect breast cancer including cancers arising from non-participants.

The aim of this study is to provide an illustration of the application of multi-state Markov models for breast cancer progression to data from the first two rounds of the Florentine screening programme (1991-1993). Three different multi-state Markov models that are widely used in the literature were applied, and three different estimation procedures (non-linear least squares, maximum likelihood, Bayesian approach) were used. Limits and advantages of these different methods are discussed.

METHODS

We modeled the natural history of breast cancer - describing the progression of the tumour from the asymptomatic phase to the clinical phase - according to three different multi-state Markov models [3, 4, 5]: two, three, and five-state models.

The Markov model provides a convenient way to model prognosis for clinical problems. The model assumes that the patient is always in one of a finite number of states of health, which are termed Markov states. All events of interest are modeled as transitions from one state to another.

In this paper, multi-state Markov stochastic processes which we apply are defined by the following properties:

- The transition rates between the states are invariant with time.
- The sojourn time in a given state follows an exponential distribution.
- The sojourn times in successive states are independent.
- The probability to transition from one state a, to another state b, is independent of the time spent in state a.
- The probability to transition from one state a, to another state b, is independent of the past states before a (Markov property).

Two-state Markov model

According to Prevost et al [3], we defined a two-state model for the natural history of the disease where the states are preclinical breast cancer that are detectable by screening (asymptomatic) and clinical breast cancer (symptomatic), and λ_{12} is the transition rate from the first state to the other (Figure 2). The inverse of the transition rate, $1/\lambda_{12}$, is the MST under the exponential distribution.

FIGURE 2

SCHEME OF THE TWO-STATE MARKOV MODEL
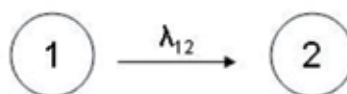
State (1) is the asymptomatic breast cancer, state (2) is the symptomatic breast cancer, and λ_{12} is the transition rate from state 1 to state 2. The model has been extensively presented by Prevost et al [3]. A more detailed presentation of the model is in the Supplementary Materials of this paper

Three-state Markov model

According to the model proposed by Duffy et al [4] and revised by Uhry et al [6] we suppose that the natural history of the disease is described by three health states: no detectable breast cancer, preclinical cancer detectable by screening (asymptomatic) and clinical breast cancer (symptomatic) (Figure 3). The transition rate λ_{01} represents the incidence rate of the preclinical disease, and λ_{12} is the transition rate from asymptomatic to symptomatic breast cancer. As in the two-state model, the MST is equal to $1/\lambda_{12}$ under the exponential assumption. The model is presented in detail in the Supplementary Materials.

Five-state Markov model:

The five-state model differentiates the progression of a tumour according to node status, i.e. node negative (PN-), and node positive (PN+). As a consequence, the overall MST in the preclinical phase can be calculated by the sum of the MST in the preclinical phase for PN- (state 1) and MST in the preclinical phase for PN+ (state 2) multiplied by the proportion of cancer moving from state 1 to state 2 (Figure 4).

Estimation procedures

For each of the models described, the parameters were estimated using three different techniques: maximum likelihood (ML), non-linear least squares (NLS), and a Bayesian approach (BA) using the Gibbs sampling technique.

ML and NLS are among the most frequently used procedures to estimate the parameters of a large variety of statistical models, including multi-state Markov models.

The likelihood function is the joint probability (density) function of observed data expressed as the function of the unknown parameters. The idea behind the maximum likelihood parameter estimation is to determine the parameters that maximize the probability of
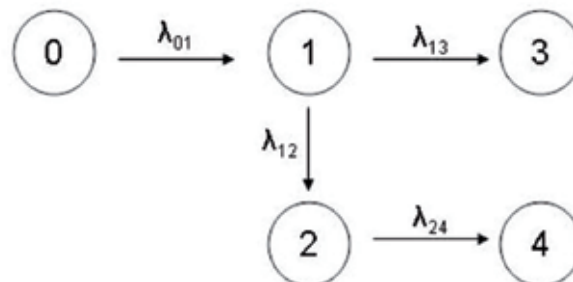
FIGURE 3

SCHEME OF THE THREE-STATE MARKOV MODEL

State (0) is where there is no detectable breast cancer, state (1) is the asymptomatic breast cancer, state (2) is the symptomatic breast cancer, and λ_{01} and λ_{12} are the transition rates between the states [6]

FIGURE 4

SCHEME OF THE FIVE-STATE MARKOV MODEL

State (0) is where there is the no detectable breast cancer, state (1) is the preclinical cancer PN- detectable by screening (asymptomatic), state (2) is the preclinical breast cancer PN+ detectable by screening (asymptomatic), state (3) is the clinical breast cancer PN- (symptomatic), and state (4) is the clinical breast cancer PN+ (symptomatic). λ_{01}, λ_{12}, λ_{13} and λ_{24} are the transition rates between the states [6]

the sample data (likelihood). The ML estimation finds the most "likely" values of the distribution parameters for a set of data. The likelihood derivation for the Markov models under study is based on assumptions regarding the densities of the available data (Supplementary Materials).

NLS regression extends the linear least squares regression in order to be used for a much larger and more general class of functions. Almost any function that can be written in closed form can be incorporated in a nonlinear regression model. The idea of using NLS technique to estimate the Markov model parameters is to calculate the observed and expected transitions, as well as to treat the equation (observed = expected + error) as a non-linear regression equation, with regression coefficients for screening Se and transition parameters to be estimated [4].

Both the NLS and ML procedures generally perform quite well, but sometimes a non-identifiability problem can occur, i.e. the impossibility of the model parameters to be uniquely determined from the distribution of the observed random variables. The latter may occur if the model does not have sufficient information to estimate all the parameters of interest. This problem may depend on the specific data used to inform the model. The non-identifiability problem may be overcome by adding external information, e.g. data from non-participants [6, 7], by further simplifying the model structure in order to reduce the number of parameters to be estimated, or by incorporating prior information in a Bayesian framework [8, 9].

The data used in our application seem to lead to a non-identifiability problem, since both the ML and NLS estimation procedures gave estimates that changed with the parameters' initial values or difficultly converged. In order to take this problem into account, this problem we used the BA by incorporating prior information on all the parameters.

In a Bayesian framework, extra information may be added to the analysis by combining the data - expressed through the likelihood function - with prior information on unknown parameters - expressed through suitable distributions - and thus deriving posterior distributions using the Bayes theorem. The posterior distribution contains updated beliefs about the values of the model parameters, after taking into account the information provided by the data.

Suitable prior distributions for the unknown parameters must be specified. In this analysis we used an uninformative flat prior Gamma(0.001, 0.001) for the transition rates. The Gamma distribution is a distribution that is continuous, flexible, and constrained between [0,+∞), and is often used to model transition rates [3, 10, 11]. The Beta distribution is a natural choice to represent the uncertainty on a probability being continuous and constrained in the interval (0,1) [12]. Since estimates of Italian breast cancer screening Se are available from several studies [13-15], an informative Beta distribution was assumed as prior distribution for the test Se. The Beta parameters were derived with the method of moments [16] from the point estimates and standard errors from previous published studies [13-15]. The selected parameters were 4.11 and 1.37 which corresponded to an average sensitivity around 75%.

The posterior distribution given the prior distributions, and data likelihood was computed by Gibbs sampling [8], providing random samples from marginal posterior densities for each parameter of interest. For the analyses presented, posterior inferences were based on 100 000 iterations of the Gibbs sampler after a burn-in of 50 000 iterations was discarded. Convergence was assessed by running multiple chains from dispersed starting values [17].

We carried out a sensitivity analysis on the prior distributions in order to evaluate the robustness of the results. In particular, both shape and inverse scale parameters of the Gamma distribution were changed from 0.1 to 0.0001. In the same way the shape parameters of the Beta distribution were changed from 1 to 5; the mean of the Se varied between 0.50 and 0.90.

We calculated the determination coefficient for the modeled outcomes from each model as a measure of goodness of fit comparing the observed number of cases to the estimated numbers [18].

The results for the ML and NLS presented above were used as initial values for the estimates from the BA.

We used the freeware program Just another Gibbs sampler (JAGS) [19] for the Bayesian modeling, and the R software for the ML and the NLS estimation procedures [20].

Data

The observed data that were used to inform the three Markov models consist of

detection rates from different screening rounds and interval cancer rates. We used data from the cohort of 50-69 years-old women who participated in the first two screening rounds of the Florentine screening programme in the period from 1991-1993, and who were followed up for breast cancer up to two years after the second round [21, 22].

The screening histories of all women from of the cohort were extracted from the local computerized screening database. Screening exposure was defined on the basis of attendance at the first and second rounds.

To identify the types of breast cancer that occurred within the cohort, a linkage with the Tuscan Cancer Registry [23] was made and the cases were classified according to:

- screen detected cancers that occurred during the first round (prevalent cases)
- screen detected cancers that occurred during the second round (incident cases)
- interval cancers that occurred between the first and second round
- interval cancers that occurred during the first year after the first round
- interval cancers that occurred during the second year after the first round
- interval cancers that occurred between the second round and the two subsequent years.

Based on the TNM classification, all cases were stratified according to the lymph-node status (PN-, PN+).

Table 1 presents the data from the Florentine screening programme. Twenty-eight thousand three hundred and ninety women participated in the first round of screening with a detection rate of 7.2‰, and 23 445 women participated in the second round with the detection rate of 3.6‰; 32 interval cancers were detected between the first and the second round, and 31 during the two years after the second round.

The detection rate was higher in the age-class 60-69 than in those aged 50-59 both during the first and the second round. The occurrence of interval cancers was quite stable in the two intervals considered and similar across age-classes.

RESULTS

We performed the analyses for the three models using three different estimation procedures on the Florentine dataset stratified for the two age classes (50-59 and 60-69). The results from the two, three, and five-state models are reported in Tables 2, 3, and 4, respectively.

Table 2 presents the results from the two-state model. The model provides a very good fit with determination coefficients around 0.99 for both age classes.

The transition rate from asymptomatic to symptomatic cancer had a very similar result with the three estimation procedures. On the contrary, estimates of screening sensitivity showed meaningful differences among the estimation procedures, especially for the parameters estimated for the 50-59 age class with very high values (around 1) for the optimization procedures. In all of the estimation procedures, MST values were higher for older women.

Table 3 presents the results from the three-state model. This model also fits very well for both age classes with determination coefficients around 0.99. Although the transition rate estimates from no detectable to asymptomatic cancer had very similar results with the three estimation procedures and the two age classes, the transition rates from asymptomatic to symptomatic cancer resulted lower in the Bayesian estimation procedure with respect to the optimization ones. The resulting MST values were therefore 1 and 1.5 years higher in the Bayesian estimation procedure respectively for 50-59 and 60-69 age classes.

Results from the five-state model are shown in Table 4. Also in this case the model had a very good fit with determination coefficients around 0.97 and 0.98 for the 50-59 and 60-69 age classes, respectively. The transition rate from no detectable to asymptomatic PN- cancer had a very similar result in all the estimation procedures and in the two age classes. All the estimates (transition rates and Se) had a higher result using the ML procedure than did the BA and NLS for both age-classes.

The estimates of Se resulted around 100% for both age-classes when the ML procedure was used, while BA and NLS Se estimates were respectively 75% and 63% for ages 50-59, and 77% and 87% for ages 60-69. In all the procedures, sensitivity seems to increase with age.

The MST estimates from the three procedures

TABLE 1

DATA FROM THE COHORT OF 50-69 YEAR-OLD WOMEN WHO PARTICIPATED IN THE FLORENTINE SCREENING PROGRAMME FROM THE PERIOD 1991-1993, STRATIFIED ACCORDING TO TWO AGE CLASSES (50-59 AND 60-69)

	AGE 50-59	AGE 60-69
ROUND 1		
Screened women	14 608	13 782
PN+ cancers at screening	21	21
PN- cancers at screening	64	99
Total cancers	85	120
ROUND 2		
Screened women	12 305	11 140
PN+ cancers at screening	11	12
PN- cancers at screening	26	36
Total cancers	37	48
INTERVAL CANCERS		
Women 1 year after round 1*	14 523	13 662
Cancers 1 year after round 1*	3	4
Women 2 years after round 1*	14 520	13 658
Cancers 2 years after round 1*	13	12
PN+ cancers between rounds 1 and 2	5	6
PN- cancers between rounds 1 and 2	11	10
PN+ cancers 2 years after round 2	6	8
PN- cancers 2 years after round 2	10	7

*Interval cancers are needed yearly to fit the two-state model

TABLE 2

RESULTS FOR THE TWO-STATE MODEL WITH THE THREE ESTIMATION PROCEDURES

ESTIMATION PROCEDURE	AGE	λ_{12}	95%CI	SE	95%CI	MST
BA	50-59	0.31	0.22, 0.39	0.91	0.74, 0.99	3.27
BA	60-69	0.20	0.14, 0.26	0.87	0.68, 0.99	4.96
ML	50-59	0.34	0.27, 0.41	0.99	-	2.94
ML	60-69	0.21	0.15, 0.29	0.91	0.46, 0.99	4.76
NLS	50-59	0.34	0.31, 0.34	1.00	0.91, 1.00	2.94
NLS	60-69	0.20	0.17, 0.23	0.88	0.77, 0.98	5.00

λ_{12}: transition rate from asymptomatic cancer to symptomatic cancer; Se: sensitivity of the screening test; 95%CI: 95% confidence interval (credible in the Bayesian estimation procedure); MST: mean sojourn time; BA: Bayesian approach; ML: maximum likelihood; NLS: non-linear least squared

had a rather similar result in the older age-class with values around 4 years for BA and NLS, and 3 years for ML. Estimates for the 50-59 age-class instead had a different result among procedures, with values around 4, 2, and 7 years for BA, ML, and NLS procedures, respectively.

DISCUSSION

We applied three multi-state Markov models to data from breast cancer screening in Florence. The case-study illustrates the use of different estimation procedures with

TABLE 3

ESTIMATION PROCEDURE	AGE	λ_{01}	95%CI	λ_{12}	95%CI	SE	95%CI	MST
BA	50-59	1.81e-3	1.12e-3, 2.36e-3	0.26	0.10, 0.39	0.81	0.47, 0.99	3.83
	60-69	2.16e-3	1.22e-3, 2.91e-3	0.20	0.07, 0.32	0.80	0.48, 0.99	4.92
ML	50-59	2.09e-3	1.73e-3, 2.53e-3	0.36	0.28, 0.45	1.00	0.99, 1.00	2.78
	60-69	2.71e-3	2.22e-3, 3.17e-3	0.30	0.24, 0.37	1.00	0.99, 1.00	3.34
NLS	50-59	2.09e-3	1.87e-3, 2.19e-3	0.36	0.26, 0.37	1.00	0.83, 1.00	2.78
	60-69	2.71e-3	2.18e-3, 287e-3	0.31	0.19, 0.33	1.00	0.75, 1.00	3.23

Title: RESULTS FOR THE THREE-STATE MODEL WITH THE THREE ESTIMATION PROCEDURES

λ_{01}: transition rate from no detectable cancer to asymptomatic cancer; λ_{12}: transition rate from asymptomatic cancer to symptomatic cancer; Se: sensitivity of the screening test; 95%CI: 95% confidence interval (credible in the Bayesian estimation procedure); MST: mean sojourn time; BA: Bayesian approach; ML: maximum likelihood; NLS: non-linear least squared

TABLE 4

ESTIMATION PROCEDURE	AGE	λ_{01}	95%CI	λ_{12}	95%CI	λ_{13}	95%CI	λ_{24}	95%CI	SE	95%CI	MST
BA	50-59	1.81E-03	0.95e-3, 2.45e-3	0.19	0.03, 0.36	0.17	0.05, 0.28	0.49	0.08, 0.99	0.75	0.37, 0.98	3.88
	60-69	2.29E-03	1.08e-3, 3.21e-3	0.17	0.03, 0.31	0.11	0.03, 0.19	0.67	0.11, 1.32	0.77	0.39, 0.99	4.50
ML	50-59	2.26E-03	1.89e-3, 2.70e-3	0.33	0.23, 0.45	0.24	0.16, 0.35	0.84	0.55, 1.20	1.00	-	2.47
	60-69	2.91E-03	2.46e-3, 3.44e-3	0.28	0.21, 0.37	0.15	0.10, 0.23	1.09	0.72, 1.52	1.00	0.99, 1.00	2.92
NLS	50-59	1.33E-03	0.70e-3, 2.01e-3	0.07	0.01, 0.26	0.12	0.05, 0.24	0.21	0.02, 0.68	0.63	0.43, 1.00	6.90
	60-69	2.28E-03	1.24e-3, 2.79e-3	0.16	0.04, 0.24	0.12	0.05, 0.19	0.73	0.18, 1.00	0.87	0.60, 1.00	4.37

Title: RESULTS FOR THE FIVE-STATE MODEL WITH THE THREE ESTIMATION PROCEDURES

λ_{01}: transition rate from no detectable cancer to asymptomatic cancer PN-; λ_{12}: transition rate from preclinical cancer PN- to preclinical cancer PN+; λ_{13}: transition rate from preclinical to clinical cancer PN-; λ_{24}: transition rate from preclinical to clinical cancer PN+; Se: sensitivity of the screening test; 95%CI: 95% confidence interval (credible in the Bayesian estimation procedure); MST: mean sojourn time; BA: Bayesian approach; ML: maximum likelihood; NLS: non-linear least squared

three multi-state Markov models for the main purpose of estimating screening Se and MST.

These estimates may further be used to make predictions for future trends of the burden of breast cancer and of the associated mortality after the introduction of screening.

The MST varied between 2 and 7 years and between 3 and 5 years for women aged 50-59 and 60-69, respectively. The two-state model gave the same MST estimates across the three estimation procedures for both age-classes. Alternatively, increasing the model complexity a higher variability in estimates was observed among the estimation procedures.

The screening Se varied between 63% and 100% and between 77% and 100% for

women aged 50-59 and 60-69, respectively. The screening sensitivity estimates had a highly variable result, both among estimation techniques and models.

The three Markov models in combination with the three estimation procedures had very good fits. The latter was found in similar applications of Markov models [6, 24].

All the models produced estimates for MST that were globally consistent with the literature [25, 26], with a slight underestimation of MST for the older age-class in the two-state model, and an overestimation of the MST in the young age-class for the NLS approach (Tables 2, 3, 4). The MST estimates showed an increase with ageing - a point that has already being raised in the

literature. Numerous studies based on Markov models for the progression of disease showed that MST becomes longer with ageing, ranging from 2 to 3.7 in the 50-59 age group and from 3.3 to 4.2 in those aged 60-69 [2, 6, 7, 26-28].

The Se estimates from ML and NLS had a result in accord with the literature that reports estimates varying from 0.82 to 1.00 in 50-59 and from 0.88 to 1.00 in those aged 60-69 [2, 6, 7, 26-28]. NLS underestimates Se in the younger age-class. Estimates of the Se based on data from observed interval cancers in a screening programme showed a slightly lower value with respect to the model-based approach - from 70% to 82% [13-15]. The BA estimates seem more plausible, perhaps because of the use of prior information on such parameters (Tables 2, 3, 4).

Increasing the model's complexity has the advantage of allowing us to describe the natural history of the disease of cancer with increasing detail, i.e. the five-state model has the advantage of estimating cancer cases by node status. However, as highlighted by Uhry et al. [6], it should be considered that the validity of some assumptions in the five-state model should be closely inspected. In particular, the MST in the two preclinical states (PN- and PN+) was assumed to be independent, whereas they might be positively correlated. The use of estimates derived from two separate dataset (screened and unscreened) has been proposed [6] as a way to avoid bias in prediction of the node status when predicting screening scenarios.

One of the several drawbacks of this kind of Markov model is that it aims to jointly estimate Se and MST that are highly correlated [6]. In our application the BA overcomes this problem by strengthening the knowledge on screening Se with the integration of prior belief. By increasing the model complexity, the BA showed an Se estimate that comes closer to the data represented in the literature.

In addition, the BA gave robust estimates with all the models. Indeed, for all the models, sensitivity analyses confirmed the stability of posterior estimates over different values of the prior hyper-parameters.

In conclusion, we estimated both natural history parameters and Se using data from the two-yearly mammographic screening programme implemented in Florence, Italy. Results were in accordance with the literature, and the BA was identified as being more appropriate in this setting, given the non-identifiably problem arising from the ML and NLS techniques, and given the more reliable parameter estimates.

References

[1] Day NE, Walter SD. Simplified models of screening for chronic disease: estimation procedures from mass screening programmes. Biometrics 1984; 40(1): 1-14

[2] Paci E, Duffy SW. Modelling the analysis of breast cancer screening programmes: sensitivity, lead time and predictive value in the Florence District Programme (1975-1986). Int J Epidemiol. 1991; 20(4): 852-8

[3] Prevost TC, Launoy G, Duffy SW, Chen HH. Estimating sensitivity and sojourn time in screening for colorectal cancer: a comparison of statistical approaches. Am J Epidemiol. 1998; 148(6): 609-19

[4] Duffy SW, Chen HH, Tabar L, Day NE. Estimation of mean sojourn time in breast cancer screening using a Markov chain model of both entry to and exit from the preclinical detectable phase. Stat Med. 1995; 14(14): 1531-43

[5] Chen HH, Duffy SW, Tabar L, Day NE. Markov chain models for progression of breast cancer. Part I: tumour attributes and preclinical screen-detectable phase. J Epidemiol Biostat 1997; 2(1): 9-23

[6] Uhry Z, Hédelin G, Colonna M, et al. Multi-state Markov models in cancer screening evaluation: a

brief review and case study. Stat Methods Med Res 2010; 19(5): 463-86

[7] Wu JC, Hakama M, Anttila A, et al. Estimation of natural history parameters of breast cancer based on non-randomized organized screening data: subsidiary analysis of effects of inter-screening interval, sensitivity, and attendance rate on reduction of advanced cancer. Breast Cancer Res Treat 2010; 122(2): 553-66.

[8] Welton NJ, Ades AE. Estimation of markov chain transition probabilities and rates from fully and partially observed data: uncertainty propagation, evidence synthesis, and model calibration. Med Decis Making 2005; 25: 633-45

[9] Myles JP, Nixon RM, Duffy SW, et al. Bayesian evaluation of breast cancer screening using data from two studies. Stat Med 2003; 22: 1661-74

[10] Carreras G, Baccini M, Accetta G, Biggeri A. Bayesian probabilistic sensitivity analysis of Markov models for natural history of a disease: an application for cervical cancer. IJPH 2012; 9: 3-12

[11] Zorzi M, Guzzinati S, Puliti D, Paci E, IMPACT working group. A simple method to estimate the episode and programme sensitivity of breast cancer screening programmes. J Med Screen 2010; 17(3): 132-8

[12] Caumo F, Vecchiato F, Pellegrini M, et al. Analysis of interval cancers observed in an Italian mammography screening programme (2000-2006). Radiol Med 2009; 114(6): 907-14

[13] Elena PM, Nehmat H, Ermes M, et al. Quality of mammography screening in the Milan programme: evidence of improved sensitivity based on interval cancer proportional incidence and radiological review. Breast 2009; 18(3): 208-10

[14] Bain LJ and Engelhardt M. Introduction to Probability and Mathematical Statistics, Duxbury Press (1991)

[15] Gelman A. Bayesian data analysis, second edition. 1995

[16] Gelman A and Rubin DB. Inference from iterative simulation using multiple sequences, Statistical Science 1992; 7: 457-511

[17] David A. Freedman. Statistical Models: Theory and Practice, Cambridge University Press (2005)

[18] Plummer M. JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling, Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003), March 20-22, Vienna, Austria. ISSN 1609-395X

[19] R Development Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria 2009. http://www.R-project.org Accessed: May 2013

[20] Giorgi D, Ambrogetti D, Bianchi S, et al. Design and preliminary results of the Florence Breast Screening Programme (Progetto Firenze Donna). Eur J Cancer Prev 1994; 3 (Suppl 1): 29: 34

[21] Ministero della Salute. Osservatorio Nazionale Screening. http://www.osservatorionazionalesc4reening.it (last access May 2013)

[22] Curado MP, Edwards B, Shin HR, et al. Cancer Incidence in five Continents. IARC Scientific Pubblication No 160. Vol 9. Lyon, France 2009

[23] Hui-Min W, Ming-Fang Y, Chen TH. SAS macro program for non-homogeneous Markov process in modeling multi-state disease progression. Comput Methods Programs Biomed 2004; 75: 95-105

[24] Tabar L, Fagerberg G, Chen HH, et al. Efficacy of breast cancer screening by age. New results from the Swedish Two-County Trial. Cancer 1995; 75(10): 2507-17

[25] Duffy SW, Day NE, Tabár L, et al. Markov models of breast tumor progression: some age-specific results. J Natl Cancer Inst Monogr 1997; 22: 93-7

[26] Chen HH, Duffy S. A Markov chain to estimate the tumour progression rate from preclinical to clinical phase, sensitivity and positive predictive value for mammography in breast cancer screening. The Statistician 1996; 45(3): 307-17

[27] Chen TH, Kuo HS, Yen MF, et al. Estimation of sojourn time in chronic disease screening without data on interval cases. Biometrics 2000; 56(1): 167-72

[28] Omlin M, Reichert P. A comparison of techniques for the estimation of model prediction uncertainty. Ecological Modelling 1999; 115: 45-59

[29] Cox DR, Miller HD. The theory of stocastic processes. London: Methuen, 1965

*