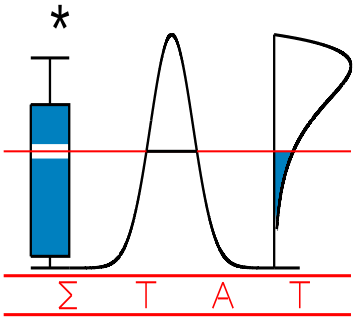
T E C H N I C A L

R E P O R T

0356

CLASSIFICATION OF LONGITUDINAL PROFILES USING NONLINEAR MIXED-EFFECTS MODELS

FIEUWS, S., VERBEKE, G. and L. J. BRANT



I A P S T A T I S T I C S

N E T W O R K

INTERUNIVERSITY ATTRACTION POLE

http://www.stat.ucl.ac.be/IAP

Classification of Longitudinal Profiles using Nonlinear Mixed-Effects Models

Abstract

In different biomedical situations markers are needed to detect the onset of a specific disease as soon as possible. Often, a series of marker measurements turns out to be a better screening tool than a single measurement. This fact urges the development of classification strategies using the longitudinal gathered information. Classical discriminant analysis can be applied whenever the data have a balanced structure. Extensions have been proposed for unbalanced data. These extensions use linear mixed effects models or linearized versions of nonlinear models to describe the longitudinal profiles in each group. Using these group-specific descriptions, posterior probabilities of group membership are calculated to classify the individual profiles. In this paper we present an extension of the proposed strategy by using linear as well as nonlinear mixed effects models for the description of group-specific evolutions. The extension will be illustrated using 342 PSA profiles collected in the Baltimore Longitudinal Study of Aging.

Key words: Classification, Disease Screening, Longitudinal profiles, Mixed Models

1 Introduction

Prostate cancer represents a very high proportion of the non-skin cancers reported annually (Carter and Coffey, 1990). Therefore, biomarkers are needed which can detect the disease in an early stage. A candidate marker which has been studied intensively, is the prostate specific antigen (PSA), since its level is related to the cancer volume. Ideally a single marker measurement should indicate the presence of a malignant cancer. However, using a single PSA measurement as a screening tool for the detection of prostate cancer is not possible since PSA is not only produced by cancerous cells but also by normal prostate cells. Therefore, an enlarged volume of the prostate tissue is not always due to the presence of a cancer, but

can also be due to a benign form of prostatic hyperplasia (BPH). According to Pearson and colleagues (1991), using a single PSA measurement would falsely identify up to 60% of the BPH patients as potential cancer cases.

As a better method for the detection of prostate cancer, clinical practice suggested the use of the rate of change in PSA. It has been shown by Pearson *et al.* (1994) that patients with prostate cancer differ from patients without prostate cancer in their evolution of PSA, suggesting that information about the PSA evolution can be used for the detection of prostate cancer in an early stage of the disease.

Since observational longitudinal data are usually unbalanced, random-effects models are a natural tool to describe the differences between the cancer and non-cancer patients. Pearson *et al.* (1994) used for all diagnostic groups a linear mixed-effects model with quadratic time effects for the logarithmic transformed PSA level. However, Morrell and colleagues (1995) suggested a nonlinear mixed-effects model as a more appropriate tool to describe the PSA evolution in the cancer patients. Their model is a piecewise model describing the transition from a slow non-cancerous growth of the prostate to an exponential phase of rapid growth due to presence of a growing tumour. This implies that a linear model is sufficient to adequately describe the PSA evolution of patients without cancer, but a nonlinear model is required to adequately describe the PSA evolution of patients with cancer. Consequently, a classification procedure is needed which can cope not only with the unbalanced structure of the data, but also with the diversity in group-specific models. In the next section we present a short overview of discriminant analysis methods for longitudinal profiles. An application of the proposed method is presented in Section 3, followed by a discussion of the obtained results in Section 4.

2 Discriminant Analysis for Longitudinal Profiles

Let $Y_i^{(j)}(t)$ denote the response for a subject i in group j, taken at time t. If the dataset with longitudinal measurements has a balanced structure, then $\boldsymbol{Y}_i^{(j)}$ denotes a random $n \times 1$ vector of complete measurements from the ith subject in the jth group, with $i = 1, 2, \ldots, N^{(j)}$ and $j = 1, 2, \ldots, G$. Assuming $\boldsymbol{Y}_i^{(j)} \sim N(X\boldsymbol{\beta}^{(j)}, V^{(j)})$, where X is the $n \times n$ design matrix with dummy variables for the time points, $\boldsymbol{\beta}^{(j)}$ the $n \times 1$ vector describing the average evolution in the jth group and $V^{(j)}$ its associated $n \times n$ unstructured covariance matrix, traditional

multivariate discriminant analysis can easily be applied. The resulting discriminant function is linear if $V^{(j)}$ is restricted to be equal for all G groups and quadratic if this restriction is abandoned. Ignoring the group structure yields the following mixture of multivariate normal models with the marginal distribution for the measurements $\boldsymbol{Y}_i$ given by:

$$\boldsymbol{Y}_i \sim \sum_{j=1}^{G} p^{(j)} N(X\boldsymbol{\beta}^{(j)}, V^{(j)}) \tag{2.1}$$

where the $p^{(j)}$'s are mixing proportions. The design matrix X_i can be augmented with m baseline covariates as age, race, and other important variables, such that its dimension becomes $n \times (n+m)$. For classification purposes the posterior probabilities for each subject to belong to each of the diagnostic groups can be calculated in the classical way. Let $p_i^{(j)}$ be the posterior probability for subject i to belong to group j, i.e.,

$$p_i^{(j)} = \frac{p^{(j)} \;\; f_i^{(j)}}{\sum_{j=1}^{G} p^{(j)} \;\; f_i^{(j)}} \tag{2.2}$$

with the $f_i^{(j)}$'s representing the densities of a multivariate normal distribution. Each patient is then classified in that group for which he has the highest posterior probability.

Albert (1983) already extended the classical concepts of discriminant analysis to multivariate response curves observed over fixed time intervals. Using interpolation or curve-fitting procedures, a time-varying distance measure between the individual curve and group-specific curves is used to allocate a subject to a group. This methodology requires that the response curves of the subjects in the training sample are fully observed over the considered time interval.

However, in longitudinal studies a balanced structure of observations is very exceptional such that the multivariate normal models with unstructured covariance matrices in mixture (2.1) cannot be used for the classification. Using the introduced notation, this implies that we have an $n_i \times 1$ vector of measurements for each subject taken at subject-specific points of time. Instead of using a multivariate normal model to describe the repeated measurements in

each group, mixed-effects models can then be used in the discriminant analysis. In general, a linear mixed-effects model for each group separately is any model which satisfies (Laird and Ware, 1982).

$$\boldsymbol{Y}_i = X_i\boldsymbol{\beta} + Z_i\boldsymbol{b}_i + \boldsymbol{\varepsilon}_i \tag{2.3}$$

where the p-dimensional vector $\boldsymbol{\beta}$ of so-called fixed effects describes the average evolution. Further, the q-dimensional vector $\boldsymbol{b}_i$ of so-called random effects describes how the profile for the ith subject deviates from the average profile, while the ε_i is the vector with the usual error components. The $\boldsymbol{b}_i$ are assumed to be normally distributed with mean vector $\mathbf{0}$ and covariance matrix D, independent of the error components in $\boldsymbol{\varepsilon}_i$ which are assumed to be independently distributed as $N(0, \sigma^2)$. It then follows that the vector $\boldsymbol{Y}_i$ of all measurements available for the ith subject is normally distributed with mean vector $X_i\boldsymbol{\beta}$ and a structured $n_i \times n_i$ covariance matrix $V_i = Z_i D Z_i' + \sigma^2 I_i$. X_i and Z_i are $(n_i \times p)$ and $(n_i \times q)$ matrices of known covariates. The number of columns of X_i depends on how the evolution over time is modelled. If for example a quadratic function is used, each row in X_i has the form $(1tt^2)$. Again other covariates can be incorporated. The resulting model is called a linear mixed model, and standard software (e.g., SAS procedure MIXED, 1992) can be used to estimate the parameters based on maximum likelihood or restricted maximum likelihood methods.

Using a linear mixed effects model to describe the evolutions in each group, the marginal distribution for $\boldsymbol{Y}_i$ presented in (2.1) is slightly modified as

$$\boldsymbol{Y}_i \sim \sum_{j=1}^{G} p^{(j)} N(X_i\boldsymbol{\beta}^{(j)}, V_i^{(j)}) \tag{2.4}$$

with X_i again the matrix containing the time effects (and eventually other covariates). The resulting classification rule based on (2.2) remains the same.

Tomasko and colleagues (1999) explored different possibilities for the structured covariance matrix in (2.4) as compared with the unstructured situation. Their modification concerns the linear discriminant function since the structured $V_i^{(j)}$ they used in (2.4) is identical

for all groups. Verbeke and Lesaffre (1996) proposed the so-called heterogeneity model, a linear mixed effects model with random effects sampled from a mixture of normal distributions. Verbeke and Molenberghs (2000) indicated that the classification rule implied by the heterogeneity model is equivalent with the discriminant function proposed by Tomasko *et al.* (1999). Brant et al. (2002) described the PSA evolution in different diagnostic groups using linear mixed effects models with group-specific covariance matrices (hereby illustrating an extension of the traditional quadratic discriminant function). They used a cross-validation framework and emphasized dynamic classification. Marshall and Barón (2000) considered nonlinear mixed effects models to describe evolutions in different groups, but used linear approximations to these models. Hereby, the resulting marginal models in each group are still multivariate normal, such that their approach is a specific example of (2.4).

This paper generalizes the above presented approach for discriminant analysis based on longitudinal profiles, such that densities implied by nonlinear mixed effects models are used and models in the different diagnostic groups are of different types (linear and nonlinear). A mixture of multivariate normal distributions as marginal distribution for the $\boldsymbol{Y}_i$, is not appropriate if a nonlinear model is used to describe the evolution in at least one group. Therefore, the mixture presented in (2.4) has to be replaced by the more general expression

$$\boldsymbol{Y}_i \sim \sum_{j=1}^{G} p^{(j)} f_i^{(j)} \tag{2.5}$$

where $f_i^{(j)}$ is the marginal distribution of the nonlinear mixed-effects model or a multivariate normal distribution if a linear mixed model has been used. For classification purposes the posterior probabilities for each subject to belong to each of the diagnostic groups have to be calculated. This is again done in the classical way, using expression (2.2) with the appropriate densities $f_i^{(j)}$. Conceptually, the procedure thus remains the same. In a first step, a description of the profiles is obtained in each group. In a second step these descriptions are used to generate for each subject a density in each group, yielding the posterior probabilities needed for the classification. Note that fitting of these models, as well as calculation of the densities $f_i^{(j)}$, increases the computational complexity severely. Indeed, the marginal densities corresponding to nonlinear mixed-effects models involve integrals with no closed-form solution. One way to solve this problem is using Gaussian quadrature methods, which are

for example available in the SAS procedure PROC NLMIXED. Details of this approximation can be found in the manual for PROC NLMIXED (SAS Institute, 1999).

3 Application : classification of longitudinal PSA profiles

The PSA dataset has already been mentioned briefly in the introduction. This dataset contains repeated PSA measures available from the Baltimore Longitudinal Study of Aging (BLSA). Started in 1958, this study accumulates information from repeated clinical examinations, including frozen blood samples. Participants in this study undergo three days of biomedical and psychological evaluations approximately every two years. The dataset contains a subset of 342 male participants with no evidence of prostate cancer at the start of the study and with a follow-up time of at least 10 years. In the diagnosis for prostate disease four different groups are distinguished. Men diagnosed as normal have no evidence of prostate disease. Men with BPH had a prostatectomy for a benign prostate enlargement. A diagnosis of local cancer is made when the presence of the cancer is limited to the prostate, while the diagnosis of metastatic cancer refers to a cancer spread beyond the prostate. Figure 1 presents the subject-specific PSA profiles of 33 BPH cases, 26 local cancer cases, 8 metastatic cancer cases and a random selection out of 275 normal cases.

Step 1 : Description of group-specific PSA profiles

In agreement with previous work (Pearson *et al.* 1994, Morrell *et al.*, 1995), linear mixed-effects models with quadratic time effects can be used to model the PSA evolutions. With $Y_i(t)$ denoting the log-transformed PSA value for subject i, taken at time t, and following the notation introduced in (2.3), the model assumes that in each non-cancer group and each cancer group $Y_i(t)$ satisfies

$$Y_i(t) \;=\; (\beta_{01} + b_{0i}) + \beta_{02} Age_i + (\beta_1 + b_{1i})t_i + (\beta_2 + b_{2i})t_i^2 + \varepsilon_i(t), \qquad (3.1)$$

where Age_i equals the age at entry of the ith subject and three random effects are used to model subject-specific deviations from the average quadratic profile.

Pearson *et al.* (1994) argued that a linear mixed model is not appropriate to describe the evolutions in all groups. Instead, they have fitted piecewise nonlinear mixed-effects models to estimate the PSA evolutions in the cancer groups. The evolution of the PSA for cancer cases is described by four parameters. Figure 2 illustrates the parameters of this model.

Each individual's longitudinal PSA trend possesses a period of slow linear change in PSA (the slope of this linear change is denoted by φ_1). This period is then followed by a period of rapid exponential increase (with φ_3 denoting the exponential rate). The unknown transition point between these two phases is denoted by φ_2, while φ_0 represents the PSA level at the transition point.

As mentioned by Pearson *et al.* (1994) the exponential phase is biologically appropriate since it is known from other studies that the doubling time of prostate tumour volume is much faster than normal or benign growth and that prostate tumour cells contribute approximately ten times more PSA to the bloodstream than normal or benign prostate cells. Since the true time of transition from the slow linear phase to the rapid exponential phase is unknown for each individual and is influenced by random factors affecting time of diagnosis, the transition time is included as a random effect in the model (denoted by u_{2i}). This random effect allows the transition times to vary among individuals. To account for natural heterogeneity in the rates of growth of prostate tumour growth, they also included the exponential growth parameter as a random effect (denoted by u_{3i}). This random effect thus allows the exponential rate to vary among individuals. The model now becomes

$$Y_i(t) = \begin{cases} \varphi_0 + \varphi_1[t - (\varphi_2 + u_{2i})] + \varepsilon_i(t) & \text{if } t \leq \varphi_2 + u_{2i} \\ \varphi_0 \exp\{(\varphi_3 + u_{3i})[t - (\varphi_2 + u_{2i})]\} + \varepsilon_i(t) & \text{if } t > \varphi_2 + u_{2i} \end{cases} \tag{3.2}$$

Further, a transition function has been used to provide a smooth transition between the two phases. More details can be found in their article.

Using the former information, we decided to describe the PSA evolution in the non-cancer groups by linear mixed models. Note that the PSA values evolve linearly in these groups such that in model (3.1) a logarithmic transformation of the PSA is not needed and the quadratic term can be dropped. To describe the PSA evolution in the cancer groups, piecewise nonlinear mixed models were used. To obtain the parameter estimates in the piecewise nonlinear mixed-effects models Pearson *et al.* (1994) used a quasi-likelihood method. Since a (Gaussian) quadrature method is more accurate, we used this approximation method in the procedure PROC NLMIXED to fit the piecewise model. To avoid variances becoming negative during the estimation process, standard deviations of the random effects were estimated. Note that in accordance with the linear mixed model, we allowed an effect of age at entry such that $\varphi_0 = \varphi_{01} + \varphi_{02} Age_i$. As opposed to the mentioned previous work on

the PSA dataset, the analyses presented in this paper express time with entry in the study instead of the moment of diagnosis (or last contact) as reference point. Indeed, the latter information is only available in retrospective studies where classification is superfluous. The resulting estimates for the non-cancer and the cancer groups are given in Table 1. Note that these estimates are obtained by fitting the linear or nonlinear mixed model in each group separately. Figure 3 shows a plot of the "average profiles" (profile of a subject with random effects equal to zero) per group.

Step 2 : Calculation of posterior probabilities

The first step in the classification procedure is the description of the longitudinal PSA profiles using linear mixed models for the non-cancer groups and nonlinear mixed models for the cancer groups. Results of this step were presented in the previous section. Using these estimates, the next step calculates the density of each subject i to belong to group j. A SAS-macro has been written for the calculation of these densities in each group separately and for generating the posterior probabilities $p_i^{(j)}$. As prior mixing proportions $p^{(j)}$, the observed proportion of patients in each diagnostic group has been used.

Classification results

Table 2 presents the classification results based on the discriminant analysis framework with a linear mixed model describing the PSA evolutions in the non-cancer groups and a nonlinear mixed model describing the evolutions in the cancer groups. The correct classification rate (CCR) equals 82.8% (283/342). However, the set of misclassified patients contains many misclassifications without a large practical impact (confusion benign with normal cases and confusion local with metastatic). Ignoring this type of misclassifications, the CCR equals 94.4% (323/342). Figure 4 gives a detailed overview of the PSA profiles of the correct and incorrect classified patients. The general impression is that the classification approach yields meaningful results. Patients with PSA profiles similar to the average trend in a specific diagnostic group are indeed classified in that diagnostic group: the PSA evolution of the correctly classified normal patients remains constant or increases slightly whereas the correctly classified BPH patients show a higher increase of PSA, but their increase is still linear. All correctly classified cancer cases are characterised by an obvious exponential phase, with the metastatic cancer cases having a higher exponential rate than the local cancer cases. Also most misclassified cases can easily be understood by comparing their profile with the

average profile of the diagnostic group they belong to. False negative cases (cancer cases classified as non-cancer cases) clearly lack any indication of an exponential phase, whereas most false positive cases (non-cancer cases classified as cancer cases) show at least a jump in their PSA profile. The observed false positive rate ($36.6\% = 15/41$) is much higher than the false negative rate ($2.7\% = 8/301$).

As indicated in the introduction, Brant *et al.* (2003) used only linear mixed models for the classification of the PSA profiles. To get an idea of the added flexibility of our approach, we have repeated their classification procedure. This implies that linear mixed models with quadratic time effects are used to describe the evolution of the (log-tranformed) PSA values of the non-cancer as well as the cancer cases. Figure 5 gives for the 308 non-cancer cases the posterior probabilities to be classified as a cancer case for both approaches. As can be derived from this figure, using non-linear mixed models does not seem to decrease the false positive rate. Note that the set of non-cancer cases classified as a cancer case is not exactly the same for both approaches. Figure 6 gives for the 34 non-cancer cases the posterior probabilities to be classified as a non-cancer case for both approaches. Clearly, the observed number of false negative cases is reduced when using non-linear mixed models. As mentioned by Brant *et al.* (2003), false positives would lead the physician to further investigate for the presence of cancer by using additional diagnostic tests, whereas the consequences of false negatives obviously can be more dramatic.

In the analysis presented sofar, the classification is based on all available information. Also verified is how the classification progressively changes as more information becomes available. Therefore, a first classification of an individual PSA profile is made 6 months after entry in the study, a second classification 1 year after entry in the study and so on. To illustrate the gained flexibility by using nonlinear instead of linear mixed models for the description of the PSA profiles in the cancer groups, we compared the results of this sequential classification procedure between the two approaches. Since it is important in a clinical context to detect the onset of the prostate cancer as soon as possible, we compared the posterior probabilities of the cancer cases between the two approaches. The higher the posterior probability, the more prone the clinician will be to suspect the presence of a cancer (and the sooner this presence will be detected in practice). Therefore, Figure 7 presents for the two approaches the evolution of the median posterior probabilities (of the 34 cancer cases) to belong to the cancer groups. These median posterior probabilities do not seem to differ much during the first period after entry in the study. However, when time evolves, the median posterior

probability becomes much higher for the approach using a nonlinear model. In practice this would be reflected in a faster detection of prostate cancer. For example, using linear models, the median posterior probability reaches the 0.50 level only 22.5 years after entry in the study, while using nonlinear models this level is reached 5 years earlier.

4 Discussion

This paper uses a flexible classification procedure for longitudinal profiles. In agreement with previous work, by using mixed-effects models, the approach is appropriate to classify longitudinal profiles of datasets with an unbalanced structure of the data. Also, covariates can be add to the classification procedure. Our approach extends the use of linear mixed models for the classification process to the use of nonlinear mixed models, this without relying on linearized versions of nonlinear models. This extension increased the computational complexity since the marginal distributions used in the classification process can no longer be derived analytically. By allowing nonlinear mixed models in the classification procedure and using mixed models of different types, the approach becomes more flexible. Using the PSA data we have shown that substantial gain is obtained in the classification results, justifying the use of this computationally more complex method. Indeed, using nonlinear mixed effects models could - by yielding higher posterior probabilities for the cancer cases - increase the speed of detection of the cancer onset and reduce the number of false negative cases. In practice, a trade-off has to be made between the added flexibility of allowing models of different types and the increased computational complexity. This paper does not use a crossvalidation framework, hereby overestimating the performance of the classification process. However, the focus of the current paper is on the use of nonlinear mixed models within a discriminant analysis, more than on the performance of the classification procedure in this particular dataset.

Another method to cope with sparseness of measurements has been recently introduced by James and Hastie (2001). These authors extend the classical discriminant analysis into a functional data analysis framework, using natural cubic spline functions to model observations of one subject. However, for the dataset under consideration, such an approach is less appealing, since the biologic motivated parametric modelling of the PSA evolutions would

be lost.

To further improve the classification performance for the PSA data, more information is needed. This can be done by inclusion of more covariates or by considering other markers besides the PSA level (thus extending the framework to a multivariate one). As indicated, adding covariate information does not change the classification approach. However, extending the framework to a multivariate one is an interesting topic for further research.

References

Albert, A. (1983), "Discriminant analysis based on multivariate response curves : a descriptive approach to dynamic allocation", *Statistics in Medicine*, **2**, 95–106.

Brant, L.J., Sheng, S.L., Morrell, C.H., Verbeke, G.N., Carter, H.B., and Lesaffre, E. (2003), "Screening for prostate cancer using random-effects models", *Journal of the Royal Statistical Society, Series A*, **166**, 51–62.

Carter, H.B. and Coffey, D.S. (1990), "The prostate : An increasing medical problem", *The Prostate*, **16**, 39–48.

James, G.M. and Hastie, T.J. (2001), "Functional linear discriminant analysis for irregularly sampled curves", *Journal of the Royal Statistical Society, Series B*, **63**, 533–550c.

Laird, N.M. and Ware, J.H. (1982), "Random-effects models for longitudinal data", *Biometrics*, **38**, 963–974.

Marshall, G. and Barón, A.E. (2000), "Linear discriminant models for unbalance longitudinal data", *Statistics in Medicine*, **19**, 1969–1981.

Morrell, C.H., Pearson, J.D., Carter, H. B., and Brant, L.J. (1995), "Estimating unknown transition times using a piecewise nonlinear mixed-effects model in men with prostate cancer", *Journal of the American Statistical Association*, **90**, 45–53.

Pearson, J.D., Kaminski, P., Metter, E.J., Fozard, J.L., Brant, L.J., Morrell, C.H., and Carter, H.B. (1991), "Modeling longitudinal rates of change in prostate specific antigen during aging", *Proceedings of the Social Statistics Section of the American Statistical Assciation, Washington D.C.*, 580–585.

Pearson, J.D., Morrell, C.H., Landis, P.K., Carter, H.B., and Brant, L.J. (1994), "Mixed-effects regression models for studying the natural history of prostate disease", *Statistics in Medicine*, **13**, 587–601.

SAS Institute (1999), *The SAS System for Windows, Release V8.0*, Cary, NC: SAS Institute Inc.

SAS Institute Inc., Cary, NC: SAS Institute Inc. (1992), *SAS Technical Report P-229, SAS/STAT Software : Changes and Enhancements, Release 6.07*.

Tomasko, L., Helms, R.W., and Snapinn, S.M. (1999), "A discriminant analysis extension to mixed models", *Statistics in Medicine*, **18**, 1249–1260.

Verbeke, G. and Lesaffre, E. (1996), "A linear mixed-effects model with heterogeneity in the random-effects population", *Journal of the American Statistical Association*, **91**, 217–221.

Verbeke, G. and Molenberghs, G. (2000), *Linear mixed models for longitudinal data*, Springer Series in Statistics, Springer-Verlag, New-York.

Table 1: *Parameter estimates from fitting a linear mixed model on the PSA values in each non-cancer group separately and a nonlinear mixed model in each cancer group, with time expressed as decades since entry.*

Non-Cancer groups

Estimate	Normal	BPH
β_{01}	-0.115	-1.624
β_{02}	0.014	0.048
β_1	0.338	1.064
$\sigma^2_{b_0}$	0.178	0.399
$\sigma^2_{b_1}$	0.099	0.510
σ_{b_0,b_1}	-0.028	0.225
σ^2_{ε}	0.135	0.490

Cancer groups

Effect	Local	Metastatic
φ_{01}	1.982	7.220
φ_{02}	-0.007	-0.067
φ_1	0.668	1.470
φ_2	1.210	1.123
φ_3	1.747	5.414
σ_{u_2}	0.694	0.820
σ_{u_3}	0.717	2.75
σ_{u_2,u_3}	0.381	1.840
σ^2_{ε}	0.406	5.997

Table 2: *Classification results using linear mixed models and nonlinear mixed models for the description of the PSA profiles in respectively the non-cancer and the cancer groups.*

	Diagnosis				
	normal	bph	local	meta	
Classification					
normal	247	18	4	0	269
bph	16	12	4	0	32
local	10	1	17	1	29
meta	2	2	1	7	12
	275	33	26	8	342

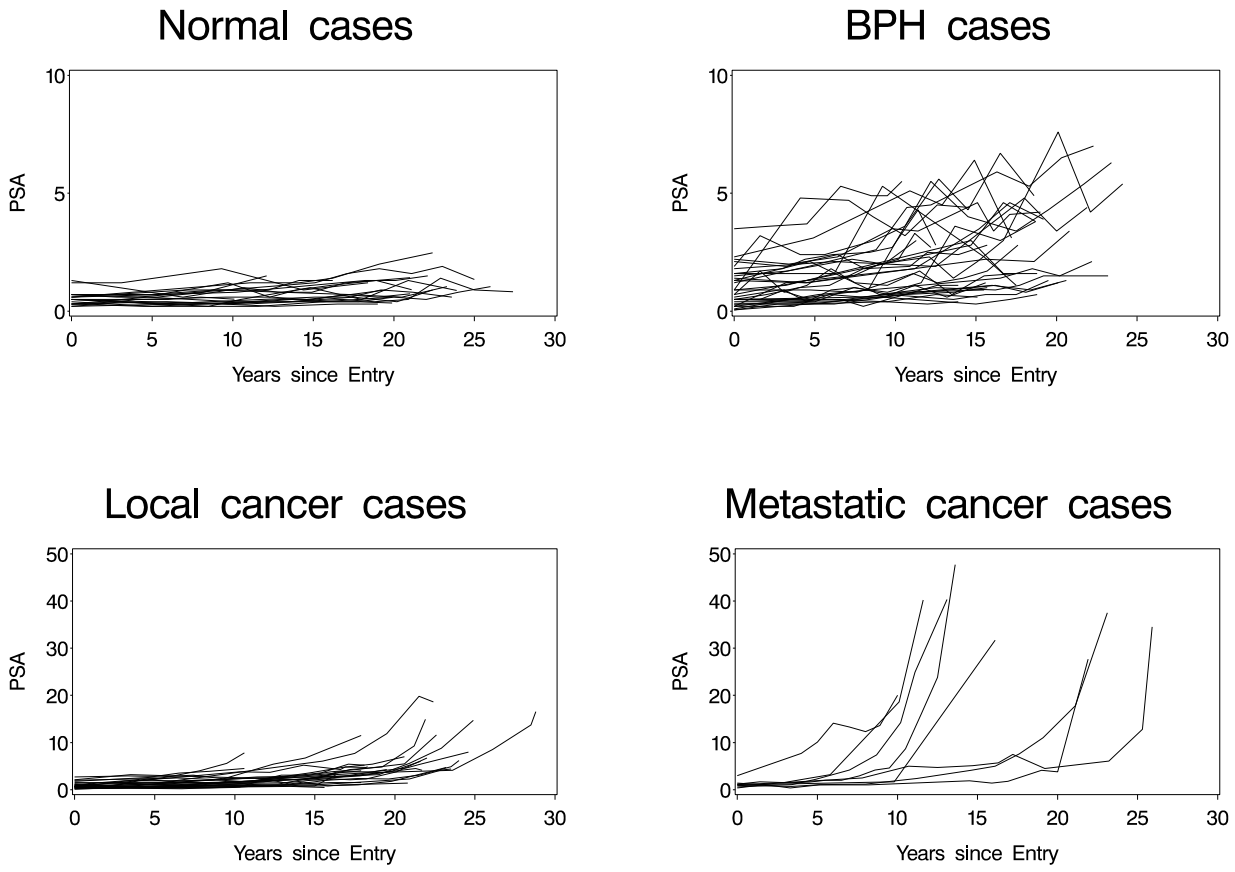


Figure 1: *Subject-specific PSA profiles for four diagnostic groups: normal men with no evidence of prostate disease (random selection of 30 out of 275 cases), men with benign prostatic hyperplasia (BPH), men diagnosed with a local cancer and men diagnosed with metastatic cancer. Note that the scale for the PSA values differs as a function of the diagnostic group.*

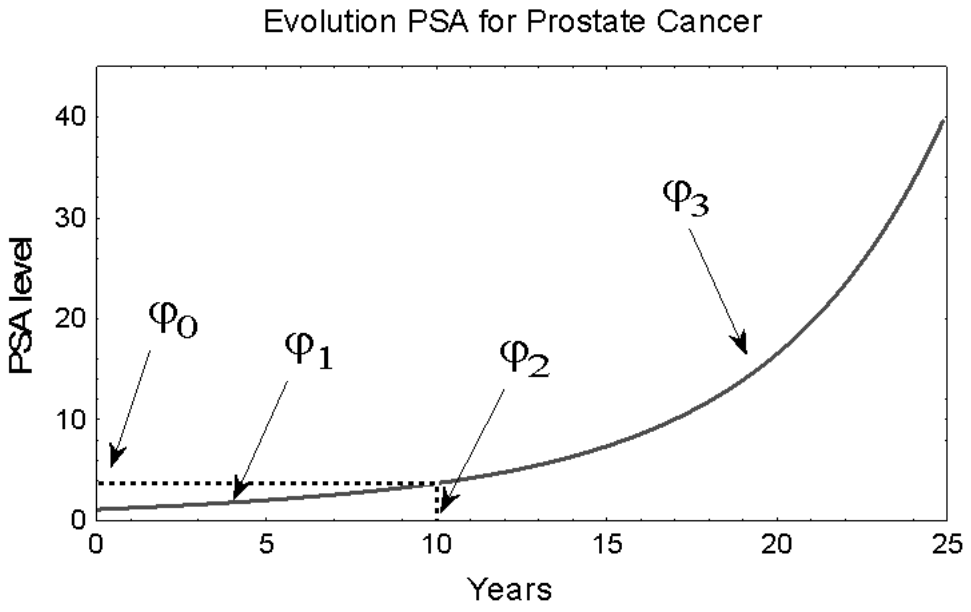


Figure 2: *The piecewise model describing the PSA evolution for the cancer cases*

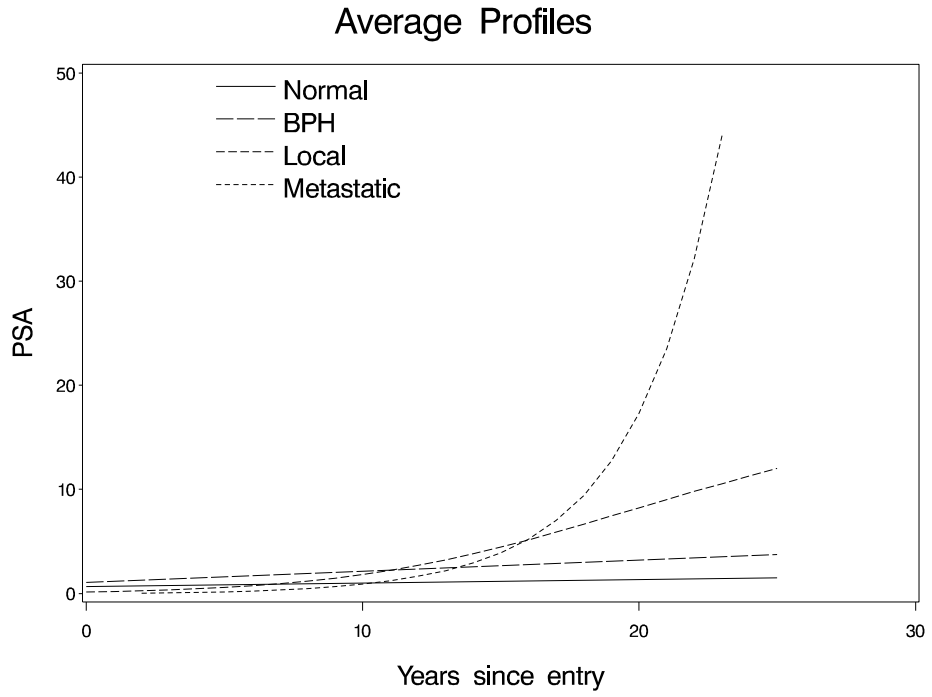


Figure 3: *Fitted "average" (a subject with random effects equal to zero) profiles from fitting a linear mixed model on the non-cancer cases and a nonlinear mixed model on the cancer cases.*

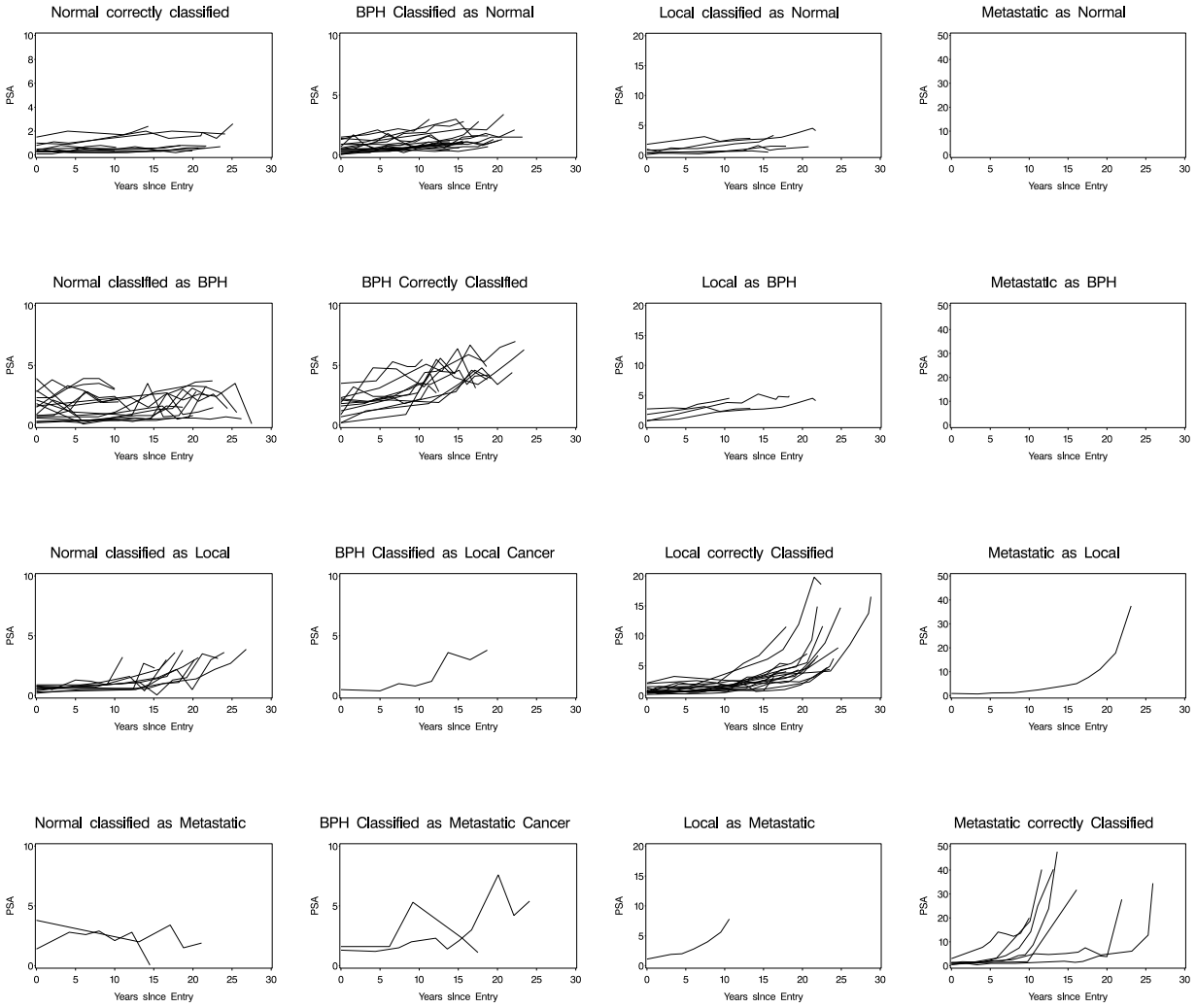


Figure 4: *PSA profiles according to diagnostic group and classification (the scale for the PSA values differs as a function of the diagnostic group and the figure with the correctly classified normals is a random selection of the 239 profiles)*

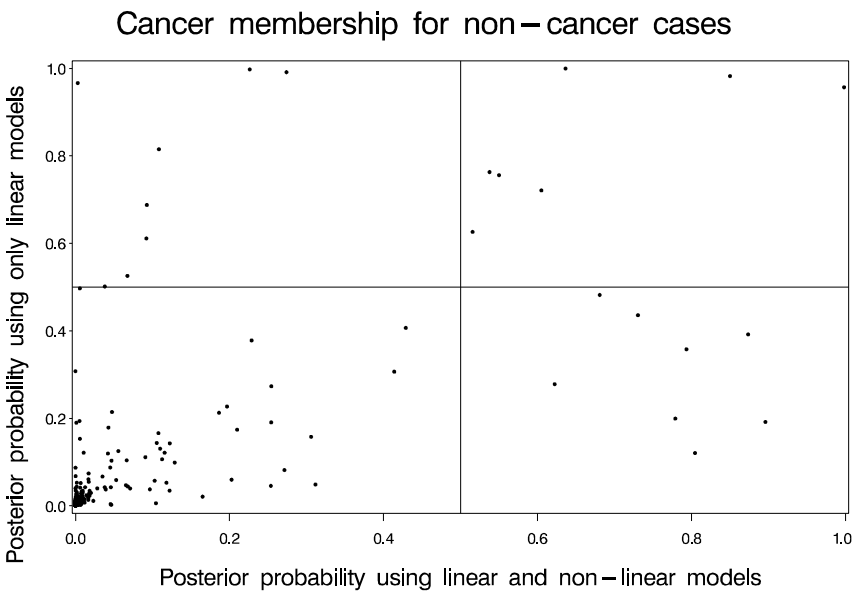


Figure 5: *Posterior probabilities to be classified as a cancer case for the 308 non-cancer cases. Posterior probabilities obtained by the two approaches are given (only using linear mixed models and using linear as well as nonlinear mixed models for respectively the non-cancer and the cancer cases).*

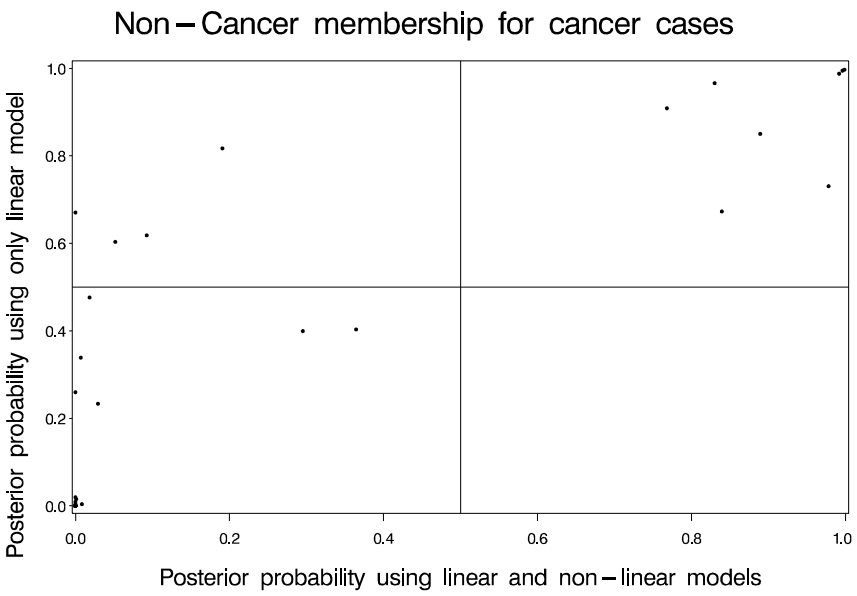


Figure 6: *Posterior probabilities to be classified as a non-cancer case for the 34 cancer cases. Posterior probabilities obtained by the two approaches are given (only using linear mixed models and using linear as well as nonlinear mixed models for respectively the non-cancer and the cancer cases).*

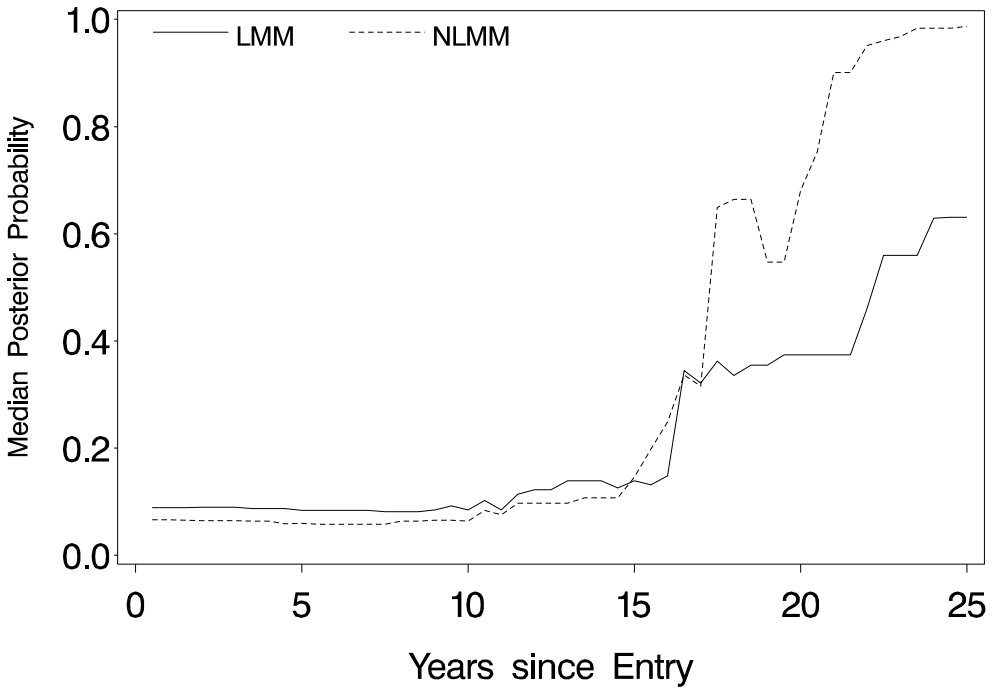


Figure 7: *Evolution of median posterior probabilities of 34 cancer cases to belong to a cancer group (LMM=Using linear mixed models to describe the cancer cases, NLMM=Using nonlinear mixed models to describe the cancer cases)*