Lead time bias correction in breast cancer screening studies

Marius Robert [1] & Brice Amadeo [2] & Marie Poiseuil [2] & Maja Pohar-Perme [3] & Virginie Rondeau [1]

[1] *Biostatistics team, Inserm U1219, University of Bordeaux, France.*
[2] *EPICENE team, Inserm U1219, University of Bordeaux, France.*
[3] *IBMI, Faculty of Medicine, University of Ljubljana, Slovenia.*

marius.robert@u-bordeaux.fr

Résumé. Le dépistage par mammographie joue un rôle crucial dans la détection et le diagnostic du cancer du sein, permettant des traitements précoces, améliorant ainsi les taux de survie. L'évaluation de l'efficacité du dépistage implique de comparer la survie des patients diagnostiqués par dépistage avec ceux diagnostiqués après l'apparition des symptômes. Cependant, cette analyse est sujette à des biais pouvant surestimer les avantages du dépistage. Le biais d'avance au diagnostic se manifeste car le dépistage permet un diagnostic précoce du cancer du sein, entraînant une survie observée plus longue sans réelle amélioration de la date de décès. Deux approches fréquemment utilisées pour remédier à ce biais ont été appliquées à des données réelles pour en faire une comparaison. La première approche repose sur le temps de séjour en phase préclinique, c'est-à-dire la période pendant laquelle la tumeur est détectable par dépistage mais ne provoque pas de symptômes. Une distribution de cette durée peut être obtenue à partir d'un modèle multi-états. La correction proposée par Duffy et al. (2008), soustrait l'espérance de cette distribution à la survie observée. Nous proposons d'améliorer la correction déjà existante en tenant compte de la densité mammaire. L'hypothèse sous-jacente est que des tissus mammaires plus denses réduisent la sensibilité du dépistage, raccourcissant ainsi la durée de la phase préclinique. La deuxième approche, développée par Abrahamsson et al. (2020), est basée sur un modèle de croissance tumorale continue. Ce modèle permet d'estimer la croissance volumique jusqu'à l'âge du patient au diagnostic. L'esprit de cette seconde approche est d'utiliser les estimations d'un modèle pour prolonger la croissance tumorale chez les patients diagnostiqués par dépistage pour estimer le temps jusqu'à un diagnostic symptomatique. Bien que cette dernière méthode soit plus précise dans l'estimation du temps d'avance, elle demeure peu utilisée en raison de sa nouveauté, de sa complexité et du manque de logiciel à disposition. Une standardisation de son utilisation serait pertinente, surtout compte tenu des critiques émises sur la correction de Duffy, jugée excessive.

Mots-clés. Cancer du sein, dépistage, biais d'avance, survie, correction.

Abstract. Mammography screening plays a crucial role in detecting and diagnosing breast cancer, enabling early treatments and thereby improving survival rates. Evaluating the effectiveness of screening involves comparing the survival of patients diagnosed through screening with those diagnosed after the onset of symptoms. However, this analysis is susceptible to biases that may overestimate the benefits of screening. The lead-time bias arises

because screening allows for the early diagnosis of breast cancer, resulting in a longer observed survival without a true improvement in the date of death. Two commonly used approaches to solve this bias were applied to real data in order to compare them. The first approach is based on the sojourn time in the preclinical phase, i.e., the period during which the tumor is detectable through screening without causing symptoms. A distribution of this duration can be derived from a multi-state model. The correction proposed by Duffy et al. (2008) subtracts the expected value of this distribution from the observed survival. We propose enhancing the existing correction by considering breast density. The underlying hypothesis is that denser breast tissues reduce screening sensitivity, thereby shortening the duration of the preclinical phase. The second approach, developed by Abrahamsson et al. (2020), is based on a continuous tumor growth model. This model allows volume growth to be estimated up to the patient's age at diagnosis. The spirit of this second approach is to use estimates from a model to prolong tumor growth in patients diagnosed by screening to estimate the time until a symptomatic diagnosis. Although this latter method is more precise in estimating lead time, it remains underutilized due to its novelty, complexity and lack of available software. Standardizing its use would be relevant, especially considering criticisms of the Duffy correction, deemed excessive.

Keywords. Breast cancer, screening, lead time, survival, bias correction.

1 Introduction

1.1 Breast cancer screening

Screening is a medical practice that involves searching for disease in an asymptomatic individual. The main objective of screening is to detect the disease at an early stage, before the onset of symptoms, which can significantly improve treatment outcomes and therefore patients' survival. Organized screening programs implemented in developed countries provide for mammograms at regular intervals for women aged over 50 years, generally every 2 years. In screening programs, cancer detection can be achieved through two distinct methods: positive screening or symptomatic identification. It is assumed that the implementation of a nationwide organized screening program, has played a significant role in improving patient survival rates for the last fifteen years. The main measure used to assess the effectiveness of breast cancer screening is the disease-specific survival rate after diagnosis, calculated using the Kaplan-Meier method. In order to estimate the survival improvement of patients detected following a screening, it is necessary to compare their survival time from diagnosis with those of patients detected following symptoms. However, screening introduces bias leading to an overestimation of survival of patients detected through screening.

1.2 Lead time bias

The lead-time bias is a well-known issue in breast cancer screening. Survival times from the diagnosis of patients following screening versus those following symptoms are not comparable. The reason is that the screening advanced the diagnosis in time. The lead time, that cannot be observed, is the duration between the date of diagnosis by screening and the hypothetical date of diagnosis by symptoms for screen-detected patients. The screening introduces an artificial extension of the survival for these patients, which is not observed in patients diagnosed by symptoms (see figure 1). This bias on the survival of patients diagnosed by screening has consequences. The longer survival observed in patients diagnosed by screening is not entirely due to the early detection, but also to the lead-time added by the screening process. Therefore, the actual benefit of screening in terms of improving patient survival is overestimated.

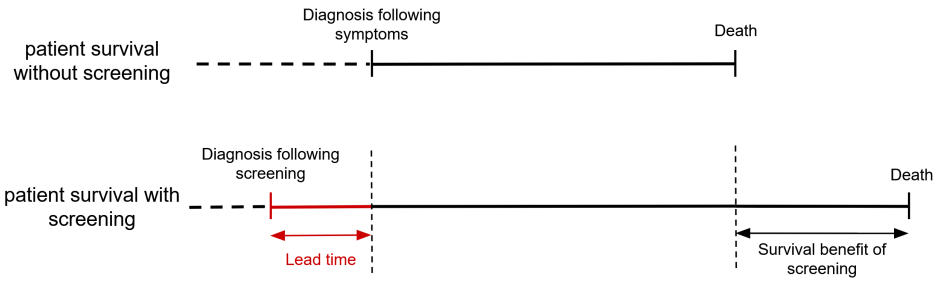

Figure 1: Illustration of the survival time comparison since diagnosis for a patient with or without screening.

2 Objective

The objective is to use two lead time correction approaches on real data and compare them. Ultimately, this project will result in the development of an R package, facilitating the application of these methods by epidemiologists.

3 Method

Two frequently used methods for correcting lead time have been compared: the first one is based on estimates from a multi-state Markov model, and the second one is based on a continuous tumor growth model. These two methods are applied in three steps. First, a model is estimated for the overall study population over the screening period, from their first mammography to their diagnosis. Then, the parameters of these models are used to determine a distribution of the additional time added due to the lead time bias. A time is then subtracted from the survival since diagnosis of each screen-detected patient.

3.1 Multi-state approach

3.1.1 Homogeneous Markov model

Duffy et al. (1995) proposed to model the tumor's progression with a three-states Markov model (white boxes in figure 2), each state corresponding to a phase in the natural history of the disease: no detectable tumor (state 1), the preclinical phase (state 2) when a tumor is detectable by screening but not symptomatic, the clinical phase (state 3) when the disease becomes symptomatic.

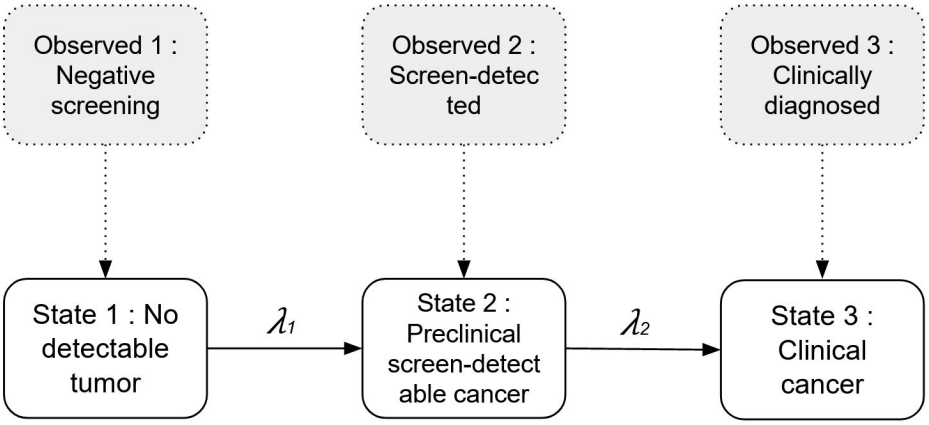

Figure 2: Three-state Markov model of the breast cancer natural history: the gray boxes represent the observed states, and the white boxes their underlying state in the model.

Negative mammography corresponds to an observation in state 1, where there is either no tumor or the tumor is not detectable. A positive screening mammography corresponds to an observation in state 2. The patient has no symptoms but the tumor is detectable. Diagnosis based on symptoms corresponds to an observation in state 3. The transition from state 1 to state 2 is interval-censored, meaning that the observations are in discrete time and do not provide information on the exact time of transition from one state to another. It is assumed that observations in state 3 occur at the exact time of transition into this state, which corresponds to diagnosis at the onset of symptoms and the clinical cancer phase in the natural history of the disease. Based on the definitions of the different states, no patient can be successively observed in states 2 and 3.

3.1.2 Breast density adjustment:

One of the main criticisms of the initial model (Duffy et al., 1998) is that it doesn't take into account patients' differences in cancer progression speed and therefore adjustments on sojourn time estimation. The underlying assumption with our adjustments is that model parameters and lead time correction are different depending on the four categories of breast density using the BI-RADS classification (Boyd et al., 2007; Choi et al., 2022). Transition

rates are estimated using the following formula with three parameters added per transition for the breast density indicator variables:

$$\lambda_j = exp(\beta_{0_j} + \beta_{1_j} br2 + \beta_{2_j} br3 + \beta_{3_j} br4)$$

for $j = 1, 2$, the transition.

3.1.3 Lead time bias correction

The correction for lead time based on the estimates of this model was published in 2008 by Duffy et al. The principle of this correction is to subtract from patients' survival time, the part of the sojourn time in preclinical phase in their post-diagnosis follow-up s. The sojourn time must follow an exponential distribution so the transition from preclinical state to clinical state has to be homogeneous in time. In the correction of Duffy et al., we define t, the time elapsed since the screening diagnosis, then there are two possibilities:
The patient passed away at the time of the latest news after diagnosis t:

$$E(s) = P(s \leq t)E(s|s \leq t) = \frac{1 - e^{-\lambda_2 t} - \lambda_2 t e^{-\lambda_2 t}}{\lambda_2(1 - e^{-\lambda_2 t})} \tag{1}$$

The patient is alive at the time of the latest news after diagnosis t:

$$E(s) = P(s \leq t)E(s|s \leq t) + P(s > t)t = \frac{1 - e^{-\lambda_2 t}}{\lambda_2} \tag{2}$$

Finally, the correction is applied to the survival of each screen-detected patient i by subtracting the average amount of sojourn time in post-diagnosis follow-up $E(s)$.

$$T_{corrected,i} = T_{observed,i} - E(s)_i \tag{3}$$

3.2 Tumor growth approach

The second correction method, which is more recent and complex, involves estimating a continuous tumor growth model. The continuous tumor growth model allows us to estimate the volumetric growth of tumors over time until a detection date through symptoms, with the assumption of exponential growth. It will then be possible to estimate the necessary time for growth from the size at diagnosis after the screening to the hypothetical detection with symptoms.

3.2.1 Continuous tumor growth model

These models are based on a tumor growth function and are estimated from latent sub-models (Isheden et al., 2017; Strandberg et al., 2019):

$$V(x) = v_0 e^{\frac{x}{r}}$$

where v_0 is the tumor volume at onset. We assume tumors are spherical, and since we define onset as the point where the tumor diameter is 0.5mm, then $v_0 \approx 0.06mm^3$. This function allows expressing the growth rate $r > 0$ as a function of the volume at detection v, the age at detection x that are observed.

In the model the individual likelihood contributions are based first on the probability of detecting a tumor of volume v. We use an exponential density function of the tumor volume at symptomatic detection ($V_{sym}|R = r \sim Exp(\eta r)$) marginalized on the gamma density function of the tumor growth rate ($R \sim Gamma(\tau_1, \tau_2)$).

The individuals contributions also take into account the probability of detecting the tumor at the screening detection and the probability of not detecting the tumor at the previous screening knowing the inverse growth rate r. For these probabilities the screening sensitivity is dependent on the tumor diameter and the BI-RADS category, and is assumed to be of logistic form.

$$S(d) = \frac{exp(\beta_0 + \beta_1 d + \beta_2 br2 + \beta_3 br3 + \beta_4 br4)}{1 + exp(\beta_0 + \beta_1 d + \beta_2 br2 + \beta_3 br3 + \beta_4 br4)}, \qquad d \geq 0$$

where d is the diameter of the tumor, which is observed at diagnosis, and which is calculated during previous screening using the tumor growth function.

3.2.2 Lead time bias correction

The model and its parameters will then be used to estimate a density function of the lead time conditionally on the tumor volume, the growth rate, and the screening sensitivity (Abrahamsson et al., 2020). The general idea is to create a customized density function for each patient tracking her follow-up from diagnosis to the date of the latest news and evaluating at each time point the probability of symptomatic detection. We denote L, a random variable representing a quantity of lead time, and l a realization of this variable. It is assumed that $L = 0$ for patients detected symptomatically, and $L > 0$ for screen-detected. The lead time is the time between the screen-detection time point t_{scr} and the hypothetical symptomatic detection time point T_{sym} ($L = T_{sym} - t_{scr}$), expressed in the model using the tumor growth function. Mathematically, we express the lead time as follows:

$$V_{sym} = v_{scr} e^{(T_{sym} - t_{scr})/R} \iff L = Rlog(\frac{V_{sym}}{v_{scr}})$$

The density function of lead time is in the following form:

$$f_{L|V=v_{scr},H}(l) \propto \int_0^\infty f_{R|V=v_{scr}}(r) \cdot f_{L|V=v_{scr},R=r}(l) \cdot P(H|V = v_{scr}, R = r)dr \qquad (4)$$

- The first part of this integral is for the probability of having an inverse growth rate r given the tumor volume at detection $V = v_{scr}$, calculated by using the conditional density function of the inverse growth rate.

- The second part is the expression of lead time in the model, given the tumor volume at detection v_{scr} and the inverse growth rate r. We derive the cumulative density function of the volume at symptomatic detection with the previously mentioned relation $V_{sym} = v_{scr}e^{L/R}$.

- The third part, depending on the screening sensitivity, is the probability of negative screening results in the patient's mammography history H. For each mammography, the tumor diameter is backward calculated as before by using the inverse growth rate r.

Similarly to the correction made by Duffy et al. (2008), the time subtracted from survival should not exceed the observed survival time for a dead patient. The lead time distribution is then in a conditional truncated form (Abrahamsson et al., 2020). The final value $\hat{l}_i$ is estimated for each woman by taking the expectancy of her lead time density function. The correction on survival is performed for each screen-detected patient by subtracting the lead time estimate from the observed survival.

$$T_{corrected,i} = T_{observed,i} - \hat{l}_i \tag{5}$$

4 Results

The data used for the analysis come from the Gironde general cancer registry. Some patients were excluded from the analyses due to unknown detection method, missing data on tumor size or the breast density at diagnosis, or the presence of an in situ tumor, which did not verify the growth hypothesis of the model.

Figure 3 shows a comparison of patient survival by diagnostic method. The probability of survival at 5 years for a woman diagnosed by symptom is 0.85. After a diagnosis by screening, the probability of survival at 5 years is 0.95. The corrected survival curves are above the curve for symptomatic patients, reflecting the fact that screening is effective on survival after correction of the lead time. The correction based on a multi-state model is the strongest of the two, and the survival curve tends to approach that of symptomatic patients. The correction based on tumor growth remains fairly close to that of survival without correction, the decreases in survival probabilities remain of the same order of magnitude.

5 Discussion

The correction based on a multi-state model is stronger, possibly even too strong, compared to the one based on tumor growth, which is more comprehensive in estimating lead time. Applying either of these corrections during survival analysis is of crucial importance for estimating the effectiveness of screening. It is necessary to keep in mind that both methods have their own assumptions, strengths, and limitations.

The definition of states is advantageous in order to estimate sojourn times and incidence rates. However, the main issue with the three-state model is that it only allows for estimating sojourn time and not lead time.

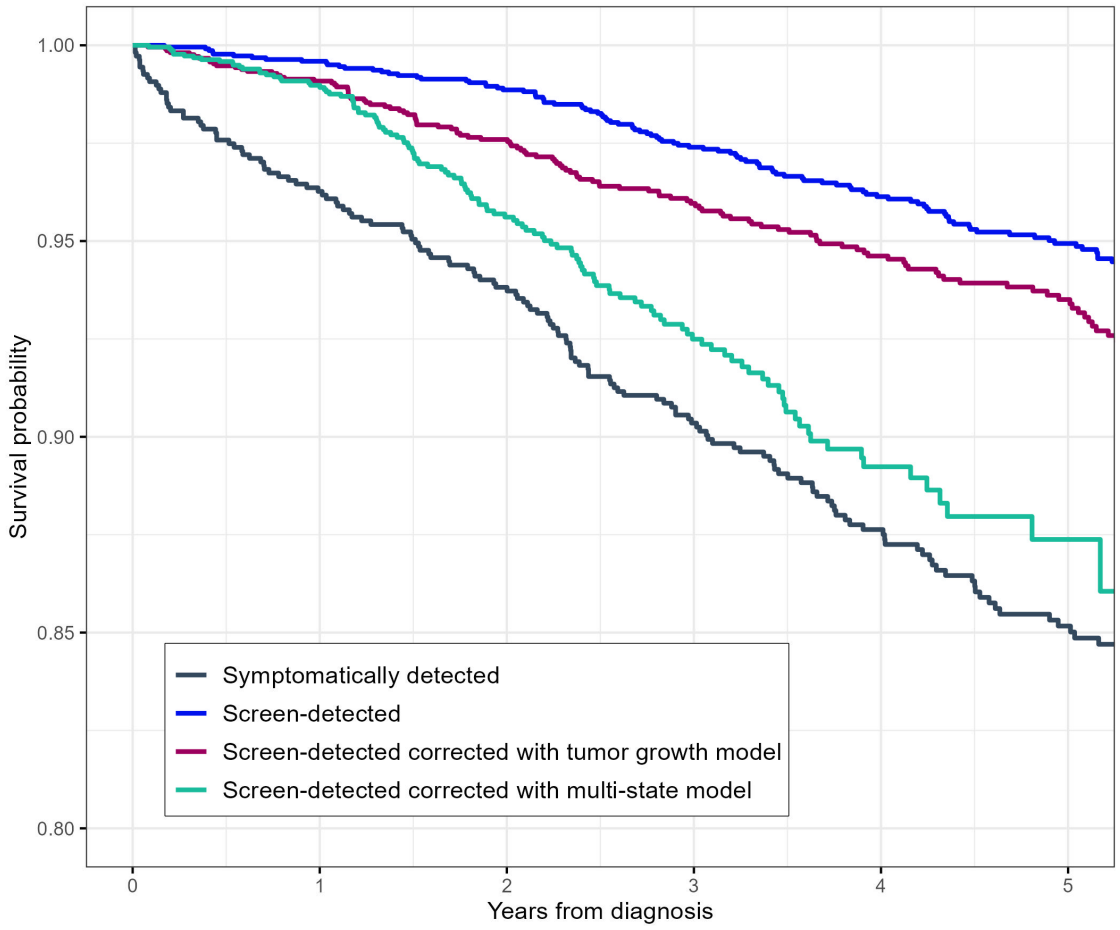

Figure 3: Survival comparison after breast cancer diagnosis for women detected by symptoms and screening according to different estimated lead time bias corrections. Symptomatic diagnosis: $n_{sym} = 1077$, screen-detected: $n_{scr} = 2212$.

The continuous tumor growth model and its associated correction directly address the limitations of the multi-state approach due to their assumptions. The distribution of lead time is personalized for each woman and adjusted on tumor size at diagnosis and screening sensitivity. However, some hypotheses constitute limitations. Particularly, the assumptions that tumors are spherical in shape and grow exponentially are not universally valid. Other shapes, such as ellipses, are possible, and tumor protrusions are common.

We assumed no modification in women's breast density over time. We were able to verify this through follow-ups of up to 10 years between the initial screenings in 2005 and the later diagnosis in 2015. However, the screening program is designed for 25 years, which is more than double the follow-up period. Therefore, breast density tends to decrease with age and, our assumption should be invalid.

6 Perspectives

Screening has a higher probability of detection for slow-growing tumors, which constitutes another bias. Extending a method towards selection biases linked to screening constitutes the continuation of this work.

References

Abrahamsson, L., Isheden, G., Czene, K., Humphreys, K. (2020). Continuous tumour growth models, lead time estimation and length bias in breast cancer screening studies. *Statistical Methods in Medical Research*, 29(2), 374-395.

Boyd, N. F., Guo, H., Martin, L. J., Sun, L., Stone, J., Fishell, E., ... Yaffe, M. J. (2007). Mammographic density and the risk and detection of breast cancer. *New England journal of medicine*, 356(3), 227-236.

Choi, E., Suh, M., Jung, S. Y., Jung, K. W., Park, S., Jun, J. K., Choi, K. S. (2022). Estimating Age-Specific Mean Sojourn Time of Breast Cancer and Sensitivity of Mammographic Screening by Breast Density among Korean Women. *Cancer Research and Treatment: Official Journal of Korean Cancer Association*, 55(1), 136-144.

Duffy, S. W., Chen, H. H., Tabar, L., Day, N. E. (1995). 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. *Statistics in medicine*, 14(14), 1531-1543.

Duffy, S. W., Nagtegaal, I. D., Wallis, M., Cafferty, F. H., Houssami, N., Warwick, J., ... Lawrence, G. (2008). Correcting for lead time and length bias in estimating the effect of screen detection on cancer survival. *American journal of epidemiology*, 168(1), 98-104.

Isheden, G., Humphreys, K. (2019). Modelling breast cancer tumour growth for a stable disease population. *Statistical Methods in Medical Research*, 28(3), 681-702.

Strandberg, J. R., Humphreys, K. (2019). Statistical models of tumour onset and growth for modern breast cancer screening cohorts. *Mathematical Biosciences*, 318, 108270.