Screening for prostate cancer by using random-effects models

Larry J. Brant and Shan L. Sheng,
National Institute on Aging, Baltimore, USA

Christopher H. Morrell,
National Institute on Aging, Baltimore, and Loyola College, Baltimore, USA

Geert N. Verbeke and Emmanuel Lesaffre
Katholieke Universiteit Leuven, Belgium

and H. Ballentine Carter
Johns Hopkins University School of Medicine, Baltimore, USA

[Received June 2000. Final revision May 2002]

Summary. Random-effects models are used to screen male participants in a long-term longitudinal study for prostate cancer. By using posterior probabilities, each male can be classified into one of four diagnostic states for prostate disease: normal, benign prostatic hyperplasia, local cancer and metastatic cancer. Repeated measurements of prostate-specific antigen, collected when there was no clinical evidence of prostate disease, are used in the classification process. An individual's screening data are considered one repeated measurement at a time as his data are collected longitudinally over time. Posterior probabilities are calculated on the basis of data from other individuals with confirmed diagnoses of each of the four diagnostic states.

Keywords: Cancer diagnosis; Classification; Disease screening; Linear mixed effects model; Longitudinal data; Prostate-specific antigen

1. Introduction

Prostate cancer is one of the common causes of deaths from cancer in men, and after lung and stomach cancer it accounts for the largest number of new non-skin-cancer cases reported worldwide per year (Pisani et al., 1999). In the USA, for example, prostate cancer is the most common clinically diagnosed non-skin cancer with about 1 in 10 American men eventually receiving a positive diagnosis. Since the chance of a diagnosis of prostate cancer increases with age, the present shift in the age distribution towards larger numbers of older men is expected to result in an even larger increase in the number of men diagnosed with prostate cancer (Carter and Coffey, 1990).

Prostate-specific antigen (PSA) is a glycoprotein that is produced by the prostatic epithelium and can be measured in serum samples by immunoassay (Wang et al., 1979). Since PSA levels correlate with the volume of cancer in the prostate, it has been found to be useful in the
management of men with prostate cancer. As PSA levels increase, the extent of cancer and its chance of detection increase (Stamey et al., 1987; Cooner et al., 1990; Catalona et al., 1991; Carter and Pearson, 1994). Brant and colleagues (Carter et al., 1992; Pearson et al., 1994; Morrell et al., 1995) have demonstrated the usefulness of longitudinal measurements of PSA levels for studying the natural history of prostate disease, and Carter and Pearson (1994) have discussed the epidemiological utility of using an average rate of change of 0.75 $\mu$g l$^{-1}$ year$^{-1}$ or greater as a screening criterion for the detection of prostate cancer. The sensitivity and specificity of this criterion were reported as 0.72 and 0.95 respectively.

Diagnoses for prostate disease include normal, benign prostatic hyperplasia (BPH), local cancer and metastatic cancer. Men diagnosed as normal have no evidence of prostate disease, whereas men with BPH had a simple prostatectomy for enlargement of the prostate. A diagnosis of local cancer is made when the cancer is determined to be confined only to the prostate and a metastatic diagnosis is made when the cancer spreads beyond the prostate. Previous studies have shown that PSA levels tend to become elevated in cancer cases with malignant prostatic tumour cells contributing approximately 10 times more PSA to the blood stream than normal or benign prostate cells (Stamey and Kabalin, 1989). Morrell and colleagues (Morrell et al., 1995) modelled this biological phenomenon by using a piecewise model to describe the transition from a background level of slow PSA change representing a normal and/or BPH growth of the prostate to an exponential PSA phase of rapid growth in the prostate due to an increase in growth of a tumour. This piecewise modelling approach suggested that the rapid increases in PSA in cancer cases begins on average at least 5 years before a clinical diagnosis, allowing a possible window of opportunity for the early detection of prostate cancer.

In health clinics, medical practitioners often collect PSA measurements sequentially or one examination at a time for use in the detection of prostate cancer. This paper uses longitudinal measurements of PSA to screen for prostate cancer. A method of classification is given based on a linear mixed effects model that classifies men with repeated PSA measurements into non-cancerous (normal or BPH) or cancer (local or regional or metastatic) diagnostic groups. On the basis of an individual’s predicted starting level of PSA (micrograms per litre) and his predicted rate of change in PSA (micrograms per litre per year) a cross-validation study of a longitudinally collected PSA data set is performed to classify the individual into a preclinical diagnostic state of non-cancerous or cancerous.

2. Data and methods

2.1. Data

The data used in this paper come from the Baltimore Longitudinal Study of Aging (BLSA), a continuing longitudinal and multidisciplinary study of normal human aging that was begun in 1958 (Shock et al., 1984). The S-PLUS programs that were used to analyse them can be obtained from

http://www.blackwellpublishers.co.uk/rss/

The study is conducted as part of the intramural research programme of the National Institute on Aging. All participants who enter the BLSA are volunteers and are given a careful health screening to ensure that they are of excellent health with no known diseases. Although our sample is not randomly taken from the US population, the sample consists of participants who were chosen from a large group of waiting-list volunteers so that the age distribution of the BLSA sample resembles the age distribution in the US population. Participants in the study
visit the Gerontology Research Center in Baltimore approximately every 2 years for 3 days of biomedical and psychological examinations, including determinations of PSA and urological examinations. The study attempts to maintain maximum participation from each individual. All participants are continually monitored to obtain information regarding their health status, especially information related to prostate disease and other disease events. This monitoring continues over time regardless of the collection of PSA measurements. In the case of death, information is received from the individual’s personal physician regarding the cause of death and autopsy information is obtained when available.

During the entire course of the BLSA, approximately 1580 men have been enrolled in the longitudinal study. Since PSA measurements have only been routinely measured in recent years, the number of men with two or more PSA measurements is 643. Many of the men who were not included in the study were either deceased or no longer participating in the study before the measurement of PSA.

The aim of this study is to investigate the prediction of the development of prostate cancer in men who started in the study without prostate disease. To achieve this, an adequate follow-up time and a sufficient number of visits will be required to be able to estimate accurately each individual participant’s PSA profile. Morrell et al. (1995) estimated that advanced cancers develop an average of 10.7 years before the clinical detection of cancer. Consequently, for this study, we require that participants have at least 10 years of follow-up as well as at least five repeated PSA measurements. This paper uses data from 342 BLSA male participants with no evidence of cancer at the start of the study and with at least 10 years of follow-up to develop a classification method for prostate cancer. The remaining 301 participants with 2–4 visits or a follow-up time of less than 10 years will be used as a check of the classification method. The 342 participants who were used in developing the classification method include 275 with no evidence of prostate disease, 26 who developed local or regional prostate cancer, eight who developed metastatic cancer and 33 who developed BPH. None of the 342 patients showed evidence of prostate disease at the beginning of the follow-up period and all PSA measurements were taken before a clinical diagnosis of prostate disease. For prostate cancer cases, the disease was clinically diagnosed sometime during the follow-up period after which no PSA measurements were considered in this study. BPH cases involved those men having a simple prostatectomy for enlargement of the prostate. Table 1 summarizes the length of follow-up, number of repeated PSA measurements per individual, time interval between repeated PSA measurements and the age of diagnosis for the normal, BPH and cancer cases. If men remained in the normal group throughout the study’s follow-up period, then their age at diagnosis was taken to be their age at their last examination. Table 1 indicates that the normal men have the widest range of age at diagnosis from 40.1 to 92.0 years, with a somewhat lower median age of 64.7 years than the other diagnostic groups. In addition, the metastatic group overall has the shortest length of follow-up (median 16.8 years), the shortest length of time between repeated PSA measurements (median 1.6 years) and the most repeated PSA measurements per individual (median 10 measurements).

Fig. 1 displays the observed repeated PSA measurements for five systematically selected normal males (Fig. 1(a)) and five systematically selected males who subsequently developed local cancer (Fig. 1(b)). In the selection process, after the means of all the repeated PSA measurements for each subject were ordered in the normal and local cancer groups, the men with the minimum, the first quartile, the median, the third quartile and maximum means of repeated PSA measurements were chosen from each group. The graph was generated on the basis of these 10 selected subjects to illustrate the different trends over time for the normal and local cancer groups. As a reference value, the graph also shows the PSA level of 4 µg l⁻¹. A PSA value
Table 1. Description of the prostate data by diagnosis group

<table>
<thead>
<tr>
<th></th>
<th>Normal</th>
<th>BPH</th>
<th>Local or regional</th>
<th>Metastatic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of participants</td>
<td>275</td>
<td>33</td>
<td>26</td>
<td>8</td>
</tr>
<tr>
<td>Age at diagnosis (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>64.7</td>
<td>72.2</td>
<td>70.0</td>
<td>76.6</td>
</tr>
<tr>
<td>Range</td>
<td>40.1–92.0</td>
<td>57.4–83.9</td>
<td>49.0–85.7</td>
<td>62.6–84.0</td>
</tr>
<tr>
<td>Length of follow-up (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>20.4</td>
<td>18.6</td>
<td>20.7</td>
<td>16.8</td>
</tr>
<tr>
<td>Range</td>
<td>10.0–29.9</td>
<td>10.4–24.1</td>
<td>10.6–28.8</td>
<td>10.0–25.9</td>
</tr>
<tr>
<td>Number of repeated PSA measurements per individual</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>7</td>
<td>7</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>Range</td>
<td>5–14</td>
<td>5–14</td>
<td>5–15</td>
<td>5–14</td>
</tr>
<tr>
<td>Interval between repeated PSA measurements (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>2.2</td>
<td>2.1</td>
<td>2.0</td>
<td>1.6</td>
</tr>
<tr>
<td>Range</td>
<td>0.2–16.6</td>
<td>0.9–11.5</td>
<td>0.1–15.2</td>
<td>0.6–9.8</td>
</tr>
</tbody>
</table>

Fig. 1. Observed repeated PSA measurements for (a) five systematically selected normal males and (b) five systematically selected males who subsequently develop local cancer

at least this large is sometimes used as a detection criterion for prostate cancer. An examination of Fig. 1 suggests that PSA measurements for approximately 25% of the local cancer cases give little evidence for the detection of prostate cancer. However, for the remaining local cancer cases and the normal cases, a classification approach based on the level and rate of change in PSA appears to be a feasible method for assisting clinicians in the diagnosis of prostate cancer. Random-effects models provide a method to capture the heterogeneity in the PSA values that can be used in the disease classification process.
2.2. Model specification

In diagnostic group \( c \), it will be assumed that the vector of longitudinal PSA measurements follows a linear mixed model as

\[
y_i = X_i\beta_c + Z_i b_i + \varepsilon_i,
\]

where \( y_i \) denotes the vector of \( \ln(1 + \text{PSA}) \) measurements on the \( i \)th male and \( X_i \) and \( Z_i \) are the corresponding design matrices for the fixed and random effects respectively. These design matrices may contain columns corresponding to age at first examination and time and time squared (in years) since the first PSA measurement. The random effects \( b_i \) have a multivariate normal distribution with mean \( \mathbf{0} \) and positive definite covariance matrix \( D_c \). The vector of measurement errors \( \varepsilon_i \) is assumed to be independent and normally distributed with mean \( \mathbf{0} \) and variance–covariance matrix \( \sigma^2_c \mathbf{I} \).

This mixed effects model for the longitudinal PSA data is employed in the cross-validation procedure to classify the participants by calculating posterior probabilities of membership in the normal, BPH, local or regional and metastatic diagnostic groups. The procedure for classifying the \( k \)th individual is

(a) fit the models (1) to the data in each diagnostic group \( c \), omitting participant \( k \), which yields estimates \( \hat{\beta}_{c(k)} \), \( \hat{D}_{c(k)} \) and \( \hat{\sigma}^2_{c(k)} \), and

(b) compute the marginal distribution for each diagnostic group for this individual \( k \).

This marginal distribution is then given by

\[
y_k|\text{group } c \sim N(X_k\hat{\beta}_{c(k)}, Z_k\hat{D}_{c(k)}Z_k^T + \hat{\sigma}^2_{c(k)} \mathbf{I})
\]

and the corresponding density is denoted by \( f_{c(k)}(y_k) \). Given prior probabilities \( p_c, c = 1, \ldots, 4 \), for the four diagnostic groups, and applying the Bayes theorem, the posterior probability that individual \( k \) belongs to group \( c \) is given by

\[
p_{kc} = p_c f_{c(k)}(y_k) \left/ \sum_{j=1}^{4} p_j f_{j(k)}(y_k) \right.,
\]

in which the prior probabilities \( p_c \) are replaced by the observed proportions of the men in each diagnostic group. Note that, although the linear mixed model (1) specifies the response vector conditionally on a vector \( b_i \) of random effects, classification is based on the marginal distribution obtained from integrating over the random effects.

The classification process proceeds for individual \( k \) by first calculating the posterior probabilities in equation (2) using the first two PSA measurements, and then sequentially repeating the process by adding one measurement at a time until the classification stopping rule is met or all the measurements have been used for individual \( k \). Fig. 2 shows the classification rule that is used for the PSA data with \( p_{k1}, p_{k2}, p_{k3} \) and \( p_{k4} \) representing the posterior probabilities for the normal, BPH, local or regional and metastatic diagnostic groups. Since the aim of this classification procedure is to determine when an individual has developed preclinical cancer, the classification rule in Fig. 2 seeks to give the greatest chance of detecting the disease. In particular, if the posterior probability of either of the cancer groups is largest, then the patient is classified as a cancer case, i.e. if the posterior probabilities for either the local or regional group (\( p_{k3} \)) or the metastatic group (\( p_{k4} \)) are the largest of the four probabilities. Next, if the combined probabilities of the cancer groups are larger than the combined probabilities of the non-cancer groups, or larger than both the normal and BPH probabilities, i.e. if \( p_{k3} + p_{k4} \gtrless p_{k1} + p_{k2} \) or \( p_{k3} + p_{k4} \gtrless p_{k1} \) and \( p_{k3} + p_{k4} \gtrless p_{k2} \), then individual \( k \) is classified as a cancer case. Otherwise,
CLASSIFICATION RULE

Fig. 2. Classification rule for the sequential PSA data from the kth individual

he is considered as a non-cancer case. If the participant has not been classified as a cancer case by his final PSA measurement, the individual is classified as either normal or BPH on the basis of his posterior probabilities at the last measurement.

2.3. Model fitting

In an earlier work, Pearson et al. (1994) described the patterns of change in PSA in the normal, BPH, local cancer and metastatic cancer groups by using a linear mixed effects model for ln(1 + PSA). The independent variables for the model included variables for the different diagnostic groups along with variables for age at diagnosis, time since first PSA measurement and interaction terms involving the diagnostic groups and time.

In this paper, the final linear mixed effects model for the prediction of prostate cancer using the PSA data is

\[
\ln(1 + PSA_{ij}) = \beta_1 \text{Age}_i + \beta_{21} \text{Normal}_i + \beta_{22} \text{BPH}_i + \beta_{23} \text{LRCancer}_i + \beta_{24} \text{MetCancer}_i + b_{1i} + (\beta_{31} \text{Normal}_i + \beta_{32} \text{BPH}_i + \beta_{33} \text{LRCancer}_i + \beta_{34} \text{MetCancer}_i + b_{2i})t_{ij} + (\beta_{41} \text{Normal}_i + \beta_{42} \text{BPH}_i + \beta_{43} \text{LRCancer}_i + \beta_{44} \text{MetCancer}_i + b_{3i})t_{ij}^2 + \epsilon_{ij},
\]

where PSA_{ij} denotes the jth measurement of PSA on the ith male. The variables Normal_i, BPH_i, LRCancer_i and MetCancer_i are binary 0–1 and denote the diagnostic group to which individual i belongs. Age_i is the age at first examination for individual i. The variable t_{ij} represents the time
In years since the first PSA measurement. This model was obtained by backward elimination from the full model where the intercept, time and time squared coefficients are allowed to vary between the four diagnostic groups. In this final model, each of the diagnostic groups has a different time squared coefficient, as well as different time coefficients and different intercepts or initial values. The random effect $b_{1i}$ allows each individual to have his own initial PSA level, and the random effects $b_{2i}$ and $b_{3i}$ allow each participant to have their own coefficient of the time and time squared terms respectively. The expected profile for each diagnostic group is then

$$E\{\ln(1 + PSA_{ij})\} = \beta_1 \text{Age}_i + \beta_{21} t_{ij} + \beta_{31} t_{ij}^2$$

for normal individuals,

$$E\{\ln(1 + PSA_{ij})\} = \beta_1 \text{Age}_i + \beta_{22} t_{ij} + \beta_{32} t_{ij}^2$$

for BPH cases,

$$E\{\ln(1 + PSA_{ij})\} = \beta_1 \text{Age}_i + \beta_{23} t_{ij} + \beta_{33} t_{ij}^2$$

for local or regional cancer cases and

$$E\{\ln(1 + PSA_{ij})\} = \beta_1 \text{Age}_i + \beta_{24} t_{ij} + \beta_{34} t_{ij}^2$$

for metastatic cancer cases.

3. Results

To illustrate how PSA levels change with time, Fig. 3 shows the predicted average trends of PSA levels with respect to years of follow-up for the four diagnostic groups. Both the normal and the BPH (non-cancerous) groups show only a small linear (constant) increase in PSA with time, whereas the local and metastatic cancer groups show evidence of an exponential trend over time. This exponential increase is larger in the metastatic group than in the local cancer group. Fig. 3 also indicates that the longitudinal trends for the BPH and local cancers are very similar for the first 12 years. Thereafter, the local cancer cases diverge from the BPH group.

Table 2 illustrates the sequential approach outlined in Sections 2.2 and 2.3 for classifying the median subject ($m$th subject) whose data are shown in Fig. 1. The posterior probabilities are first...
Table 2. Sequential classification results for the patient representing the median of the PSA measurements for the local cancer group

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>PSA</th>
<th>Normal</th>
<th>BPH</th>
<th>Local</th>
<th>Metastatic</th>
</tr>
</thead>
<tbody>
<tr>
<td>52.4</td>
<td>0.7</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>53.8</td>
<td>0.8</td>
<td>0.812</td>
<td>0.102</td>
<td>0.069</td>
<td>0.017</td>
</tr>
<tr>
<td>55.7</td>
<td>0.8</td>
<td>0.823</td>
<td>0.097</td>
<td>0.069</td>
<td>0.011</td>
</tr>
<tr>
<td>57.2</td>
<td>0.8</td>
<td>0.840</td>
<td>0.087</td>
<td>0.068</td>
<td>0.005</td>
</tr>
<tr>
<td>58.8</td>
<td>0.8</td>
<td>0.860</td>
<td>0.076</td>
<td>0.063</td>
<td>0.001</td>
</tr>
<tr>
<td>60.4</td>
<td>1.0</td>
<td>0.851</td>
<td>0.081</td>
<td>0.067</td>
<td>0.001</td>
</tr>
<tr>
<td>61.9</td>
<td>1.8</td>
<td>0.728</td>
<td>0.128</td>
<td>0.142</td>
<td>0.002</td>
</tr>
<tr>
<td>63.6</td>
<td>1.8</td>
<td>0.652</td>
<td>0.147</td>
<td>0.200</td>
<td>0.001</td>
</tr>
<tr>
<td>65.3†</td>
<td>2.2</td>
<td>0.375</td>
<td>0.138</td>
<td>0.487</td>
<td>0.000</td>
</tr>
<tr>
<td>67.0</td>
<td>3.1</td>
<td>0.309</td>
<td>0.124</td>
<td>0.567</td>
<td>0.000</td>
</tr>
<tr>
<td>68.4</td>
<td>3.8</td>
<td>0.254</td>
<td>0.106</td>
<td>0.640</td>
<td>0.000</td>
</tr>
<tr>
<td>70.0</td>
<td>4.3</td>
<td>0.270</td>
<td>0.113</td>
<td>0.617</td>
<td>0.000</td>
</tr>
<tr>
<td>73.0</td>
<td>5.5</td>
<td>0.391</td>
<td>0.168</td>
<td>0.441</td>
<td>0.000</td>
</tr>
<tr>
<td>75.0</td>
<td>11.6</td>
<td>0.225</td>
<td>0.082</td>
<td>0.693</td>
<td>0.000</td>
</tr>
</tbody>
</table>

†Predicted as local cancer at age 65.3 years (10.6 years before clinical diagnosis).

calculated on the basis of the first two measurements taken at ages 52.4 and 53.8 years. After the first two observations for subject \( m \), \( p_{m1} + p_{m2} = 0.914 \), supporting an initial classification as non-cancer. The sequential approach continues one measurement at a time until the ninth measurement taken at age 65.3 years gives \( p_{m3} = 0.487 \), the largest of the four posterior probabilities, and a classification into the local cancer group. The posterior probabilities are also shown for the five PSA measurements taken after the classification was determined for this patient.

Using the stopping rule given in Fig. 2 and the sequential classification procedure, the classification results for prostate cancer from the longitudinal PSA measurements of the 342 men described in Table 1 are presented in Table 3, which shows that overall 277 of 342 (81.0%) males are correctly classified. Since the primary interest is to distinguish between cancer and non-cancer cases, if the BPH and normal groups are combined as non-cancer cases, then 297 of 342 (86.8%) males are correctly classified. In addition, if no distinction is made between local and metastatic cancer, then the classification rate becomes 88.3% (302/342). Thus 21 of 34 (61.8%)

Table 3. Classification results for prostate cancer based on longitudinal PSA measurements

<table>
<thead>
<tr>
<th>Classification</th>
<th>Cases with the following diagnoses:</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal</td>
<td>BPH</td>
</tr>
<tr>
<td>Normal</td>
<td>261</td>
<td>19</td>
</tr>
<tr>
<td>BPH</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Local</td>
<td>7</td>
<td>11</td>
</tr>
<tr>
<td>Metastatic</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>275</td>
<td>33</td>
</tr>
</tbody>
</table>
of the cancer cases (sensitivity) and 281 of 308 (91.2%) of the non-cancer cases (specificity) were correctly classified in the sequential analysis.

Next, for each of the 10 out of 26 correctly classified local cancer cases and the six out of eight correctly classified metastatic cancer cases, the mean (and standard deviation) lead times were determined between the individual’s age at classification and the actual age at the clinical diagnosis of cancer. The estimated mean lead times using the sequential PSA measurements are 6.1 (standard deviation 2.9) years and 7.7 (standard deviation 3.9) years for the local and metastatic cancers respectively.

Finally, for the 301 males with 2–4 PSA measurements or a follow-up time of less than 10 years that were used as a check of the classification method, 246 were diagnosed as normal, 51 as BPH cases and four as local cancer cases. No metastatic cancer cases were diagnosed in this group of men. The overall correctly classified rate was 225/301 (74.8%). Since these 301 males have a median follow-up time of 5.9 years, a more useful value might be the rate correctly classified between cancer and non-cancer cases. This rate is 259/301 (86.0%), a value that is very similar to the corresponding rate (86.8%) obtained from the 342 men in the cross-validation analysis.

4. Discussion

In reviewing statistical methods in diagnosis, Hand (1992) stated that ‘diagnosis is classification’. He went on to state that there are two classes of problem: one where the diagnostic groups are not known \textit{a priori} and the other where the objective is to classify individuals into already known or identified disease states.

In this paper, random-effects models, along with the use of the Bayes theorem to obtain posterior probabilities of prostate disease outcomes, are used to screen males from a long-term longitudinal study for prostate cancer. Independent or predictor variables used in the models include age at first examination and time since the first PSA measurement. Other possible predictors of PSA concentration are diet and family history. These variables were not included in the model because of a lack of availability or insufficient information in our data. For example, it is known that measuring diet accurately is often difficult. In general, the inclusion of important predictors will reduce the random-effects variability which is believed to improve classification results.

The classification rule that is presented in this paper is derived by using the posterior probabilities that were calculated from a linear mixed effects model for modelling the individual rates of change in PSA measurements. A cross-validation study of PSA levels was performed, where diagnostic information was used for all the males except the one being classified, to classify that person successively (one visit at a time) into a non-cancer, local cancer or metastatic cancer group. Overall, 86.8% or 88.3% were correctly classified by using the longitudinally collected PSA measurements, depending on whether or not a distinction is made between local and metastatic cancer. The sensitivity of the method is not particularly high since it correctly classifies about 62% of the true cancers. This is not surprising since about a third of all men measured for PSA with prostate cancer show little increase in PSA levels or have PSA levels that are below the clinical detection cut-off of 4.0 \(\mu\)g l\(^{-1}\) (Carter and Pearson, 1994). There is also evidence of this in our data from Fig. 1. Hence, it is very unlikely that any classification method that is based on either the absolute level or the rate of change in PSA would correctly classify these individuals. Finally, the specificity of the method is high (over 90%), and approximately half of the individuals classified as cancer cases are actually false positive results. In clinical practice, these false positive classifications would lead the physician to investigate further for the presence of cancer by requesting additional
diagnostic tests, such as a prostate biopsy, beyond the usual physical examination and PSA determination.

With regard to the practice of screening for prostate cancer, medical practitioners will collect PSA measurements sequentially or one examination at a time, and so a classification approach as described in the cross-validation study would be quite natural. Zelen and Feinleib (1969) used a stochastic approach to estimate the mean lead time between being in the preclinical disease state and the clinical diagnostic state for individuals participating in an early detection programme. In this study, using the diagnostic data for the 10 of 26 men correctly classified with local cancer and the six of eight men correctly classified with metastatic cancer, the observed mean lead time by using the sequential PSA measurements is 6.1 (standard deviation 2.9) years before the actual diagnosis for local cancer and 7.7 (standard deviation 3.9) years for metastatic cancer. Thus, there appears to be an important window of opportunity for intervention or secondary prevention in the case of prostate cancer.

Similar conclusions were obtained by Morrell et al. (1995) who hypothesized that the longitudinal PSA trends in each cancer case represent a period of slow linear change in PSA followed by a period of rapid exponential increase in peripheral PSA levels. Morrell et al. (1995) used a piecewise mixed effects model with linear and non-linear components and a transition point between the two components to model trends in PSA levels in the local and metastatic cancer groups. They showed that their model provides for different average transition times and levels between the linear and exponential components for the two cancer groups. Morrell et al. (1995) found that there is a difference of 7.7 years before diagnosis in the average estimated transition times between the local and metastatic cancer cases, suggesting an opportunity to detect prostate cancer before it metastasizes.

The linear mixed model (1) specifies the response vector conditionally on a vector $b_i$ of random effects. Classification, however, is based on the marginal distribution, obtained from integrating over the random effects. For linear models, this marginal distribution can be derived analytically. This would no longer be the case for non-linear or generalized linear mixed models. Hence, if the present approach were to be generalized to non-linear models (such as the models proposed by Morrell et al. (1995)), or to classification based on discrete longitudinal profiles, numerical integration methods would be required for the calculation of the posterior probabilities given in equation (2).

Recently many references have considered both longitudinal changes in a variable and the associated effect on the length of time to the occurrence of an event (Altman and De Stavola, 1994; Bycott and Taylor, 1998; De Stavola and Christensen, 1996; Dafni and Tsiatis, 1998; De Gruttola and Tu, 1994; Hogan and Laird, 1997a, b; Tsiatis et al., 1995; Wulfsohn and Tsiatis, 1997; Xu and Zeger, 2001). Some of these references present models that jointly account for the repeated measurements over time and their association with the outcome event. To fit the joint models, the EM algorithm may be used to estimate the parameters from a suitably defined likelihood, or Markov chain Monte Carlo methods may be applied. These methods require extensive time to develop the software routines to fit these models. Furthermore, any model reformulation would require additional time to modify the procedure.

Hogan and Laird (1997a) presented a mixture model for the joint modelling of longitudinal repeated measures and event times, in which the event represents drop-out from the study. A similar approach could be considered for our PSA data. The survival outcome of interest would be the moment that cancer starts to develop in the subject. Obviously, the event time would be censored to the right for all non-cancer cases. However, as no exact time for the onset of cancer is available for the cancer cases, their event times are censored as well, i.e. it is only known to
be before the time that the cancer was detected by actual clinical diagnosis. Also, in this study the prediction of the actual disease status is of primary interest rather than the actual time of onset of cancer. In view of these comments, we believe that applying these methods to our data would yield no additional relevant information.

A considerable number of computations were performed, especially in the cross-validation part of this paper. The use of S-PLUS 2000 for Windows has made this computational burden feasible. In practice, any real screening method to detect preclinical disease will require a large training data set to compute the fixed effects and variance components in the model. Once the fixed effects estimates have been determined from this large data set, each time that a future patient is screened or measured on the screening marker the posterior probabilities can be calculated for this individual with a relatively simple formula. This classification approach offers a practical method to identify preclinical disease and can lead to potential benefits of improved health for the individual and a reduction in treatment for severe cases of illness.

Acknowledgements

The authors thank the Joint Editor and two referees for useful comments that helped to improve this paper.

References


