A Bayesian decision analysis approach to assess voice disorder risks by using acoustic features

Article *in* Biometrical Journal · November 2018
DOI: 10.1002/bimj.201700233

4 authors:

María Jesús Rufo
Universidad de Extremadura

Jacinto Martín
Universidad de Extremadura

Carlos J. Pérez
Universidad de Extremadura

Sandra Paniagua
Universidad de Extremadura

A Bayesian decision analysis approach to assess voice disorder risks by using acoustic features

María J. Rufo[*,1], **Jacinto Martín** [2], **Carlos J. Pérez** [3], and **Sandra Paniagua** [4]

[1] School of Technology, University of Extremadura *(Department of Mathematics, Avda de la Universidad s/n, 10003, Cáceres, Spain)*

[2] Faculty of Science and ICCAEx, University of Extremadura *(Department of Mathematics, Avda de Elvas s/n , 06006, Badajoz, Spain)*

[3] Faculty of Veterinary, University of Extremadura *(Department of Mathematics, Avda de la Universidad s/n, 10003, Cáceres, Spain)*

[4] Faculty of Nursing, University of Extremadura *(Department of Nursing, Calle Sta. Teresa Jornet, 38, 06800, Mérida, Spain)*

Received zzz, revised zzz, accepted zzz

Vocal fold nodules are recognized as an occupational disease for all collective of workers performing activities for which maintained and continued use of voice is required. Computer-aided systems based on features extracted from voice recordings have been considered as potential non-invasive and low cost tools to diagnose some voice-related diseases. A Bayesian decision analysis approach has been proposed to classify university lectures in three levels of risk: low, medium and high, based on the information provided by acoustic features extracted from healthy controls and people suffering from vocal fold nodules. The proposed risk groups are associated with different treatments. The approach is based on the calculation of posterior probabilities of developing vocal fold nodules and considers utility functions that include the financial cost and the probability of recovery for the corresponding treatment. Maximization of the expected utilities is considered. By using this approach, the risk of having vocal fold nodules is identified for each university lecturer, so he/she can be properly assigned to the right treatment. The approach has been applied to university lecturers according to the Disease Prevention Program of the University of Extremadura. However, it can also be applied to other voice professionals (singers, speakers, coaches, actors...).

Key words: Acoustic features; Bayesian decision analysis; Disease prevention program; Utility function; Voice disorder

1 Introduction

Voice disorders affect communication and have important implications for public health (see, Roy *et al.* (2005), Cohen and Garret (2007) and Bhattacharyya (2014), among others). Some of the most common vocal pathologies are laryngitis, vocal fold nodules and polyps. They can severely affect the vocal function. A current study about voice therapy in vocal fold lesions is provided by Tang and Thibeault (2017).

Voice disorders are recognized as an occupational disease for all collective of workers performing activities for which maintained and continued use of voice is required (singers, speakers, teachers...). Voice disorders in adults are common among these professionals, see, for instance, De Medeiros *et al.* (2012), Cantor Cutiva *et al.* (2013), Ubillos *et al.* (2015) and Pestana *et al.* (2017). The professional use of voice is mainly characterized by an excessive vocal load, leading, in many cases, to dysphonia. Dysphonia is one of

*Corresponding author: e-mail: mrufo@unex.es, Phone: +34 927257220, Fax: +34 927257103

the most prevalent disorders among teachers (see Niebudek-Bogusz *et al.* (2007)). The recognition of this health problem as an occupational disease requires to identify hazards and establish preventive measures.

Voice recordings have been considered as a potential non-invasive and low cost biomarker to diagnose some voice-related diseases. Baghai-Ravary and Beet (2013) provided a view of automatic speech signal analysis for clinical diagnosis and assessment of speech disorders. Most researches consider protocols with phonations of sustained vowel, mainly /a/, but also words and phrases are considered. The recorded sounds are transformed into electrical signals by using a microphone. Later, they are numerically quantified by using feature extraction algorithms and used in pattern recognition algorithms to classify individuals by similitude, allowing to discriminate healthy voices from diseased ones (see, e.g, Ali *et al.* (2016) or Jalalinajafabad (2016)).

Computer-Aided Diagnosis (CAD) is a broad concept that integrates signal processing, artificial intelligence and statistics into computerized techniques that assist health professionals in their decision-making processes. These techniques seek to maximize the information that may be automatically extracted from medical tests via objective and quantitative computations (see, e.g, Calle-Alonso *et al.* (2013)). In this context, a CAD system can be designed considering voice recordings as the ingredient to extract information based on signal processing from healthy and diseased people.

Medical institutions, insurance companies, policy makers and clinicians are faced with the decision of replacing an existing procedure with a new one or using complementary information, and how to update protocols. These decisions are primarily based on a trade-off between accurate diagnostic and cost-effectiveness (see Kornak and Lu (2011)). The economic costs derived from sick leaves caused by vocal diseases in voice professionals could be reduced by applying prevention programs based on automatic non-invasive and low cost tools as the one proposed here. Hence, the determination of appropriate treatment may bring benefits not only to the individual, but also to the society.

The statistical decision theory deals with scenarios where decisions have to be made under a state of uncertainty. The goal is to provide a rational framework for dealing with such situations. These scenarios are typical in medical decision-making (see, for instance Barbini *et al.* (2013) and Stallard *et al.* (2017)). In addition, Ashby and Smith (2000) argued that the natural statistical framework in this context is a Bayesian approach, since it allows to incorporate an integrated summary of the available evidence and associated uncertainty with assessment of utilities.

This paper proposes a Bayesian decision analysis approach that allows an automatic classification of university lecturers into three different voice risk groups (low, medium and high risks) by using acoustic features extracted from their voice recordings. These groups involve both the disease status and the corresponding treatment based on the cost and the recovery rate. The approach considers the maximization of the expected utilities. Firstly, the posterior probabilities of developing a vocal disorder is provided based on healthy subjects and people suffering from vocal fold nodules (modelling uncertainty). These probabilities are obtained by using a probit regression model. Next, the utility functions are taken into account. They involve the costs for the different treatments and the recovery utility (modelling preferences). A Monte Carlo approach is proposed to estimate the expected utilities. Thus, the identification of medium-risk subjects will lead them to a speech therapy program, whereas those allocated to the high-risk group will be candidates for a more advanced treatment that could include surgery. Finally, the results show that this methodology may be useful for Prevention and Health Promotion Services of institutions and companies whose workers use the voice as a work tool. Although the approach has been used with university lecturers, it is applicable to other voice professionals by following the same framework.

The outline of the paper is as follows. Information on participants, speech recordings and feature extraction is presented in Section 2. Section 3 presents the approach by describing the modelling of uncertainty and preferences. The experimental results are shown in Section 4. Section 5 presents a simulation-based analysis to address a theoretical scenario where the real status of the subjects are known. Finally, the conclusion is presented in Section 6.

2 Acoustic feature database

In this section, information on participants, speech recordings and feature extraction necessary to build the acoustic feature database is presented.

2.1 Participants

A total of 90 university lecturers (59 men and 31 women) from the University of Extremadura (Cáceres, Spain) were involved in the study. The mean ($\pm$ standard deviation) age was 46.38 $\pm$ 8.73. A protocol consisting of a physical examination, a survey and voice recordings was applied to the subjects within the Program for Health Promotion of the University of Extremadura. The protocol was approved by the Bioethical Committee. All subjects signed an informed consent.

2.2 Speech recordings

The vocal task was the sustained phonation of the /a/ vowel at comfortable pitch and loudness, as constant as possible. This phonation had to be kept for at least 5 seconds and on one breath. Three valid phonations were obtained per individual.

The speech data were recorded using a portable computer with an external sound card (TASCAM US322) and a headband microphone (AKG 520) featuring a cardiod pattern. The digital recording was performed at a sampling rate of 44.1 KHz and a resolution of 16 bits/sample by using Audacity software (release 2.0.5).

In addition, the voice recordings of 53 healthy controls and 19 subjects suffering vocal fold nodules were obtained from the commercial database MEEI (Massachusetts Eye and Ear Infirmary) of KayPENTAX Inc. (see Daoudi and Bertrac (2014)).

2.3 Feature extraction

The study is based on three acoustic features and the gender. Local Shimmer as an amplitude perturbation measure was considered (see, e.g., Baken and Orlikoff (2000)). Harmonic-to-noise ratio (HNR) is a measure of the relative level of noise present in speech that also have been considered (see, e.g., Shue *et al.* (2010)). Finally, the existence of nonlinear phenomena in the production process of the speech signal has been theoretically and experimentally established. Specifically, the Recurrence Period Density Entropy (RPDE) is used here (see, e.g., Little *et al.* (2007)).

After feature extraction, the three replications of each feature were averaged. Then, a matrix with 162 rows (one for each of the 53 healthy controls, 19 patients and 90 university lecturers) and 4 columns (sex, Shimmer, HNR, and RPDE) was obtained.

3 Model description

Suppose that the voice quality status of a set of university lecturers have to be analyzed. The main aim is to classify university lecturers according to three different treatments related to the three voice risk groups (low, medium and high). A cost-utility model will be used for this purpose. The proposed model includes modelling uncertainty, that is associated with the probability that a university lecturer develops vocal problems. This uncertainty will be modelled by using the information available from previous subjects which are known to be healthy or diseased.

The proposed Bayesian approach could be adequately represented by using the influence diagram shown in Figure 1. Influence diagrams are powerful graphical tools for dealing with Bayesian decision problems under uncertainty (see, e.g., Ríos and Ríos Insua (2009), and Bielza *et al.* (2011)). They are acyclic directed graphs with three types of nodes and two types of arcs. Decision nodes, which are represented by a square or rectangle, chance or uncertainty nodes that are symbolized by a circle or an ellipse and value nodes that are represented as diamonds or hexagons. In relation to the arcs, they can be conditional arcs which are directed towards a chance node or a value node and informational arcs that are directed toward a decision node. Arcs towards a value or uncertainty node indicate functional and probabilistic dependence, respectively. Finally, arcs into a decision node indicate that when the decision is made, the values of the preceding nodes are known. Deeper studies about influence diagrams and analysis of medical decision problems can be found in Owens *et al.* (1997) and Pauker and Wong (2005).

For the influence diagram in Figure 1, depending on each university lecturer's *state*, their *acoustic predictors* are obtained. This last node includes the covariates that consist of acoustic features and gender from the subjects previously evaluated. These predictors are observed, for each university lecturer, before making *decisions* and they consist of assigning a medical treatment to each one of them. Such node is discrete. Once the decision is made and, by taking into account that it has influence on the *recovery* of each university lecturer, the utility for each one of them can be evaluated. These *utilities* encloses both the corresponding cost of the treatment and the utility of the university lecturer' state after the treatment.

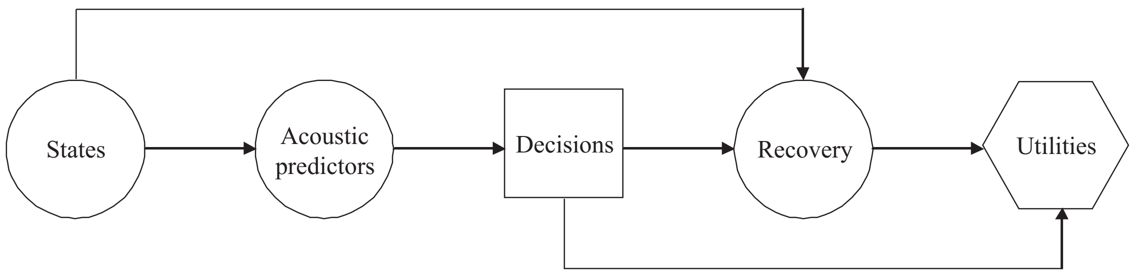


Figure 1 Influence diagram for the process.

In order to solve the influence diagram in Figure 1 (see, for instance, Shachter (1986)), firstly, the arc between the nodes *states* and *acoustic predictors* is inverted. Thus, the posterior probability that a university lecturer suffers from vocal nodule is obtained for each one of them. As a consequence, the node corresponding to *acoustic predictors* can be removed. Once the decision for each university lecturer and the posterior probability of having the disease are known, then the probabilities of recovering are calculated. Finally, by taking into account both the previous decision and the probability of recovering, the expected utility is also obtained for each university lecturer.

In the next two subsections the model is described in detail.

3.1 Modelling uncertainty

Following the influence diagram in Figure 1, firstly, the posterior distributions of suffering from vocal fold nodules is calculated for each university lecturer based on the probit regression model. The process will consists of two steps which will be described next.

Suppose that m independent binary random variables $Y_1, Y_2, \ldots, Y_m$ are observed, where Y_j is Bernoulli distributed with success probability $P(Y_j = 1) = p_j$, $j = 1, 2, \ldots, m$. The probabilities p_j are related to a set of known covariates $\boldsymbol{x}_j^t = (x_{j1}, x_{j2}, \ldots, x_{jK})$ through the following model:

$$\begin{aligned} Y_j &\sim \text{Bernouilli}\,(p_j)\,, \\ p_j &= \Psi\left(\boldsymbol{x}_j^t \boldsymbol{\beta}\right), \end{aligned} \tag{1}$$

where $\boldsymbol{\beta}$ is a $K \times 1$ vector of unknown parameters and $\Psi(\cdot)$ is the standard Gaussian cumulative density function.

The first step deals with the obtention of a random sample from the regression vector $\boldsymbol{\beta}$. Therefore, the task is to infer $\boldsymbol{\beta}$ given the vector of binary data $\boldsymbol{y} = (y_1, y_2, \ldots, y_m)^t$, $y_j \in \{0, 1\}$ and the matrix of explanatory variables $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m)^t$.

The likelihood is given as:

$$\mathcal{L}\left(\boldsymbol{\beta}|\boldsymbol{y},\boldsymbol{x}\right) = \prod_{j=1}^{m} \left[\Psi\left(\boldsymbol{x}_j^t\boldsymbol{\beta}\right)\right]^{y_j} \ \left[1-\Psi\left(\boldsymbol{x}_j^t\boldsymbol{\beta}\right)\right]^{1-y_j}.$$

Latent variables are introduced based on the proposal of Albert and Chib (1993), i.e. $z_1, z_2, \ldots, z_m$ are considered, where z_j is distributed as Normal$\left(\boldsymbol{x}_j^t\boldsymbol{\beta}, 1\right)$ and it is defined:

$$Y_j = \begin{cases} 1 \text{ if } z_j > 0 \\ 0 \text{ if } z_j \leq 0 \end{cases}$$

In addition, by considering a multivariate normal distribution, $\boldsymbol{\beta} \sim$Normal$_K\left(\boldsymbol{b}, \boldsymbol{B}\right)$, as the prior distribution for the regression vector, $\pi\left(\boldsymbol{\beta}\right)$, then the posterior distribution has the following expression:

$$\pi\left(\boldsymbol{\beta}|\boldsymbol{y},\boldsymbol{x},\boldsymbol{z}\right) \propto \pi\left(\boldsymbol{\beta}\right) \prod_{j=1}^{m} \int_{-\infty}^{\infty} (2\pi)^{-1/2} \exp\left\{-\frac{1}{2}\left(z_j - \boldsymbol{x}_j^t\boldsymbol{\beta}\right)^2\right\} dz_j.$$

This posterior density is not analytically tractable, so it must numerically be estimated. MCMC methods are computing tools that can be used in this context (see Naranjo *et al.* (2016) and references therein). Specifically a Gibbs sampling algorithm is considered. Thus, the full conditional distributions are given by:

$$z_j|\boldsymbol{y},\boldsymbol{x},\boldsymbol{\beta} \sim \begin{cases} \text{Normal}\left(\boldsymbol{x}_j^t\boldsymbol{\beta}, 1\right) I_{\{z_j>0\}} \text{ if } y_j = 1 \\ \\ \text{Normal}\left(\boldsymbol{x}_j^t\boldsymbol{\beta}, 1\right) I_{\{z_j\leq 0\}} \text{ if } y_j = 0 \end{cases},$$

$$\boldsymbol{\beta}|\boldsymbol{z},\boldsymbol{y},\boldsymbol{x} \sim \text{Normal}_K\left(\boldsymbol{b}^*, \boldsymbol{B}^*\right),$$

where $\boldsymbol{B}^* = \left[\boldsymbol{x}^t\boldsymbol{x} + \boldsymbol{B}^{-1}\right]^{-1}$ and $\boldsymbol{b}^* = \boldsymbol{B}^*\left[\boldsymbol{x}^t\boldsymbol{z} + \boldsymbol{B}^{-1}\boldsymbol{b}\right]$.

The final Gibbs sampling-based algorithm consists of choosing an initial value $\boldsymbol{\beta}^{(0)}$, and iteratively sampling $z^{(l)}$ and $\boldsymbol{\beta}^{(l)}, l = 1, 2, \ldots, N$, from the full conditional distributions.

In the second step, a sample corresponding to the posterior probabilities of suffering from vocal nodules for each university lecturer ($i = 1, 2, \ldots, n$) is obtained. In order to do it, the parameter values, $\boldsymbol{\beta}^{(l)}$, are considered together with the acoustic features of the university lecturers that have to be classified. Hence, by taking into account the probit model (1), the probabilities are given by:

$$p_i^{(l)} = \Psi\left(\boldsymbol{x}_i^t\boldsymbol{\beta}^{(l)}\right), i = 1, 2, \ldots, n,\ l = 1, 2, \ldots, N. \tag{2}$$

3.2 Modelling preferences

Once the posterior probabilities have been obtained, the next step is to select a suitable utility function. Observe that the set of alternatives contains three possible decisions: no treatment (NT) and the type of treatment, which will depend on the university lecturer's needs and are called medium (MT) or high (HT) treatments. In addition, by taking into account the previously calculated probabilities, the states of nature are healthy (H) or suffering from nodules (SN). In this context, a multicriteria additive utility function is considered for all university lecturers. Those criteria are the cost of the corresponding treatment and the utility of recovery. This information is obtained from doctors and health financial managers.

Specifically, the following utility functions are considered:

$$\begin{aligned}\mathcal{U}_i\left(\cdot,\mathrm{SN}\right) &= q_{\cdot_i}\, 0 + (1-q_{\cdot_i})\, A + C_{\cdot_i},\\ \mathcal{U}_i\left(\cdot,\mathrm{H}\right) &= 0,\end{aligned}$$

where $\cdot$ represents the different decisions (treatments), and $q_{\cdot_i}$ is the probability of recovery with MT or HT for the patient i. Note that the probability of recovery with NT is 0. A is a parameter which denotes the cost of not improving after the treatment (he/she still sick). The constants $C_{\cdot_i}$ include the mean costs and the inconvenience of the treatment. In addition, they depend on the considered treatment: MT or HT. As a consequence, the values for the parameter A and the constants C_i are negatives.

Therefore, a classification approach based on the maximization of the expected utility is proposed and applied to each university lecturer. The expressions for the corresponding expected utilities are obtained. If π_i denotes the posterior distribution corresponding to the probability of suffering from vocal nodules and Θ the states of nature, then:

$$\begin{aligned}\mathcal{U}_i\left(\mathrm{NT}\right) &= E_{\pi_i}\left[\mathcal{U}_i\left(\mathrm{NT},\Theta\right)\right] = (1-p_i)\ \mathcal{U}_i\left(\mathrm{NT},\mathrm{H}\right) + p_i\,\mathcal{U}_i\left(\mathrm{NT},\mathrm{SN}\right) =\\ &= (1-p_i)\ 0 + p_i\ A = p_i\ A,\end{aligned} \tag{3}$$

where p_i is the mean for the posterior probabilities of suffering from nodules.

In the same way, the expressions for the utilities with the MT and HT are, respectively,

$$\mathcal{U}_i\left(\mathrm{MT}\right) = E_{\pi_i}\left[\mathcal{U}_i\left(\mathrm{MT},\Theta\right)\right] = p_i\,(1-q_{\mathrm{MT}_i})\,A + C_{\mathrm{MT}_i}, \tag{4}$$

$$\mathcal{U}_i\left(\mathrm{HT}\right) = E_{\pi_i}\left[\mathcal{U}_i\left(\mathrm{HT},\Theta\right)\right] = p_i\,(1-q_{\mathrm{HT}_i})\,A + C_{\mathrm{HT}_i}, \tag{5}$$

where q_{MT_i} and q_{HT_i} are provided by doctors based on previous information. This framework can be also used when information about possible relationships between the probabilities p_i and q_{MT_i} and the probabilities p_i and q_{HT_i} are available.

3.3 A Monte Carlo approach

The expectations given in (3), (4) and (5) can be easily estimated by considering the posterior probabilities obtained in (2). Hence, the corresponding estimates for each individual, $i = 1, 2, \ldots, n$ are given by the following expressions:

$$\begin{aligned}\widehat{\mathcal{U}}_i\left(\mathrm{NT}\right) &= A\,\overline{p_i^{(l)}},\\ \widehat{\mathcal{U}}_i\left(\mathrm{MT}\right) &= A\,\overline{p_i^{(l)}}\,(1-q_{\mathrm{MT}_i}) + C_{\mathrm{MT}_i},\\ \widehat{\mathcal{U}}_i\left(\mathrm{HT}\right) &= A\,\overline{p_i^{(l)}}\,(1-q_{\mathrm{HT}_i}) + C_{\mathrm{HT}_i},\end{aligned}$$

where $\overline{p_i^{(l)}} = \dfrac{1}{N}\displaystyle\sum_{l=1}^{N} p_i^{(l)}$.

On the other hand, when information that allows to relate the probabilities q_{MT_i} and q_{HT_i} to the probability p_i is available, then functions $q_{\mathrm{MT}_i}(p_i)$ and $q_{\mathrm{HT}_i}(p_i)$ could be considered. Thus, the previous approximations are given by:

$$\widehat{\mathcal{U}}_i\left(\mathrm{NT}\right) = A\,\overline{p_i^{(l)}}, \tag{6}$$

$$\widehat{\mathcal{U}}_i\left(\mathrm{MT}\right) = A\,\overline{p_i^{(l)}\left(1-q_{\mathrm{MT}_i}(p_i^{(l)})\right)} + C_{\mathrm{MT}_i}, \tag{7}$$

$$\widehat{\mathcal{U}}_i\left(\mathrm{HT}\right) = A\,\overline{p_i^{(l)}\left(1-q_{\mathrm{HT}_i}(p_i^{(l)})\right)} + C_{\mathrm{HT}_i}, \tag{8}$$

where $\overline{p_i^{(l)}\left(1-q_{\mathrm{MT}_i}(p_i^{(l)})\right)} = \frac{1}{N}\sum_{l=1}^{N} p_i^{(l)}\left(1-q_{\mathrm{MT}_i}(p_i^{(l)})\right)$ and a similar expression is obtained for $\overline{p_i^{(l)}\left(1-q_{\mathrm{HT}_i}(p_i^{(l)})\right)}$.

Next section shows how the proposed approach is applied in a real scenario and the experimental results obtained.

4 Experimental results

The response variable Y takes the value $y = 1$ for people suffering from vocal fold nodules and $y = 0$ for healthy subjects. The prior distribution of $\boldsymbol{\beta}$, $\pi(\boldsymbol{\beta})$, is a noninformative normal distribution with mean $\boldsymbol{b} = (0,0,0,0,0)^t$ and covariance matrix $\boldsymbol{B} = \text{diag}_5(100)$.

The proposed approach is applied to obtain a random sample corresponding to the 5×1 vector $\boldsymbol{\beta}$ of unknown regression parameters. The Gibbs sampling-based algorithm is run for 10,000 iterations with a burn-in of 500 samples. The sample values are taken by considering intervals of ten units length. By using these specifications, the chain seems to have converged. Therefore, a sample of size $N = 950$ is obtained, i.e., $\boldsymbol{\beta}^{(l)} = \left(\beta_0^{(l)}, \beta_1^{(l)}, \beta_2^{(l)}, \beta_3^{(l)}, \beta_4^{(l)}\right)^t$, $l = 1, 2, \ldots, 950$.

Once the random sample for the regression vector $\boldsymbol{\beta}$ has been obtained by using previous information available from subjects which are known to be healthy or diseased, then, in the following step, the posterior probabilities of suffering nodules for university lecturers are calculated considering the collected database of their acoustic features $\boldsymbol{x}_i^t = (1, x_{i1}, x_{i2}, x_{i3}, x_{i4})$, $i = 1, 2, \ldots, 90$, where $x_{i1}, x_{i2}, x_{i3}, x_{i4}$ denote the values for the variables sex, shimmer, HNR and RPDE. The expression for the posterior probabilities in (2) becomes:

$$p_i^{(l)} = \Psi\left(\boldsymbol{x}_i^t \boldsymbol{\beta}^{(l)}\right), \; i = 1, 2, \ldots, 90, \; l = 1, 2, \ldots, 950. \tag{9}$$

Now, the expected utilities given in the equations (6), (7) and (8) are estimated. Following the recommendations provided by the Prevention and Health Promotion Service of the University of Extremadura, the following values are considered: $A = -2500$, $C_{\mathrm{MT}_i} = -200$ and $C_{\mathrm{HT}_i} = -500$. These considered costs are averaged costs based on the information on previous university lecturers.

The probabilities of recovery with the MT (q_{MT_i}) and HT (q_{HT_i}) depend on the estimated posterior probabilities of suffering nodules $p_i^{(l)}$ given in expression (9). In addition, it is also considered that both q_{MT_i} and q_{HT_i} are varying according to a threshold. This value has been taken as 0.8 for the two recovery probabilities in this application. Hence, the functions $q_{\mathrm{MT}_i}(p_i)$ and $q_{\mathrm{HT}_i}(p_i)$ in the Monte Carlo approach given in (7) and (8) have the following expressions:

$$q_{\mathrm{MT}_i}(p_i^{(l)}) = \begin{cases} 0 & \text{if } p_i^{(l)} > 0.8 \\ 0.8 & \text{if } p_i^{(l)} \leq 0.8 \end{cases} \quad \text{and } q_{\mathrm{HT}_i}(p_i^{(l)}) = \begin{cases} 0.5 & \text{if } p_i^{(l)} > 0.8 \\ 1 & \text{if } p_i^{(l)} \leq 0.8 \end{cases}$$

Note that the smaller the values of $p_i^{(l)}$ are, the larger the values of q_{MT_i} and q_{HT_i} are. These specifications have also been provided by the members of the Prevention and Health Promotion Service of the University of Extremadura, and the estimated expected utilities for each university lecturer, $i = 1, 2, \ldots, 90$, are:

$$\begin{aligned} \widehat{\mathcal{U}}_i\,(\text{NT}) &= -2500\,\overline{p_i^{(l)}}, \\ \widehat{\mathcal{U}}_i\,(\text{MT}) &= 2500\,\overline{p_i^{(l)}\left(q_{\mathrm{MT}_i}(p_i^{(l)})-1\right)} - 200, \\ \widehat{\mathcal{U}}_i\,(\text{HT}) &= 2500\,\overline{p_i^{(l)}\left(q_{\mathrm{HT}_i}(p_i^{(l)})-1\right)} - 500. \end{aligned}$$

where $\overline{p_i^{(l)}\left(q_{\mathrm{MT}_i}(p_i^{(l)})-1\right)} = \sum_{l=1}^{950} p_i^{(l)}\left(q_{\mathrm{MT}_i}(p_i^{(l)})-1\right)\Big/950$. The same expression is obtained for $\overline{p_i^{(l)}\left(q_{\mathrm{HT}_i}(p_i^{(l)})-1\right)}$.

Figure 2 shows the estimated expected utilities for each university lecturer and its maximum value. It can be observed that the maximum expected utility corresponds to NT for 46 lecturers, to MT for 23 lecturers, and to HT for 21 lecturers. Therefore, these university lectures are classified in those groups.

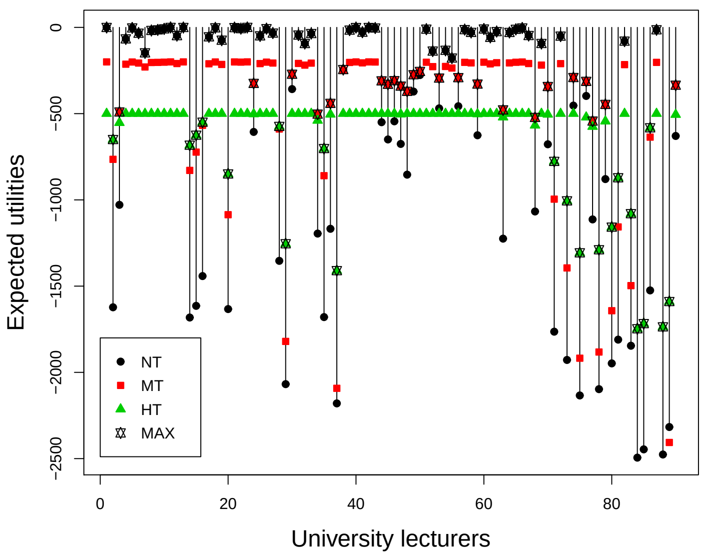


Figure 2 Expected utilities and maximum value.

The posterior predictive distributions for the university lecturers are shown in Figure 3. Note that all distributions are asymmetric. Besides, the probabilities of suffering from vocal fold nodules are increasing according to the considered group in the following order: NT, MT, and HT. In particular, for patients into NT group, these probability values are small. Table 1 presents the quartiles of the estimated posterior probabilities in (9) for each group. Observe that the interquartile ranges are very different from one group to another.

	NT	MT	HT
Q_1	0.00005	0.12448	0.62481
Q_2	0.00181	0.24719	0.78559
Q_3	0.01532	0.40898	0.94099

Table 1 Quartiles for the posterior probabilities corresponding to NT, MT and HT.

Finally, a university lecturer from each group is considered. The generated posterior probabilities in (1) are presented in Figure 4. In can be observed that lecturer 7 has lower probability values of suffering from nodules than lecturer 90. At the same time, lecturer 2 has higher probability values of suffering the disease than lecturer 90. As it is expected, they have been classified into non treatment, medium and high treatments, respectively.

The source code to reproduce the results is available as Supporting Information on the journal's web page ($http://onlinelibrary.wiley.com/doi/xxx/suppinfo$).

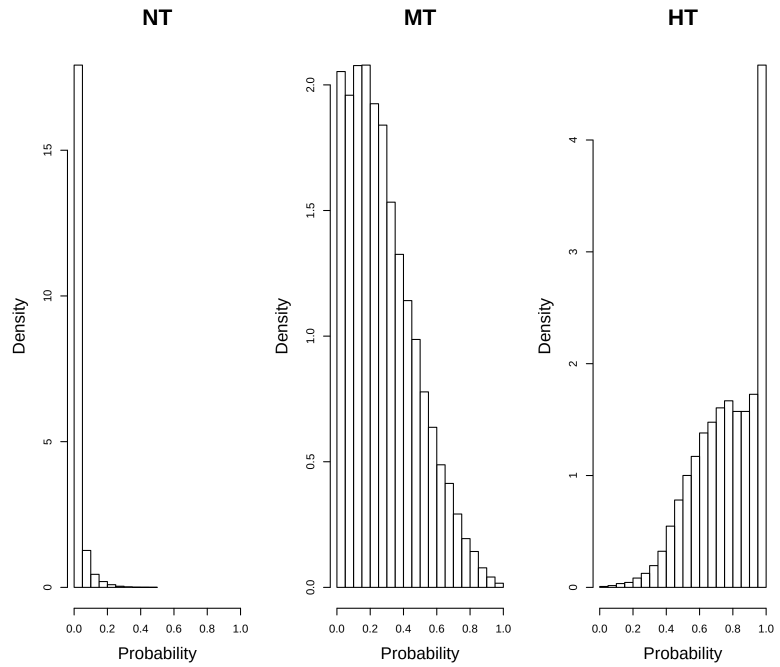


Figure 3 Posterior predictive distributions for the university lecturers classified into NT, MT and HT groups.

This approach has been implemented and used in the Prevention and Health Promotion Service of the University of Extremadura in the context of the Disease Prevention Program. The medium treatment has been mainly focussed on logopaedics interventions and the high treatment considers phoniatrics and otorhinolaryngology interventions.

Next, a simulation-based analysis is considered to address a theoretical scenario where the real status of the subjects are known.

5 Simulation-based analysis

A simulation-based experiment is conducted to analyze the model performance. A total of $N = 100$ simulations have been carried out in a random stratified cross-validation. For each simulation, a sample corresponding to 92 subjects has been generated partially based on the experimental data considered in the previous section. Specifically, the gender variable is generated by using a Bernoulli distribution with parameter $p = 0.5$. With respect to the acoustic variables, Table 2 shows the way they have been generated so that the distributions corresponding to the commercial database MEEI in Subsection 2.2 can be approximated.

Variables	**Healthy distributions**	**Diseased distributions**
Shimmer	$Gamma(2, scale = 120)$	$Gamma(5, scale = 100)$
HNR	$Normal(25, 3.5)$	$Normal(19, 4.5)$
RPDE	$Normal(0.25, 0.17)$	$Normal(0.33, 0.1)$

Table 2 Distributions to generate acoustic covariates for healthy and diseased subjects.

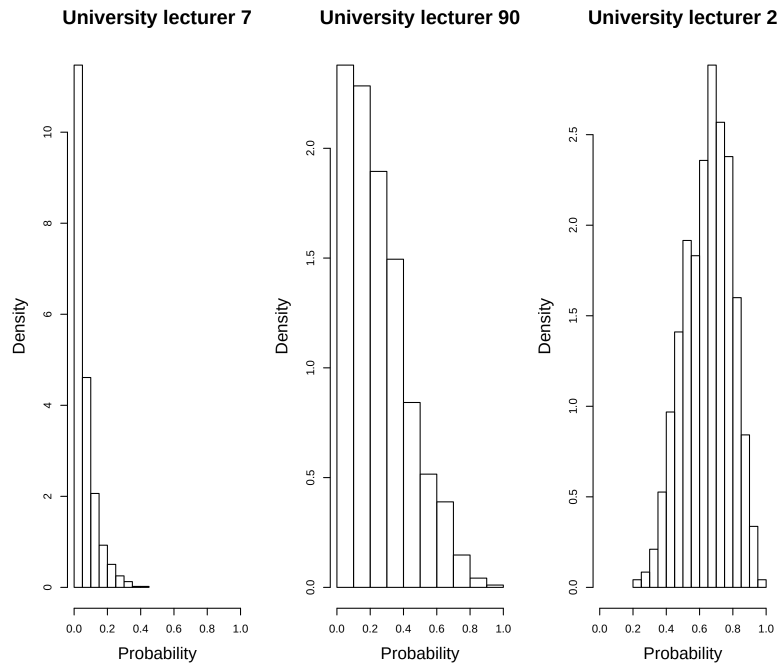


Figure 4 Posterior probabilities of suffering the disease.

Each generated dataset is randomly split into a training subsample composed by 53 individuals drawn from *healthy distributions* and 19 people drawn from *diseased distributions* in Table 2. The remaining 20 subjects (15 and 5 individuals distributed according to *healthy* and *diseased distributions*, respectively) constitute the testing subsample. The proposed approach learns from the training subsample and the accuracy is computed from the testing subsample.

Once the different samples have been obtained, a probit model is applied for each simulation. Thus, it is obtained a random sample for the regression vector $\boldsymbol{\beta}$ by using the training subsample. Then, the main goal is to classify the 20 subjects from the testing subsample into three different treatments: NT, HT, and MT. By using the sample corresponding to the regression vector, the posterior probabilities of suffering from nodules given in (2) are calculated. Then, the expected utilities given in equations (6), (7) and (8) are estimated in the same way as in the previous application. In addition, given that the subjects have been artificially generated, the expected utility corresponding to the different treatments can be analytically calculated. That is, for each subject i, $i = 1, 2, \ldots, 20$, the probability of suffering from vocal nodules is exactly known. Hence, these values can be used in order to calculate the expected utilities $\mathcal{U}_i$ (NT), $\mathcal{U}_i$ (MT) and $\mathcal{U}_i$ (HT) given in expressions (3), (4) and (5), respectively.

Next, three measures are proposed to evaluate the approach. Two of them are based on the estimation of the expected utilities and the other one is related to the estimation corresponding to the allocated treatment.

For each subject i, $i = 1, 2, \ldots, 20$, the errors between the exact and the estimated values obtained for the expected utilities are denoted by:

$$\begin{aligned}
\mathcal{U}_{i||}\,(\mathrm{NT}) &= \left|\mathcal{U}_i\,(\mathrm{NT}) - \widehat{\mathcal{U}}_i\,(\mathrm{NT})\right|, \\
\mathcal{U}_{i||}\,(\mathrm{MT}) &= \left|\mathcal{U}_i\,(\mathrm{MT}) - \widehat{\mathcal{U}}_i\,(\mathrm{MT})\right|, \\
\mathcal{U}_{i||}\,(\mathrm{HT}) &= \left|\mathcal{U}_i\,(\mathrm{HT}) - \widehat{\mathcal{U}}_i\,(\mathrm{HT})\right|,
\end{aligned}$$

where $\mathcal{U}_i$ (NT), $\mathcal{U}_i$ (MT) and $\mathcal{U}_i$ (HT) are calculated by considering the probability of suffering from nodules whereas $\widehat{\mathcal{U}}_i$ (NT), $\widehat{\mathcal{U}}_i$ (MT) and $\widehat{\mathcal{U}}_i$ (NT) are computed by using the mean of the generated posterior predictive probabilities.

Thus, once the exact values have been calculated for each subject i and, as a consequence thereof, the treatment with the highest value for the expected utility is known. Then, the corresponding error $\mathcal{U}_{i||}$ (NT), $\mathcal{U}_{i||}$ (MT) or $\mathcal{U}_{i||}$ (HT) is computed. Finally, it is considered:

$$\widehat{\mathcal{U}}_{\max} = \max_{i=1,2,\ldots 20} \left\{ \widehat{\mathcal{U}}_{i\,\max} \right\},$$

where $\widehat{\mathcal{U}}_{i\,\max} = \left| \mathcal{U}_i^* (\cdot) - \widehat{\mathcal{U}}_i (\cdot) \right|$, with $\mathcal{U}_i^* (\cdot) = \max \left\{ \mathcal{U}_i \,(\mathrm{NT}), \mathcal{U}_i \,(\mathrm{MT}), \mathcal{U}_i \,(\mathrm{HT}) \right\}$ and $\cdot$ denotes the corresponding treatment for the maximum in $\mathcal{U}_i^* (\cdot)$. That is, for each subject i, it is computed the difference between the *real* expected utility and the estimated one for the optimal treatment. Hence, the maximum of the previous values is taken as a measure of these errors.

On the other hand, it is also taken:

$$\widehat{\overline{\mathcal{U}}}_{\max} = \sum_{i=1}^{20} \frac{1}{20} \widehat{\mathcal{U}}_{i\,\max}.$$

Therefore, it is shown the mean distance corresponding to the errors obtained as the differences between the maximum analytical value and the corresponding estimated one for the expected utilities.

Observe that, the two measures above-mentioned are related to errors associated to the expected utility of the optimal decision. The first one, $\widehat{\mathcal{U}}_{\max}$, only focuses on the subject i with the highest error value, $\widehat{\mathcal{U}}_{i\,\max}$, whereas the second measure, $\widehat{\overline{\mathcal{U}}}_{\max}$, is an average of the highest error value for each subject i, $i = 1, 2, \ldots, 20$.

Finally, a third measure is based on the optimal decision. Thereby, MP refers to the rate of misclassified subjects. Table 3 shows the mean of these three measures together with the standard deviation, for $N = 100$ replications of the procedure:

Measures	$\widehat{\mathcal{U}}_{\max}$	$\widehat{\overline{\mathcal{U}}}_{\max}$	MP
Mean of the values	623.35	89.91	0.109
Standard deviation	344.43	50.24	0.085

Table 3 Evaluation measures.

Observe that, the first measure in Table 3 presents a high value, since it is based on the misclassified subject for whom $\widehat{\mathcal{U}}_{\max}$ takes the highest value in each replication. When the mean for the previous differences, $\widehat{\mathcal{U}}_{i\,\max}$, are considered, then it obtained that the average value for all replications is 89.91. Note that, the values for these reference measures are varying from around 2000 to around 2500 when absolute value is considered (see, Figure 2). Thus, the obtained values for $\widehat{\mathcal{U}}_{\max}$ and $\widehat{\overline{\mathcal{U}}}_{\max}$ are not large as compared with the previous ones. In addition, the average percentage of subjects correctly classified is about 90%.

6 Conclusion

This paper presents an approach for the classification of university lecturers according to a treatment associated with the risk of suffering vocal problems. This procedure takes into account the probabilities of recovery, the posterior probability of developing vocal problems and the utility function that includes the different criteria, i.e. the cost of the corresponding treatment and the utility of recovery. Although, the approach have been applied to university lecturers for which an experiment has been conducted in the Prevention and Health Promotion Service of the University of Extremadura, the approach can be applied to other professionals that use the voice as a work tool.

To the best of the authors' knowledge, this is the first approach based on acoustic features considering costs to classify professionals of voice in three risk groups with different actions. The way the assignment to groups is performed is based on automatic feature extraction, which is a non-invasive and low-cost technique that can be implemented in Prevention and Health Promotion services of companies and public administrations.

Acknowledgments

We thank Antonio Moreno and Gloria Gragales, as members of the Prevention and Health Promotion Service of the University of Extremadura, for their help in the experimental part of this study. We also thank the volunteers participating in the Disease Prevention Program of the University of Extremadura.

This research has been supported by projects MTM2014-56949-C3-3-R and MTM2017-86875-C3-2-R (MINECO), and projects IB16054 and GR15106 (Junta de Extremadura/European Regional Development Funds, EU).

References

Albert, J.H. and Chib, S. (1993) Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association*, **88**, 669–679.

Ali, Z., Alsulaiman, M., Elamvazuthi, I., Muhammad, G., Mesallam, T.A., Farahat, M., and Malki, K.H. (2016) Voice pathology detection based on the modified voice contour and svm. *Biologically Inspired Cognitive Architectures*, **15** (Supplement C), 10–18.

Ashby, D. and Smith, A.F.M. (2000) Evidence-based medicine as Bayesian decision-making. *Statistics in Medicine*, **19** (23), 3291–3305.

Baghai-Ravary, L. and Beet, S.W. (2013) *Automatic speech signal analysis for clinical diagnosis and assessment of speech disorders*, Springer Briefs in Electrical and Computer Engineering, Springer.

Baken, R.J. and Orlikoff, R.F. (2000) *Clinical Measurement of Speech and Voice*, Singular Thomson Learning, San Diego, 2nd edn..

Barbini, E., Manzi, P., and Barbini, P. (2013) *Current Topics in Public Health, Dr. Alfonso Rodriguez-Morales (Ed.)*, InTech, DOI: 10.5772/52402, https://www.intechopen.com/books/current-topics-in-public-health/bayesian-approach-in-medicine-and-health-management, chap. Bayesian Approach in Medicine and Health Management, pp. 17–35.

Bhattacharyya, N. (2014) The prevalence of voice problems among adults in the United States. *Laryngoscope*, **124**, 2359–2362.

Bielza, C., Gómez, M., and Shenoy, P.P. (2011) A review of representations issues and modeling challenges with influence diagrams. *Omega*, **39** (3), 227–241.

Calle-Alonso, F., Pérez, C.J., Arias-Nicolás, J.P., and Martín, J. (2013) Computer-aided diagnosis system: A Bayesian hybrid classification method. *Computer Methods and Programs in Biomedicine*, **112** (1), 104 – 113.

Cantor Cutiva, L.C., Vogel, I., and Burdorf, A. (2013) Voice disorders in teachers and their associations with work-related factors: a systematic review. *Journal of Communication Disorders*, **46**, 143–155.

Cohen, S.M. and Garret, C.G. (2007) Utility of voice therapy in the management of vocal fold polyps and cysts. *Otolaryngol Head Neck Surgery*, **136** (5), 742–746.

Daoudi, K. and Bertrac, B. (2014) *On classification between normal and pathological voices using the MEEI-KayPENTAX database: Issues and consequences*, INTERSPEECH, Singapore, chap. INTERSPEECH-2014.

De Medeiros, A.M., Assunção, A.A., and Barreto, S.M. (2012) Absenteeism due to voice disorders in female teachers: A public health problem. *International Archives of Occupational and Environmental Health*, **85** (8), 853–864.

Jalalinajafabad, F. (2016) *Computerised GRBAS assessment of voice quality*, PhD dissertation, Faculty of Engineering and Physical Sciences, University of Manchester.

Kornak, J. and Lu, Y. (2011) Bayesian decision analysis for choosing between diagnostic/prognostic prediction procedures. *Stat Interface*, **4**, 27–36.

Little, M.A., McSharry, P.E., Roberts, S.J., Costello, D.A.E., and Moroz, I.M. (2007) Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection. *BioMedical Engineering OnLine*, **6** (23), 1–19.

Naranjo, L., Pérez, C.J., Campos-Roca, Y., and Martín, J. (2016) Addressing voice recording replications for Parkinson's disease detection. *Expert System with applications*, **46**, 286–292.

Niebudek-Bogusz, E., Kotylo, P., and Śliwinńska-Kowalska, M. (2007) Evaluation of voice acoustic parameters related to the vocal-loading test in proffessionally active teachers with dysphonia. *International Journal of Occupational Medicine and Environmental Health*, **20**, 25–30.

Owens, D.K., Shachter, R.D., and Nease, R.F.J. (1997) Representation and analysis of medical decision problems with influence diagrams. *Medical decision making*, **17** (3), 241–262.

Pauker, S. and Wong, J. (2005) The influence of influence diagrams in medicine. *Decision analysis*, **2**, 238–244.

Pestana, P.M., Vaz-Freitas, S., and Manso, M.C. (2017) Prevalence of voice disorders in singers: Systematic review and meta-analysis. *Journal of Voice. In Press. https://doi.org/10.1016/j.jvoice.2017.02.010*.

Ríos, J. and Ríos Insua, D. (2009) Supporting negotiations over influence diagrams. *Decision Analysis*, **6** (3), 153–171.

Roy, N., Merrill, R.M., Gray, S.D., and Smith, E.M. (2005) Voice disorders in the general population: Prevalence, risk factors, and occupational impact. *Laryngoscope*, **115** (11), 1988–1995.

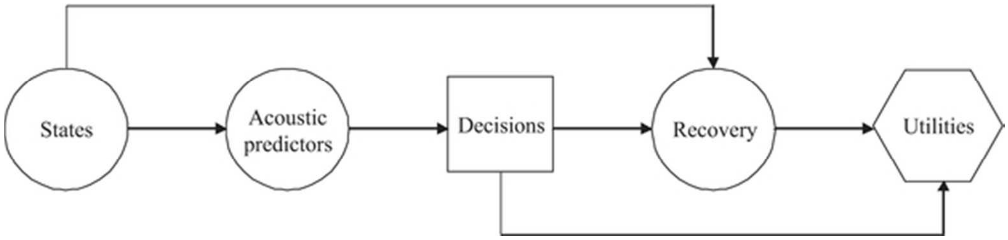
Shachter, R.D. (1986) Evaluating influence diagrams. *Operations Research*, **34** (6), 871–882.

Shue, Y.L., Keating, P., Vicenik, C., and Yu, K. (2010) VoiceSauce: A program for voice analysis. *Energy*, **1** (H2), H1–A1.

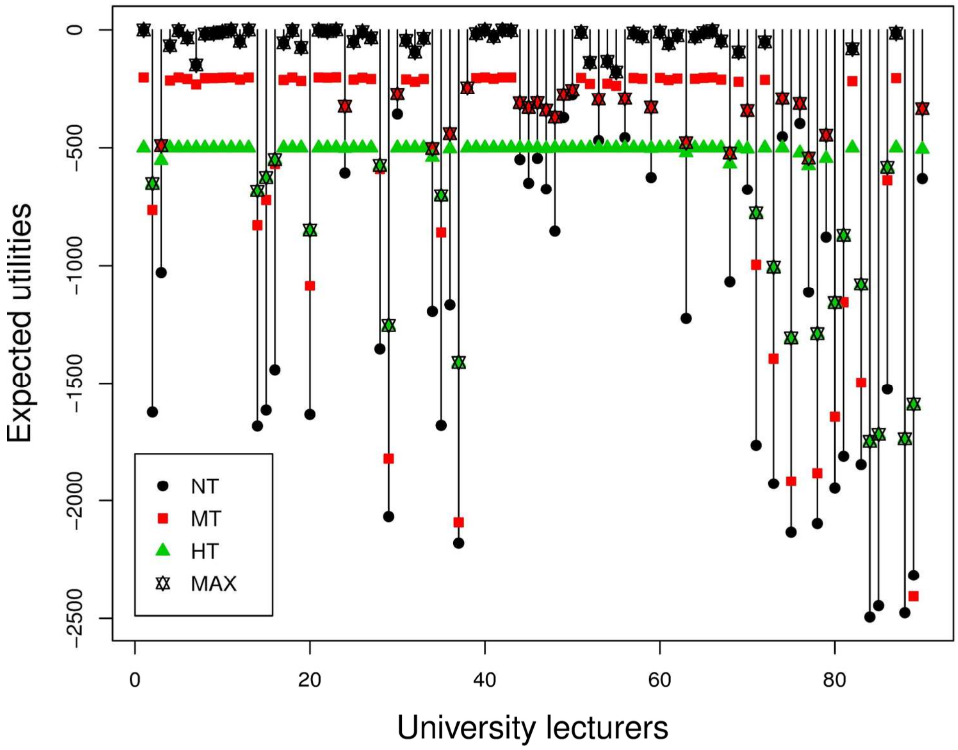
Stallard, N., Miller, F., Day, S., Hee, S.W., Madan, J., Zohar, S., and Posch, M. (2017) Determination of the optimal sample size for a clinical trial accounting for the population size. *Biometrical Journal*, **59**, 609–625.

Tang, S.S. and Thibeault, S.L. (2017) Timing of voice therapy: A primary investigation of voice outcomes for surgical benign vocal fold lesion patients. *Journal of Voice*, **31** (129), 129.e1–129.e7.

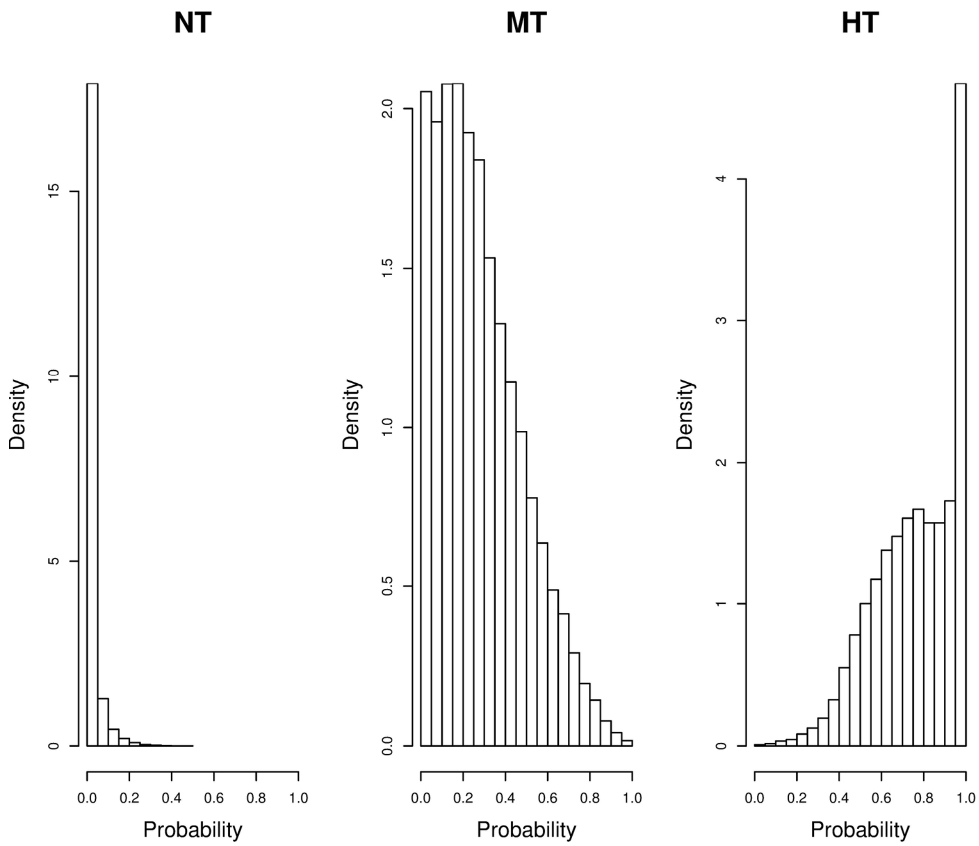
Ubillos, S., Centeno, J., Ibañez, J., and Iraurgi, I. (2015) Protective and risk factors associated with voice strain among teachers in Castile and Leon, spain: Recommendations for voice training. *Journal of Voice*, **29** (2), 261.e1–261.e12.



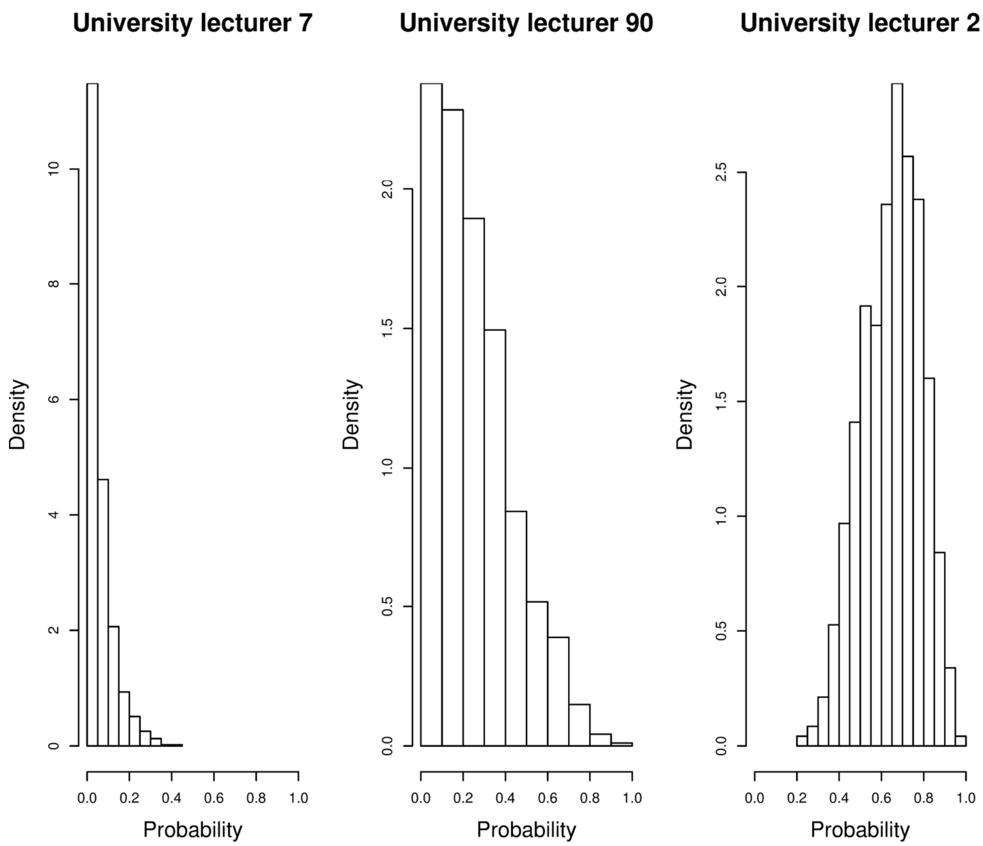
Influence diagram for the process.



Expected utilities and maximum value.



Posterior predictive distributions for the university lecturers classified into NT, MT and HT groups.



Posterior probabilities of suffering the disease.