Dipartimento di Statistica
"Giuseppe Parenti"

Dipartimento di Statistica "G. Parenti" – Viale Morgagni 59 – 50134 Firenze - www.ds.unifi.it

WORKING PAPER 2010/02

Probabilistic classification of age by third molar development: the use of soft-evidence

Vilma Pinchi, Fabio Corradi,
Iljà Barsanti, Stefano Garatti

Università degli Studi
di Firenze

Probabilistic classification of age by third molar development: the use of soft-evidence

Vilma Pinchi *, FabioCorradi **, Iljà Barsanti **, Stefano Garatti

*** Department of Anatomy, Histology and Legal-Medicine, University of Florence, viale Morgagni 85, 50134 Firenze (Italy)**
**** Department of Statistics, University of Florence, viale Morgagni 59, 50134 Firenze (Italy)**

Abstract
The aim of the paper is to classify individuals according to age through dental development of their third molars. Such teeth were classified by Demirjian's 8-stages dental maturity scale, but we introduced a new and relevant variation. In fact the odontologist is allowed to classify a tooth representing the uncertainty about the stage attribution, using the soft-evidence, which is included in the parametric learning. We used a modified Naïve Bayes to classify 559 Italian youths (307 males and 252 females) aged between 16 and 23, according to dichotomous and trichotomous classifications. Results show the importance of the expert's skill in reading the OPG and the ability to express their beliefs about the dental maturity stage.

Keywords: soft-evidence, parametric learning, modified naïve Bayes, age identification, forensic odontology, third molar.

1. Introduction

In recent decades, the immigration of people without regular identification papers has often required age assessment by courts and other public authorities. Age assessment becomes an issue in cases concerning crimes, helping refugees and fulfilling scholastic duties. This need has originated a new, specific and autonomous field in forensic science (1,2).

In most European Union countries and in the United States, the important legal thresholds are 14 and 18 years, whereas in Germany, Spain and extra-European countries, 21 years is also relevant. These legal thresholds motivate the special interest in age assessment concerning young individuals.

Methods using dental mineralization evidence observed through orthopantomographs (OPG), have been proved useful to assess an adolescent's age. Many researchers consider the third molars' development the most effective in detecting ages over 15-16 years since the dentition up to the second molars is almost completed at that time. Nevertheless some problems such as inclusion, malformation or agenesis, make the third molar more problematic than other teeth (3), the special relevance on this form of evidence (4-12) has made the wisdom tooth the referral dental element in discriminating age for 16-22 years old.

There is a consensus about the relevance of gender in dental development(4,6,7,13), as well as other characteristics - like ethnicity (14-17), general health (18-22) and nutrition (23) although the influence of these factors are not definitely ascertained. Also doubts arise about the relevance, in age assessment, in all third molars versus some subset of them - lower or upper arch, right or left part (8)- and the role played by the radiographic technology of the OPGs.

Regarding the statistical approach used, two main alternatives have been proposed in the literature: regression models and polynomial function (24-26), estimating the individual's age based on the dental development and some covariates, and supervised classification methods (27), ascertaining if an individual is younger or older than the age threshold of interest.

Our aim is to propose a probabilistic model dealing with forensic age assessment, using the third molars according to the Demirjian's classification (28) and assessing the probability of the individual's age as over or below a specific threshold. For this reason, we concentrated on the class of the supervised classification models, and decided to use a *Bayesian Naive Classifier* , or simply *Naïve Bayes* (29,30), modified to cope with uncertain observations.

Uncertain observations arise when the observers find difficulty in classifying the third molars in only one of the stages of the Demirjian's scale. For this reason, we allowed the observers to specify more than one state with an associated degree of uncertainty. This form of observation is called *soft-evidence* and has been introduced in the statistical literature by Pearl (31) and employed in Bayesian networks by Bilmes (32). Shapiro also considered this feature for medical diagnosis (33).

After an attempt to classify individuals as adults or minors, we introduced a non-decision class including the age threshold of 18 years. Taking into account the trade-off between the reduction of misclassification errors and the information loss for classified individuals, we propose to evaluate the models performance and consider the OPG's technology.

Finally, because of the uncertain nature of the observations, we expected discrepancies between the evaluations of the different experts and between the evaluations provided on the same OPGs by the same observer at different times. Concerning this issue we proposed to evaluate the inter-observers and intra-observer reproducibility.

2. Material and methods

To estimate the model we employed a sample of 559 Caucasians Italian youths (307 males and 252 females) aged between 16 and 22 years. Differently to other Authors (34,35) we prefer to consider all available third molars (upper and lower) and to stage them by Demirjian's scale. In fact Demirjian's classification does not require the evaluation of a length fraction of an incomplete root, that implies a sort of prediction of its final length and it was criticized in some previous studies (36). Ages of the subjects were recorded based on their birth year. The OPGs were analogical or digital. The analogical OPGs were obtained by direct exposure from an X-ray photographic film and scanned in a jpg file (200 dpi) by an EPSON Expression 1680 Pro scanner. Digital OPGs were obtained by an electronic method called CDD acquisition, directly exporting files in jpg format from a radiographic system. All the resulting files were stored in a restricted area of a web site and available for the experts' evaluation.

Two experts provided the evaluations on all of the third molars' developments for each OPG, working independently and never exchanging information; they knew only the gender of the subjects. The experts are dentists of different experience and training: *A* is an experienced forensic odontologist while *B* has extensive clinical experience, but little forensic training. *A* and *B* also provided evaluations twice on a randomly drawn sub-sample of 77 OPGs to measure their intra-observers reproducibility.

3. The model

3.1. Variables and notations

Let $T = t$ represent the variable indicating the age of an individual and $\{\tau_1, \tau_2, ..., \tau_{Q-1}\} \subset \mathbb{R}$ the set of age thresholds of possible interest. The elements of the set define the states of the random *Class variable* $C = \{c_1, c_2, ..., c_Q\}$, where $c_1 = \{t : t < \tau_1\}$, $c_2 = \{t : \tau_1 \leq t < \tau_2\}$, …, $c_Q = \{t : t \geq \tau_{Q-1}\}$.

Let $\mathbf{S}$ represent a set of opportunely chosen covariates and let N attributes $\mathbf{X}_h$, with $1 \le h \le N$, represent the dental developments of the h^{th} third molars defined accordingly to the Demirjian's scale. So that $\mathbf{X}_h = (X_{h1}, X_{h2}, ..., X_{hk})$ is a multinomial r. v. on a single observation, $Mu_k\left(\mathbf{x}_h \mid \boldsymbol{\theta}_h, 1\right)$, where:

$$X_{h,j} = \begin{cases} 1 \text{ if the state } j \text{ occurred in the } h^{th} \text{ third molar} \\ 0 \text{ otherwise} \end{cases}.$$

The parametric vector $\boldsymbol{\theta}_h = (\theta_{h1}, \theta_{h2}, ..., \theta_{hk})$, with $0 \le \theta_{hj} \le 1 \quad \forall h, j$ and $\sum_{j=1}^{k} \theta_{hj} = 1$, represents the probabilities associated with each state of Demirjian's scale of the h^{th} third molar and it is assumed to have a Dirichlet prior distribution, $Dir_k(\boldsymbol{\theta}_h \mid \boldsymbol{\alpha}_h)$, with known hyperparameters $\boldsymbol{\alpha}_h = (\alpha_{h1}, \alpha_{h2}, ..., \alpha_{hk})$. To simplify the notation for $\boldsymbol{\theta}_h \mid \boldsymbol{\alpha}_h$ we hereafter omit the hyperparameters $\boldsymbol{\alpha}_h$.

3.2. Conditional independence assumptions

Consider the class variable *C*, the four third molars joint distribution $\mathbf{X} = (\mathbf{X}_1, ..., \mathbf{X}_N)$ and the set of covariates with $\mathbf{S}$ representing the main influential variables on $\mathbf{X}_h$. Let $\boldsymbol{\Theta} = (\boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_N)$, then a natural factorization of $P(\mathbf{X}, C, \mathbf{S} \mid \boldsymbol{\Theta})$ is:

(i) $\quad P(\mathbf{X}, C, \mathbf{S} \mid \boldsymbol{\Theta}) = P(\mathbf{X} \mid C, \mathbf{S}, \boldsymbol{\Theta}) P(C, \mathbf{S})$.

Now assume that, conditionally on the class variable *C* and on the set of covariates $\mathbf{S}$, each tooth grows independently with respect to the others, $\mathbf{X}_i \perp \mathbf{X}_j \mid C, \mathbf{S}, \boldsymbol{\Theta}$. So that:

(ii) $\quad P(\mathbf{X} \mid C, \mathbf{S}, \boldsymbol{\Theta}) = \prod_{h=1}^{N} P(\mathbf{X}_h \mid C, \mathbf{S}, \boldsymbol{\theta}_h)$,

The assumption (ii) allows to estimate $\boldsymbol{\theta}_h$ separately for each tooth, conditionally to each class of age and to each possible configuration of the covariates in $\mathbf{S}$. This model, a *Naïve Bayesian Classifier* (29,30), leads to the factorization $P(\mathbf{X}, C, \mathbf{S} \mid \boldsymbol{\Theta}) = \prod_{h=1}^{N} P(\mathbf{X}_h \mid C, \mathbf{S}, \boldsymbol{\theta}_h) P(C, \mathbf{S})$.

We also allowed independence among the prior distributions of parameters $\boldsymbol{\theta}_{h|qs}$ so that:

(iii) $\quad f(\boldsymbol{\theta}_{1|qs}, ..., \boldsymbol{\theta}_{N|qs}) = \prod_{h=1}^{N} f(\boldsymbol{\theta}_{h|qs})$.

To simplify the notation let $\mathbf{X}_{h|qs} \mid \boldsymbol{\theta}_{h|qs}$ indicate $\mathbf{X}_h \mid C, \mathbf{S}, \boldsymbol{\theta}_h$ with a Multinomial probability density $Mu_k\left(\boldsymbol{\theta}_{h|qs}, 1\right)$, where $\boldsymbol{\theta}_{h|qs}$ represents the vector of probabilities for the h^{th} third molar conditionally to $C = c_q$ and to $\mathbf{S} = \mathbf{s}$, one of the possible joint realizations of the covariates in $\mathbf{S}$.

3.3. The soft-evidence

Since $\mathbf{X}_h$ is often observable with uncertainty we introduced a further random variable $\mathbf{E}_h = (E_{h1}, E_{h2}, ..., E_{hk})$ with $1 \le \sum_{j=1}^{k} E_{hj} \le k$, indicating which development stages of the h^{th} third molar possibly occurred. So that:

$$E_{hj} = \begin{cases} 1 \text{ if the state } j \text{ possibly occurred} \\ 0 \text{ otherwise} \end{cases}$$

The observer provides the evaluation of the h^{th} third molar's development stages and makes use of the vector of *beliefs*, $\mathbf{b}_h = (b_{h1}, b_{h2}, ..., b_{hk})$, with $0 \le b_{hj} \le 1$ and $\sum_{j=1}^{k} b_{hj} = 1$ to weight them.

If the observer indicates only one state, we have got *hard evidence* on j^*, with $b_{hj^*} = 1$ and $b_{hj} = 0 \;\; \forall j \ne j^*$, which is equivalent to a direct observation on $\mathbf{X}_h$. Otherwise, if more than one state has a belief $0 < b_{hj} < 1$ *soft evidence* has occurred. The case of maximum uncertainty on the observation is when $b_{hj} = 1/k$ per $\forall j$. This possibility is used to cope with *missing data*. In these circumstances, missing data is considered not informative about the class variable, i.e. they are assumed *missing completely at random* (*MCAR).*

The probabilistic connection between the variables $\mathbf{X}_h$ and $\mathbf{E}_h$ is given by:

$$P(\mathbf{X}_{h|qs}, \mathbf{E}_{h|qs} \mid \boldsymbol{\theta}_{h|qs}, \mathbf{b}_{h|qs}) = P(\mathbf{X}_{h|qs} \mid \mathbf{E}_{h|qs}, \boldsymbol{\theta}_{h|qs}) P(\mathbf{E}_{h|qs} \mid \mathbf{b}_{h|qs}) \propto \mathbf{b}_{h|qs} \boldsymbol{\theta}_{h|qs}, \quad (1)$$

where $P(\mathbf{X}_h \mid \mathbf{E}_h, \boldsymbol{\theta}_h) \propto \boldsymbol{\theta}_h \mathbf{E}_h$.

Finally we extend the assumption (*ii*) to include the soft evidence, i.e. $\forall i \ne j$ $\left(\mathbf{X}_i, \mathbf{E}_i\right) \perp \left(\mathbf{X}_j, \mathbf{E}_j\right) \mid C, \mathbf{S}, \boldsymbol{\Theta}, \mathbf{B}$, where $\mathbf{B} = (\mathbf{b}_1, ..., \mathbf{b}_N)$. So that:

(iv) $$P(\mathbf{X}, \mathbf{E} \mid C, \mathbf{S}, \boldsymbol{\Theta}, \mathbf{B}) = \prod_{h=1}^{N} P(\mathbf{X}_h, \mathbf{E}_h \mid C, \mathbf{S}, \boldsymbol{\theta}_h, \mathbf{b}_h).$$

4. Parametric learning

The likelihood function for $\boldsymbol{\theta}_{h|qs}$ based on the i^{th} observation beliefs $\mathbf{b}_{i,h|qs} = \left(b_{i,h1|qs}, b_{i,h2|qs}, ..., b_{i,hk|qs}\right)$ is the marginalization of $P(\mathbf{X}_{h|qs}, \mathbf{E}_{h|qs} \mid \boldsymbol{\theta}_{h|qs}, \mathbf{b}_{h|qs})$ with respect to the random variable $\mathbf{X}_{h|qs}$, hereafter considered unobserved:

$$L(\boldsymbol{\theta}_{h|qs}; \mathbf{b}_{i,h|qs}) = \sum_{j=1}^{k} P(X_{hj|qs}, \mathbf{E}_{h|qs} \mid \boldsymbol{\theta}_{h|qs}, \mathbf{b}_{i,h|qs}) = \sum_{j=1}^{k} b_{i,hj|qs} \theta_{hj|qs}. \quad (2)$$

For *n* conditional independent observations, the likelihood function is:

$$L(\boldsymbol{\theta}_{h|qs};\mathbf{B}_{h|qs}) = \prod_{i=1}^{n_{qs}} L(\boldsymbol{\theta}_{h|qs};\mathbf{b}_{i,h|qs}) = \prod_{i=1}^{n_{qs}} \sum_{j=1}^{k} b_{i,hj|qs}\theta_{hj|qs} \, . \tag{3}$$

A more compact representation of (3), achievable by grouping all of the likely polynomials with equal base and by adding the corresponding coefficients, is:

$$L(\boldsymbol{\theta}_{h|qs};\mathbf{B}_{h|qs}) = \sum_{m=1}^{\tilde{m}_{h|qs}} \tilde{b}_{m,h|qs} \prod_{j=1}^{k} \theta_{hj|qs}^{p_{m,hj|qs}} \, , \tag{4}$$

where $\tilde{m}_{h|qs}$ represents the number of polynomials with a different base, $\mathbf{P}_{h|qs} = [p_{m,hj|qs}]_{\tilde{m}_{h|qs} x k}$ is the matrix whose rows contain the powers of the parameters $\theta_{hj|qs}$ for each of the $\tilde{m}_{h|qs}$ polynomials and $\tilde{b}_{m,h|qs}$ is the coefficient associated to each of them.

Let $\mathbf{B}_{qs} = [\mathbf{b}_{1|qs},...,\mathbf{b}_{N|qs}]$ stand for the vector of conditional beliefs for all third molars, then from the assumption (iv) we factorize the likelihood function as:

$$L(\boldsymbol{\Theta}_{qs};\mathbf{B}_{qs}) = \prod_{h=1}^{N} L(\boldsymbol{\theta}_{h|qs};\mathbf{B}_{h|qs}) \, , \tag{5}$$

so that the posterior probability density function of $\boldsymbol{\Theta}_{qs} \,|\, \mathbf{B}_{qs}$ can be written as:

$$f(\boldsymbol{\Theta}_{qs} \,|\, \mathbf{B}_{qs}) = \frac{L(\boldsymbol{\Theta}_{qs};\mathbf{B}_{qs})\, f(\boldsymbol{\Theta}_{qs})}{\int_{\boldsymbol{\Theta}_{qs}} L(\boldsymbol{\Theta}_{qs};\mathbf{B}_{qs})\, f(\boldsymbol{\Theta}_{qs}) d\boldsymbol{\Theta}_{qs}} = \frac{\prod_{h=1}^{N} L(\boldsymbol{\theta}_{h|qs};\mathbf{B}_{h|qs}) \prod_{h=1}^{N} f(\boldsymbol{\theta}_{h|qs})}{\int_{\boldsymbol{\theta}_{1|qs}} ... \int_{\boldsymbol{\theta}_{N|qs}} \prod_{h=1}^{N} L(\boldsymbol{\theta}_{h|qs};\mathbf{B}_{h|qs}) \prod_{h=1}^{N} f(\boldsymbol{\theta}_{h|qs}) d\boldsymbol{\theta}_{1|qs} ... d\boldsymbol{\theta}_{N|qs}} =$$

$$= \prod_{h=1}^{N} \frac{L(\boldsymbol{\theta}_{h|qs};\mathbf{B}_{h|qs})\, f(\boldsymbol{\theta}_{h|qs})}{\int_{\boldsymbol{\theta}_{h|qs}} L(\boldsymbol{\theta}_{h|qs};\mathbf{B}_{h|qs})\, f(\boldsymbol{\theta}_{h|qs}) d\boldsymbol{\theta}_{h|qs}} = \prod_{h=1}^{N} f(\boldsymbol{\theta}_{h|qs} \,|\, \mathbf{B}_{h|qs}) \, . \tag{6}$$

This implies the possibility to work separately on each posterior probability density of $\boldsymbol{\theta}_{h|qs} \,|\, \mathbf{B}_{h|qs}$, being:

$$f(\boldsymbol{\theta}_{h|qs} \,|\, \mathbf{B}_{h|qs}) = \frac{L(\boldsymbol{\theta}_{h|qs};\mathbf{B}_{h|qs})\, f(\boldsymbol{\theta}_{h|qs})}{\int_{\boldsymbol{\theta}_{h|qs}} L(\boldsymbol{\theta}_{h|qs};\mathbf{B}_{h|qs})\, f(\boldsymbol{\theta}_{h|qs}) d\boldsymbol{\theta}_{h|qs}} \, . \tag{7}$$

Since the likelihood (4) is a mixture of a multinomial r. v. and a Dirichlet prior density $f(\boldsymbol{\theta}_{h|qs})$, the posterior density (7) is a mixture of $\tilde{m}_{h|qs}$ Dirichlet r. v. Each mixture component is a random variable $Dir_{\tilde{m}_{h|qs}}(\boldsymbol{\theta}_{m,h|qs} \,|\, \boldsymbol{\alpha}_{h|qs} + \mathbf{p}_{m,h|qs})$, whose vector of hyperparameters, is determined by the $\boldsymbol{\alpha}_{h|qs}$ and the correspondent row of the matrix $\mathbf{P}_{h|qs}$. More explicitly:

$$f(\boldsymbol{\theta}_{h|qs} \,|\, \boldsymbol{\alpha}_{h|qs}, \mathbf{B}_{h|qs}) = \sum_{m=1}^{\tilde{m}_{h|qs}} q_{m,h|qs} Dir_{\tilde{m}_{h|qs}}(\boldsymbol{\theta}_{m,h|qs} \,|\, \boldsymbol{\alpha}_{h|qs} + \mathbf{p}_{m,h|qs}) \, . \tag{8}$$

where the weights $q_{m,h|qs}$ are obtained, considering the normalization constant in (7), so that:

$$q_{m,h|qs} = \frac{\tilde{b}_{m,h|qs} \prod_{j=1}^{k} \Gamma(\alpha_{hj|qs} + p_{m,hj|qs})}{\sum_{m=1}^{\tilde{m}_{h|qs}} \tilde{b}_{m,h|qs} \prod_{j=1}^{k} \Gamma(\alpha_{hj|qs} + p_{m,hj|qs})} . \tag{9}$$

5. Prediction

5.1. Hard evidence

Let $\mathbf{X}^{n+1} = \mathbf{x}^{n+1}$ and $\mathbf{S}^{n+1} = \mathbf{s}^{n+1}$ the attributes and covariates observed on the $(n+1)^{th}$ subject and C^{n+1} the correspondent unobserved class variable. Let $\mathbf{X}_{h|qs}^{n+1}$ indicate the r. v. $\mathbf{X}_h^{n+1} \,|\, c_q^{n+1}, \mathbf{s}^{n+1}$, then the predictive probability for $C^{n+1} = c_q^{n+1}$ is:

$$P(c_q^{n+1} \,|\, \mathbf{x}^{n+1}, \mathbf{s}^{n+1}) = \frac{P(\mathbf{x}^{n+1} \,|\, c_q^{n+1}, \mathbf{s}^{n+1}) P(c_q^{n+1} \,|\, \mathbf{s}^{n+1})}{P(\mathbf{x}^{n+1} \,|\, \mathbf{s}^{n+1})} \propto P(c_q^{n+1} \,|\, \mathbf{s}^{n+1}) \prod_{h=1}^{N} P(\mathbf{x}_{h|qs}^{n+1}) , \tag{10}$$

where $P(\mathbf{x}_{h|qs}^{n+1})$ is the probability of the realization $\mathbf{x}_{h|qs}^{n+1}$ of a multinomial r. v. $Mu_k \left(\mathbf{x}_{h|qs}^{n+1} \,|\, \boldsymbol{\theta}_{h|qs}, 1 \right)$.

Since the parametric vector $\boldsymbol{\theta}_{h|qs}$ is unknown, first we derive the distribution of the r. v. $\mathbf{X}_{h|qs}^{n+1}$, using the results (8) and marginalizing respect to $\boldsymbol{\theta}_{h|qs}$:

$$P(\mathbf{X}_{h|qs}^{n+1} \,|\, \boldsymbol{\alpha}_{h|qs}, \mathbf{B}_{qs}) = \int_{\boldsymbol{\theta}_{h|qs}} P(\mathbf{X}_{h|qs}^{n+1} \,|\, \boldsymbol{\theta}_{h|qs}) f(\boldsymbol{\theta}_{h|qs} \,|\, \boldsymbol{\alpha}_{h|qs}, \mathbf{B}_{qs}) d\boldsymbol{\theta}_{h|qs} =$$

$$= \int_{\boldsymbol{\theta}_{h|qs}} Mu_k \left(\mathbf{x}_{h|qs}^{n+1} \,|\, \boldsymbol{\theta}_{h|qs}, 1 \right) \cdot \left(\sum_{m=1}^{\tilde{m}_{h|qs}} q_{m,h|qs} Dir_{\tilde{m}_{h|qs}} (\boldsymbol{\theta}_{m,h|qs} \,|\, \boldsymbol{\alpha}_{h|qs} + \mathbf{p}_{m,h|qs}) \right) d\boldsymbol{\theta}_{h|qs} =$$

$$= \sum_{m=1}^{\tilde{m}_{h|qs}} q_{m,h|qs} Md_k (\mathbf{x}_{h|qs}^{n+1} \,|\, \boldsymbol{\alpha}_{h|qs} + \mathbf{p}_{m,h|qs}, 1) , \tag{11}$$

The vector (11) represents the probability to observe the states of the h^{th} attribute for the $(n+1)^{th}$ individual, conditionally to the q^{th} class of age and s^{th} joint realization of the covariates in $\mathbf{S}$ and it corresponds to a weighted mean of probabilities determined by $\tilde{m}_{h|qs}$ Multinomial-Dirichlet r. v. $Md_k (\mathbf{x}_{h|qs}^{n+1} \,|\, \boldsymbol{\alpha}_{h|qs} + \mathbf{p}_{m,h|qs}, 1)$.

Then the (10) becomes:

$$P(c_q^{n+1} \,|\, \mathbf{x}^{n+1}, \mathbf{s}^{n+1}, \boldsymbol{\alpha}_{h|qs}, \mathbf{B}_{qs}) \propto P(c_q^{n+1} \,|\, \mathbf{s}^{n+1}) \prod_{h=1}^{N} P(\mathbf{x}_{h|qs}^{n+1} \,|\, \boldsymbol{\alpha}_{h|qs}, \mathbf{B}_{qs}) , \tag{12}$$

where the vector of probabilities $P(\mathbf{x}_{h|qs}^{n+1} \,|\, \boldsymbol{\alpha}_{h|qs}, \mathbf{B}_{qs})$ in (10) is easily derived:

$$P(\mathbf{x}_{h|qs}^{n+1} \,|\, \boldsymbol{\alpha}_{h|qs}, \mathbf{B}_{qs}) = \sum_{m=1}^{\tilde{m}_{h|qs}} q_{m,h|qs} Md_k (\mathbf{x}_{h|qs}^{n+1} \,|\, \boldsymbol{\alpha}_{h|qs} + \mathbf{p}_{m,h|qs}, 1) = \frac{\boldsymbol{\alpha}_{h|qs} + \sum_{m=1}^{\tilde{m}_{h|qs}} q_{m,h|qs} \mathbf{p}_{m,h|qs}}{\alpha_{0,h|qs} + n} \tag{13}$$

where $\alpha_{0,h|qs} = \sum_{j=1}^{k} \alpha_{hj|qs}$.

As it concerns the vector of prior conditional probabilities $P(c_q^{n+1} \mid \mathbf{s}^{n+1})$, it can be possibly estimated as proportion of n_{qs}, numbers of individuals in the age class c_q^{n+1} with covariates $\mathbf{s}^{n+1}$, observed on the n_s observations:

$$\hat{P}(c_q^{n+1} \mid \mathbf{s}^{n+1}) = \frac{n_{qs}}{n_s} . \qquad (14)$$

5.2. Soft evidence

Now consider the variable $\mathbf{X}_{h|qs}^{n+1}$ is observed with uncertainty so that we make use of the expert's beliefs $\mathbf{b}_h^{n+1}$. This does not allow the use of predictive probability (13) but we must consider the linear combination $\sum_{j=1}^{k} b_{hj}^{n+1} P(x_{hj|qs}^{n+1} \mid \boldsymbol{\alpha}_{h|qs}, \mathbf{B}_{qs})$.

So that, the probability the $(n+1)^{th}$ individual belongs to the class of age $c_{q^*}^{n+1}$, given the expert's beliefs on the dental developments $\mathbf{b}_h^{n+1}$, the observed set of covariates $\mathbf{s}^{n+1}$ and the conditional beliefs $\mathbf{B}_{qs} = [\mathbf{b}_{1|qs}, ..., \mathbf{b}_{N|qs}]$ of the training data set, is:

$$\hat{P}(c_q^{n+1} \mid \mathbf{b}^{n+1}, \mathbf{s}^{n+1}, \boldsymbol{\alpha}_{h|qs}, \mathbf{B}_{qs}) \propto n_{qs} \prod_{h=1}^{N} \sum_{j=1}^{k} b_{hj}^{n+1} \frac{\alpha_{hj|qs} + \sum_{m=1}^{\tilde{m}_{h|qs}} q_{m,hj|qs}\, p_{m,hj|qs}}{\alpha_{0,h|qs} + n} . \qquad (15)$$

6. Classification

To determine the individual's age we need to explicit a classification rule *f*. A possibility is classifying the $(n+1)^{th}$ subject into the class of age $c_{q^*}^{n+1}$ with highest estimated probability, i.e.:

$$\hat{c}^{n+1} = f(n+1) = \underset{q}{\operatorname{argmax}}\, P(c_q^{n+1} \mid \mathbf{b}^{n+1}, \mathbf{s}^{n+1}, \boldsymbol{\alpha}_{h|qs}, \mathbf{B}_{qs}) \qquad (16)$$

Alternatively, introducing a probabilistic classification threshold π, we can use a more refined decision rule, $f_{2,\pi}$ or $f_{3,\pi}$ accordingly, as we consider respectively two or three classes of age. Let $P_q = P(c_q^{n+1} \mid \mathbf{b}^{n+1}, \mathbf{s}^{n+1}, \boldsymbol{\alpha}_{h|qs}, \mathbf{B}_{qs})$:

$$f_{2,\pi}(n+1)=\begin{cases} c_2 & \text{if } \underset{q}{\operatorname{argmax}} P_q = 2 \text{ and } P_2 \geq \pi \\ c_1 & \text{otherwise} \end{cases} \quad (17)$$

$$f_{3,\pi}(n+1)=\begin{cases} c_3 & \text{if } \underset{q}{\operatorname{argmax}} P_q = 3 \text{ and } P_3 \geq \pi \\ \mathit{unclassified} & \text{otherwise} \end{cases} \quad (18)$$

7. Expert's reproducibility

A measure of the *inter-observers* and *intra-observes reproducibility* we propose is:

$$\mu_h^{t_1,t_2} = 1 - \frac{\sum_{i=1}^{n}\sum_{j=1}^{k}\left|b_{i,hj}^{t_1} - b_{i,hj}^{t_2}\right|}{2n} \quad (18)$$

regarding the evaluations of the h^{th} tooth and where the beliefs $b_{i,hj}^{t_1}$ and $b_{i,hj}^{t_2}$ correspond to the j^{th} state of the h^{th} teeth for the i^{th} observation, with t_1 and t_2 represent the times the same expert evaluates the same OPG (intra-observer reproducibility) or label the experts themselves (inter-observers reproducibility).

8. The application

In the set of covariates **S** we included only the variable *Gender* while the variable *Technology* was employed only to compare the experts' performances according to the OPG's technology.

First, we produced a dichotomous analysis with only one age threshold, $\tau_1 = 18$ years. This defined two classes of age, $c_1 = \{t : t < 18\}$ and $c_2 = \{t : t \geq 18\}$. Then we considered two age thresholds, $\tau_1 = 17$ and $\tau_2 = 19$ years, to originate three meaningful classes of age: $c_1 = \{c : c < 17\}$, $c_2 = \{c : 17 \leq c < 19\}$ and $c_3 = \{c : c \geq 19\}$.

The learning procedure was based on a *training data set* of 447 observations, randomly chosen among the 559 observations and stratified for age, gender, expert's expertise and evidence type (hard, soft or missing). From this data set we inferred on the $\theta_{h,qs}$'s distributions (8), setting $\boldsymbol{\alpha}_{h|qs} = \mathbf{1}$, and we produced the classification probabilities (15) on the remaining 112 observations, composing the *test data set*. Finally, we classified the subjects making use of the classification rule (17) in the dichotomous case or the classifications rules (16) and (18) in the trichotomous case. All the results shown hereafter are obtained by averaging 1,000 replications of this procedure, differing for the training and test data sets randomly drawn .

9. Results

9.1. Dichotomous analysis

The dichotomous analysis's aim is to evaluate the probability that a subject is an adult or a minor, conditionally on the gender and the third molars' evaluation provided by the expert. In *Tab. 1* we show, for Expert A, who leads to better findings respect to Expert B, the percentages of individuals correctly classified by different combination of teeth[1] according to the decision rule (17) for some probabilistic classification thresholds π.

Table 1 about here

Tab. 1 shows how the higher the probabilistic classification threshold, the lower the percentage of correctly classified individuals will be. This is especially true if a reduced amount of evidence is employed and the probabilistic classification thresholds is low. For instance, if only a single tooth is considered, not one of the adults overcome the threshold equaling to 0,99 so that the percentage of correctly classified individuals is only 27,7%, corresponding to the percentage of minors in the training data set. The benefit of considering all third molars is clear if we use a probabilistic classification threshold in the range $\pi \geq 0,80$.

Tab. 2 shows the percentages of misclassified minors based on the combination of teeth and probabilistic classification threshold:

Table 2 about here

Looking at *Tab. 2*, if the probabilistic threshold π increases then the percentage of misclassified minors decreases. Once again, the benefit of considering all third molars is clear if we use a probabilistic classification threshold but in the complementary range, $\pi \leq 0,80$.

All these findings suggest that the trade-off between the percentage of correct classification and the percentage of misclassified minors strongly depend on the probabilistic classification threshold π. Combining the results of the *Tab. 1* and *Tab. 2,* the four third molars evidence leads to the best performances if we use a probabilistic classification threshold equaling to 0,80. However, these elements are not completely satisfactory since a high percentage of individuals are misclassified so, we explored the data and introduced a third class of age.

9.2. Trichotomous analysis

[1] The notation T_1, T_2, T_3, T_4 correspond, in odontology, respectively to the third molars T_{18}, T_{28}, T_{38}, T_{48}

The change of status from minors to adults obviously does not correspond to an immediate change in dental development. This consideration raises doubts about the opportunity to use the dichotomous analysis because of the difficulties to discriminate individuals aged around the threshold of 18 years.

A more sensible approach consists in the introduction of a *not-decision* class surrounding that threshold, presumably, including many individuals misclassified by the dichotomous analysis.

In *Tab. 3* and *4* we show how individuals aged less than 17 years have been classified, by expert and technology. An individual is classified according to the highest predictive probability, making use of the classification rule (16):

Table 3 about here

Table 4 about here

Averaging for technology, Expert *A* (*Tab. 3*) assessed correctly 79,1% of individuals aged <17 years, obtaining a 81,3% of the correct classification if the considered OPGs are analogical. Expert B provided slightly worse results: 75,0% were correctly assessed without taking technology into account and only a slight improvement if digital OPGs were employed. The main results of the analysis are represented by the percentage of misclassification. Averaging for technology, false adults were 2,5% and 6,9% for the experts *A* and *B*, respectively, and even better results are obtained if the experts are considered for their preferred technologies.

In *Tab. 5* and *6* we show how individuals aged $\geq$ 19 years have been classified.

Table 5 about here

Table 6 about here

In this case, the performances of the model estimated on the experts *A* and *B* are very similar: 74,3% and 74,5% of the cases were correctly classified with some improvement if the experts are allowed to choose the OPG’s technology. Individuals erroneously classified as minors occurred in 7,2% of the cases for expert *A* and in 7,9% for expert *B*. All of these findings support the trichotomous analysis for minor’s age assessment by using dental evidence deriving from all four third molars.

9.3. Percentage of false adults versus percentage of non-classified individuals

In the trichotomous analysis we can employ the classification rule (18), introducing a probabilistic classification threshold, which can be diversely specified according to the case and to the judge. This means that, if the predictive probability for an individual does not overcome the threshold, then the individual is not classifiable.

In *Fig.1* and *Fig.2* model performances are illustrated for Expert A and B. First, we show the relation between the percentage of non-classified individuals and the probabilistic classification threshold (*Fig.1*), then, for classified individuals, we show the false adults percentage (*Fig.2*).

Figure 1 about here

Figure 2 about here

As it was expected, there is a direct, almost linear, relationship between the classification threshold and the percentage of non-classified individuals. If all of the thresholds are equal, the model based on Expert A’s evaluation provides better performances than those deriving from Expert B. Based on the percentage of false adults produced, Expert A’s performances are not very much affected by the classification probabilistic threshold, unlike in Expert B’s performances. More specifically, Expert *A* can choose a classification threshold that equals 0,50 with a very small percentage of non-classified individuals (2,96%), among which only 5‰ are false adults. To produce the same result, Expert B needs to make use of a threshold that equals 0,85 which obviously implies a higher percentage of non-classified cases, around 60%.

9.4. Evaluations’ reproducibility

Concerning the reproducibility of the experts’ evaluations, we evaluated the inter-observers reproducibility index (19) for each third molar and technology as it is shown in *Tab.7*:

Table 7 about here

The divergences among Expert *A* and *B* shown in *Tab.7* justifies the differences in their models’ performances.

To assess the intra-observer reproducibility we drew two different samples, of 77 OPGs (44 analogical and 33 digital) each per expert stratified by gender and age class.. Then we compared the new dental evaluations with the previous correspondent by means of the index (19):

Table 8 about here

Table 9 about here

Expert *A*, who has a forensic background and uses soft-evidence extensively, has a higher reproducibility compared to the clinician expert *B*, who uses mostly hard-evidence. It is also clear that using the preferred technology can improve the reproducibility for both experts. Furthermore they obtain better performances on the lower teeth (T_{12}) than when the upper teeth were used.(T_{34}).

In conclusion, building a model based on expert evaluations, we should include a source of uncertainty concerning intra-observers reproducibility. The higher the variability of the same observations the expert provides, the greater the uncertainty on the predictive probability. Hence, it is convenient that each expert uses his or her preferred radiographic technology (analogical or digital) allowing the highest reproducibility.

10. Conclusions

In this paper, we showed a methodology to deal with the age assessment of non adult individuals making use of dental evidence for forensic purposes.

The interest is on a specified age threshold, in this case 18 years old, since this age causes different application of laws, grants and other social interventions. Considering the importance of predicting the individual's age it is more appropriate to make use of a classifier instead of a regression model. More specifically, we chose a Bayesian naive classifier, opportunely modified to cope with soft evidence and missing data and, despite its simplicity, produced verifiable and satisfactory results.

According to the literature on dental evidence, we concentrated on observing the third molars for the 16-22 age ranges since they still exhibit an appreciable growth. Observations were provided with respect to Demirjian classification scale, one of the most reliable and widespread dental classification methods.

Our main contribution takes into consideration that the observers are often unable to classify a tooth in only one of the eight Demirjian stages. Therefore, they were allowed to make use of soft evidence providing the opportunity to classify in more than one state and with an associated belief. In addition to differing skill levels and experience of the experts, providing the choice of the preferred technology can improve the model performance.

A second important feature of the method is to take into consideration the conflict between continuous dental development, and the arbitrary age thresholds set. To cope with this problem we introduced a third, intermediate non-decision age class including the threshold itself. This new age class obviously produces a reduction of misclassified individuals but, also, decreases the percentage of classifiable subjects.

Based on these ideas and results concurred from the articulated experiment, the experts hold a crucial role specifically concerning their skill in reading the OPG probabilistically.

In the experiment we noticed appreciable differences in the experts' performances and consequently on the value of the inter-observers reproducibility. Furthermore, Expert A who is more experienced in forensic field, has provided the best results and has showed a large coherence in the evidence evaluations (37).

The results clearly show when soft evidence arises, it is fundamental to assess the experts abilities in advance. Future studies will involve extending the analysis and including less homogenous individuals in a training set to take into account additional covariates. More importantly, we would like to produce more detailed age data, and to propose an optimized non-decision age class, to act as a compromise between the percentage of false adults and the percentage of unclassified cases.

References

1. Schmeling A., Reisinger W., Geserick G., Olze A, Age estimation of unaccompanied minors. Part I. General considerations, Forensic Sci. Int. 2006; **159**: S61–4.
2. Olze A, Reisinger W, Geserick G, Schmeling A. Age estimation of unaccompanied minors. Part II. Dental aspects.Forensic Sci Int. 2006 May 15;159 ,S65-7.
3. Friedrich RE, Ulbricht C, von Maydell LA, Scheuer HA, ,The impact of wisdom teeth topography on chronology of root formation--forensic consequence for forensic-odontologic age estimation of adolescents and young adults. Radiographic investigations using orthopantomography, Arch Kriminol. 2005;216(1-2):15-35
4. Gunst K., Mesotten K, Carbonez A., Willems G. Third molar root development in relation to chronological age: a large sample sized retrospective study, Forensic Sci. Int. 2003; 136: 52–7.
5. Harris MJP., Nortje CJ, The mesial root of the third mandibular molar. A possible indicator of age. J Forensic Odontostomatol, 1984; 2:39-43
6. Mincer H.H., Harris E.F., Berryman H.E., The A.B.F.O. study third molar development and its use an estimator of chronological age, J. Forensic Sci. 1993; 38 : 379–90.
7. Pinchi V, Franchi E, Manetti G, Norelli G-A. La maturazione radiografica dell'ottavo nello studio dell'età del vivente. Zacchia, 2000; 18:261- 278
8. Willershausen B., Loffler N., Schulze R., Analysis of 1202 orthopantomograms to evaluate the potential of forensic age determination based on third molar developmental stages. Eur. J. Med. Res. 2001; 6: 377–84.
9. Solari A.C., Abramovitch K., The accuracy and precision of third mandibular molar development as an indicator of chronological age in Hispanics, J. Forensic Sci. 2002, 47: 531–5.
10. Bolaños M.V., Moussa H., Manrique M.C, Bolaños, Radiographic evaluation of third molar development in Spanish children and young people, Forensic Sci. Int.2003, 133: 212

11. Orhan K, Ozer L, Orhan AI, Dogan S, Paksoy CS. Radiographic evaluation of third molar development in relation to chronological age among Turkish children and youth.Forensic Sci Int. 2007;165(1):46-51.
12. Liversidge HM. Timing of human mandibular third molar formation, Ann Hum Biol. 2008 May-Jun;35(3):294-321.
13. Lavesque G.Y., Demirjian R., Tanguay R., Sexual dimorphism in the development, emergence and agenesis of the mandibular third molar,J. Dent. Res., 1981, 60 , 1735-1741
14. Maber M., Liversidge H.M., Hector M.P., Accuracy of age estimation of radiographic methods using developing teeth,For Sci Int 2006; 159S: S68 – S73
15. Chaillet N, Nystrom M, Demirjian A. Comparison of dental maturity in children of different ethnic origins: international maturity curves for clinicians. J Forensic Sci 2005;50(5):1164–74.
16. Harris EF. Mineralization of the mandibular third molar: a study of American blacks and whites.Am J Phys Anthropol. 2007 Jan;132(1):98-109.
17. Olze A, Schmeling A, Taniguchi M, Maeda H, van Niekerk P, Wernecke KD, Geserick G. Forensic age estimation in living subjects: the ethnic factor in wisdom tooth mineralization.Int J Legal Med. 2004;118(3):170-3.
18. Liversidge HM, Kosmidou A, Hector MP, Roberts GJ Epidermolysis bullosa and dental developmental age. Int J Paediatr Dent. 2005 Sep;15(5):335-41.
19. Seow WK. Dental development in amelogenesis imperfecta: a controlled study. Pediatr Dent. 1995;17(1):26-30.
20. Kotilainen J, Pirinen S. Dental maturity is advanced in Fragile X syndrome. Am J Med Genet. 1999; 83(4):298-301.
21. Hazza'a AM, Al-Jamal G. Dental development in subjects with thalassemia major, J Contemp Dent Pract. 2006;7(4):63-70.
22. Uslenghi S, Liversidge HM, Wong FS. A radiographic study of tooth development in hypodontia.Arch Oral Biol. 2006 Feb;51(2):129-33.
23. Cameriere R, Flores-Mir C, Mauricio F, Ferrante L. Effects of nutrition on timing of mineralization in teeth in a Peruvian sample by the Cameriere and Demirjian methods.Ann Hum Biol. 2007 Sep-Oct;34(5):547-56.
24. Chaillet N., Nystrom M, Kataja M., Demirjian A, Dental maturity curves in Finnish children: Demirjian's method revisited and polynomial functions for age estimation. J Forensic Sci. 2004; 49:1324-31
25. Pinchi V, Franchi E, Manetti G, Norelli G-A. Identificazione dell’età e mineralizzazione dell’ottavo: verifica del metodo in una casistica italiana. Zacchia, 2005;78:484-498
26. Cameriere R, Ferrante L, De Angelis D, Scarpino F, Galli F. The comparison between measurement of open apices of third molars and Demirjian stages to test chronological age of over 18 year olds in living subjects.Int J Legal Med. 2008 Nov;122(6):493-7.
27. Braga J., Heuze Y., Chadabel O., Sonan N.K. , Gueramy A., Non-adult dental age assessment: correspondence analysis and linear regression versus Bayesian predictions, Int. J. Legal Med. 2005; 119: 260–74
28. Demirijan A., Golstein H., Tanner JM, A new system of dental age assessment, Hum Biol, 1973;45: 211-27
29. Duda R. O., Hart E. (1973), Pattern classification and scene analysis, Wiley, New York.
30. Langley P., Iba W., Thompson K., An analysis of Bayesian classifiers, Proceedings, Tenth National Conference on Artificial Intelligence, 1992; 223–228. San Jose: AAAI Press.
31. Pearl J., Probabilistic reasoning in intelligent system: Networks of plausible inference, 1988, Morgan Kaufmann Publishers Inc., United States of America
32. Bilmes J., On Soft Evidence in Bayesian Networks, Tech. Rep. UWEETR-2004-0016, University of Washington, Dept. of Electr. Engineering

33. Shapiro M.D., The evaluation of clinical predictions. A method and initial application, N Eng J Med 1977, 296(26):1509-1514
34. Bhat VJ, Kamath GP. Age estimation from root development of mandibular third molars in comparison with skeletal age of wrist joint. Am J Forensic Med Pathol. 2007; 28(3):238-41.
35. Meinl T A. Tangl S., Huber C., Maurer B Watzek G. The chronology of third molar mineralization in the Austrian population—a contribution to forensic age estimation For Sci Int 2007; 169: 161-7
36. Olze A., Bilang D., Schmidt S., Wernecke K.D., Geserick G., Schmeling A., Validation of common classification system for assessing the mineralization of third molars, Int. J. Legal Med. 2005;119: 22–6
37. Soomer H., Lincoln M.J., Ranta H., Penttila A., Leibur E., Dentist's Qualifications Affect the Accuracy of Radiographic Identification, J Forensic Sci, 2003; 48, 1121-1126.

Tables and Figures

Tab.1 Percentages of correctly classified individuals, based on the combination of teeth and different classification probabilistic thresholds π - Expert A.

Probabilistic classification threshold π	T_1	T_2	T_3	T_4	$T_{1,2}$	$T_{3,4}$	$T_{1,4}$	$T_{2,3}$	$T_{1,3}$	$T_{2,4}$	$T_{1,2,3,4}$
0,50	80,8	80,0	80,9	81,8	78,9	80,6	82,1	81,5	82,3	81,5	80,8
0,70	79,1	78,3	76,4	77,4	78,7	79,1	80,2	79,1	80,0	79,5	79,8
0,80	67,8	64,9	66,9	65,1	73,4	75,1	76,7	73,6	76,4	74,7	78,2
0,90	59,6	61,7	58,3	57,9	68,3	64,7	69,0	68,4	68,9	69,5	74,3
0,95	31,6	47,5	44,0	48,5	62,4	60,7	59,0	61,3	59,4	61,9	71,1
0,99	27,7	27,7	27,7	27,7	44,3	49,2	41,2	50,3	39,7	48,8	63,9

Tab.2 Percentages of misclassified minors, based on combination of teeth and different classification probabilistic thresholds π - Expert A.

Probabilistic classification threshold π	T_1	T_2	T_3	T_4	$T_{1,2}$	$T_{3,4}$	$T_{1,4}$	$T_{2,3}$	$T_{1,3}$	$T_{2,4}$	$T_{1,2,3,4}$
0,50	44,6	44,4	51,6	46,3	32,0	35,5	30,6	32,1	33,2	32,2	25,8
0,70	27,0	26,4	30,4	28,4	24,5	26,2	24,6	23,3	23,5	23,1	22,5
0,80	16,3	8,3	13,4	13,1	18,3	20,7	20,2	15,3	19,1	16,0	20,4
0,90	8,7	6,7	5,3	5,8	12,4	8,9	12,2	10,0	10,9	11,0	14,3
0,95	0,6	1,8	2,1	1,8	8,2	6,0	6,4	6,7	6,1	6,0	11,4
0,99	0	0	0	0	1,5	2,3	1,4	1,8	1,9	1,8	7,5

Tab.3 Percentages of individuals aged < 17 years according technology - Expert A

Individuals aged < 17 years (Expert A)	Classified as		
	< 17 (correctly)	17 o 18 (no decision)	≥ 19 (uncorrectly)
Digital	72,1	22,7	5,3
Analogic	81,3	16,9	1,7
Average	79,1	18,4	2,5

Tab.4 Percentages of individuals aged < 17 years according technology - Expert B

Individuals aged < 17 years (Expert B)	Classified as		
	< 17 (correctly)	17 o 18 (no decision)	≥ 19 (uncorrectly)
Digital	76,1	19,4	4,5
Analogical	74,7	17,8	7,5
Average	75,0	18,1	6,9

Tab.5 Percentages of individuals aged ≥ 19 years according technology - Expert A

Individuals aged ≥ 19 years (Expert A)	Classified as		
	<17 (uncorrectly)	17 o 18 (no decision)	≥ 19 (correctly)
Digital	5,9	13,9	80,2
Analogical	7,6	19,9	72,5
Average	7,2	18,5	74,3

Tab.6 Percentages of individuals aged ≥ 19 years according technology - Expert B

Individuals aged ≥ 19 years (Expert B)	Classified as		
	<17 (uncorrectly)	17 o 18 (no decision)	≥ 19 (correctly)
Digital	8,0	13,5	78,6
Analogical	7,9	18,9	73,2
Average	7,9	17,7	74,5

Tab.7 Inter-observers reproducibility by tooth and technology

INTER-observers reproducibility	T_1	T_2	T_3	T4
Analogical	0,659	0,673	0,688	0,738
Digital	0,622	0,650	0,668	0,648
Average	0,651	0,668	0,685	0,722

Tab.8 Intra-observers reproducibility by tooth and technology - Expert A

INTRA-observers reproducibility (A)	T_1	T_2	T_3	T_4
Analogical	0,844	0,814	0,842	0,857
Digital	0,776	0,790	0,867	0,865
Average	0,817	0,804	0,851	0,860

Tab.9 Intra-observers reproducibility by tooth and technology - Expert B

INTRA-observers reproducibility (B)	T_1	T_2	T_3	T_4
Analogical	0,689	0,719	0,714	0,829
Digital	0,754	0,673	0,759	0,791
Average	0,714	0,700	0,731	0,815

Fig.1 Relationship between percentage of non-classified individuals and probabilistic classification threshold

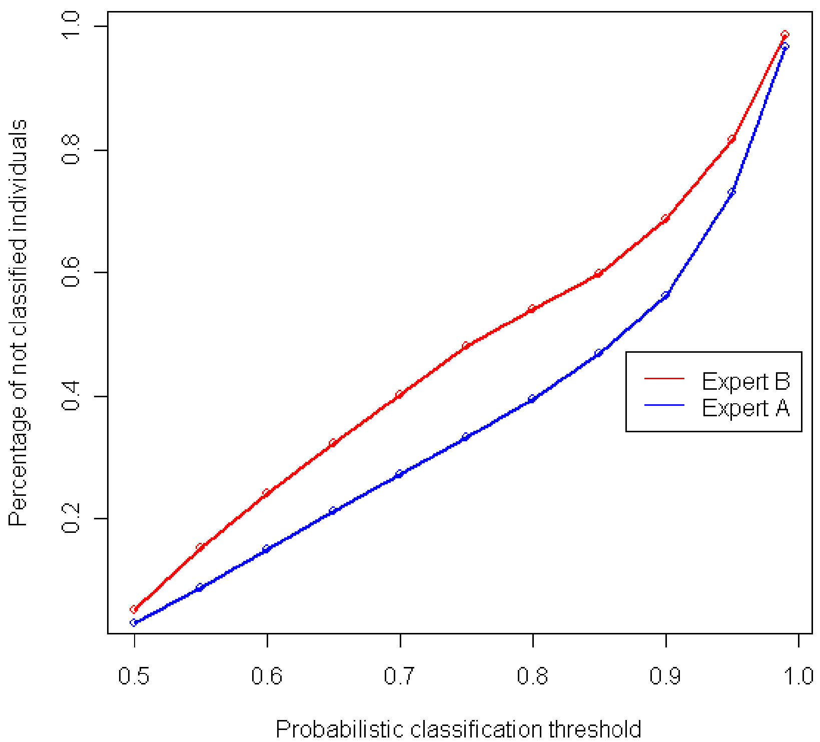
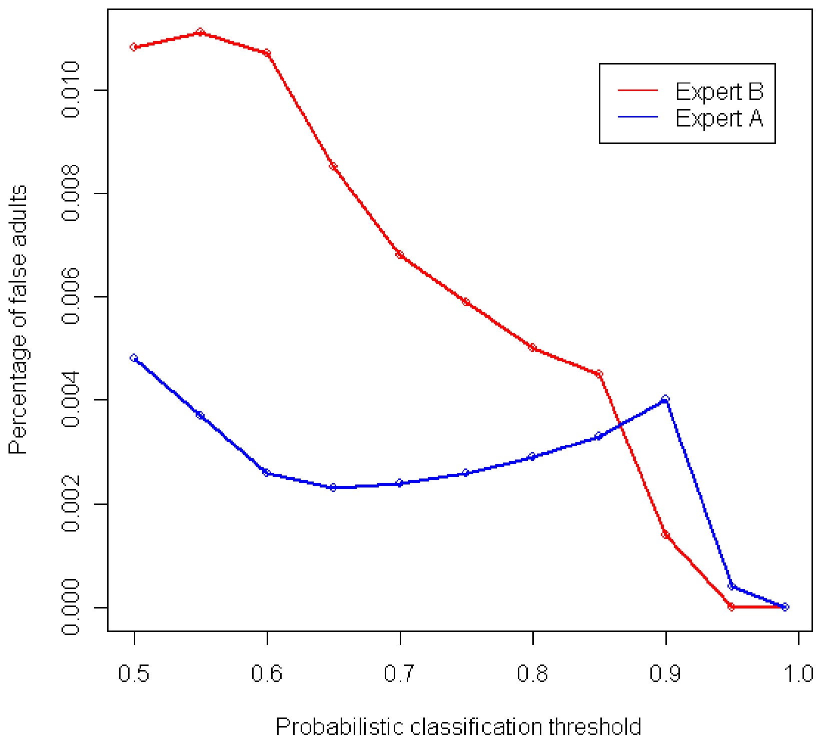


Fig.2 Relationship between percentage of false adults and probabilistic classification threshold


Expert B
Expert A
Percentage of false adults
0.000
0.002
0.004
0.006
0.008
0.010
0.5
0.6
0.7
0.8
0.9
1.0
Probabilistic classification threshold