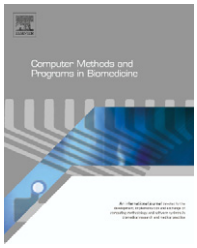
COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE 90 (2008) 104–116

journal homepage: www.intl.elsevierhealth.com/journals/cmpb



Bayesian classification for the selection of in vitro human embryos using morphological and clinical data

Dinora Araceli Morales[a,*], *Endika Bengoetxea*[b], *Pedro Larrañaga*[a,e], *Miguel García*[c], *Yosu Franco*[c], *Mónica Fresnada*[a], *Marisa Merino*[d]

[a] *Department of Computer Science and Artificial Intelligence, University of the Basque Country, Paseo Manuel Lardizabal 1, E-20018 Donostia-San Sebastián, Spain*
[b] *Department of Computer Architecture and Technology, University of the Basque Country, Paseo Manuel Lardizabal 1, E-20018 Donostia-San Sebastián, Spain*
[c] *Clínica del Pilar, Donostia-San Sebastián, Spain*
[d] *Hospital Donostia, Donostia-San Sebastián, Spain*
[e] *Department of Artificial Intelligence, Technical University of Madrid, Campus de Montegancedo, Boadilla del Monte 28660 Madrid, Spain*

ARTICLE INFO

Article history:
Received 25 April 2007
Received in revised form
27 November 2007
Accepted 28 November 2007

Keywords:
Bayesian classifiers
Decision support systems
Feature subset selection
IVF outcome prediction

ABSTRACT

In vitro fertilization (IVF) is a medically assisted reproduction technique that enables infertile couples to achieve successful pregnancy. Given the uncertainty of the treatment, we propose an intelligent decision support system based on supervised classification by Bayesian classifiers to aid to the selection of the most promising embryos that will form the batch to be transferred to the woman's uterus. The aim of the supervised classification system is to improve overall success rate of each IVF treatment in which a batch of embryos is transferred each time, where the success is achieved when implantation (i.e. pregnancy) is obtained.

Due to ethical reasons, different legislative restrictions apply in every country on this technique. In Spain, legislation allows a maximum of three embryos to form each transfer batch. As a result, clinicians prefer to select the embryos by non-invasive embryo examination based on simple methods and observation focused on morphology and dynamics of embryo development after fertilization.

This paper proposes the application of Bayesian classifiers to this embryo selection problem in order to provide a decision support system that allows a more accurate selection than with the actual procedures which fully rely on the expertise and experience of embryologists. For this, we propose to take into consideration a reduced subset of feature variables related to embryo morphology and clinical data of patients, and from this data to induce Bayesian classification models. Results obtained applying a filter technique to choose the subset of variables, and the performance of Bayesian classifiers using them, are presented.

© 2007 Elsevier Ireland Ltd. All rights reserved.

* *Corresponding author*. Tel.: +34 943018070; fax: +34 943015590.
E-mail address: dinora-morales@ehu.es (D.A. Morales).
0169-2607/$ – see front matter © 2007 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.cmpb.2007.11.018

1. Introduction

Infertility of couples is nowadays considered an important social problem which is subject of special interest by medical doctors and biologists. Intensive research is being done in this field in order to improve both the techniques and treatments applied to improve the results and welfare of patients. Research trends include developments in medical technology on assisted human reproduction in aspects such as equipment technology, medical treatments, and also on artificial intelligence. More precisely, this research area combines both advances on clinician and embryologist knowledge together with data mining techniques. Data mining techniques allow developing an intelligent system which could support the decision of the embryologist expert in order to choose the most promising embryos that will form the batch to be transferred to the woman's uterus. Since the embryo selection problem faces diverse legislative restrictions that are different in most of the countries, the procedures and methodologies that are applied in clinical practice vary from a country to another. In our case, we deal with data obtained from Clínica del Pilar, which is subject to Spanish legislation. Spanish regulations changed recently in May 2005, when the Parliament restricted to a maximum of three, the number of embryos transferred in order to reduce the incidence of multifetal pregnancies, but removed restrictions to the number of oocytes that can be fecundated in each IVF treatment, leaving this decision at the criterion of the responsible biomedical team at each clinic.

Embryologists, who handle human germ cells and embryos, are familiarized with non-invasive and precise techniques of embryo evaluation. Again due to ethical and legislative reasons, protocols of human embryo manipulation are very restricted in aspects such as the period of time for the follow-up of promising embryos after fertilization. This point is critical for a successful embryo selection since precise examination of embryos on particular days after fertilization by human assisted reproduction methods facilitate selection of the most promising embryos for transfer.

Other important aspect to improve the success rates is related not only to the improved embryo examination and selection, but also to age, presence of oocyte dimorphisms, sperm quality, fertilization rate, cleavage rate and number of embryos transferred, endometrial thickness and number of previous cycles of treatment [1]. Moreover, other factors need to be taken into account for the final decision, among them the type of oocyte insemination method such as in vitro fertilization (IVF) and intra-cytoplasmic sperm injection (ICSI), embryo transfer, preimplantation genetic diagnosis, and finally the composition of culture media.

The recent literature shows examples of applying artificial intelligence methods to improve success rates of IVF programs. Saith et al. [2] analyse the potential contribution of artificial intelligence to the embryo selection process and proposes a data mining model based on decision trees to investigate the relationship between the features of the embryo, oocyte and follicle to the successful outcome of the embryo transfer. Jurisica et al. [3] present a case-based reasoning system in the form of an intelligent decision support system for IVF practitioners that, in some situations, is able to suggest possible treatments to improve the success rate. Trimarchi et al. [4] provide a study of models based on data mining techniques, in particular the C5.0 algorithm, to infer classification trees. Patrizi et al. [5] present a pattern recognition algorithm to select embryos from images, which classifies the objects given into a number of classes and formulate from these a general rule. Manna et al. [6] compare the precision in the recognition of viable embryos by a group of experts to that of a machine recognition procedure.

This paper contributes to this research field by presenting a novel intelligent decision support system for IVF treatment based on detailed analysis of human embryo morphology and clinician data of patients. Our proposal is a Bayesian supervised classification system that can assist on the selection of the most promising embryos for implantation in IVF treatment.

The rest of the paper is organized as follows: the next section describes the embryo selection decision of IVF treatments as a supervised classification problem, describes the characteristics of the clinician database used in our research and introduces the Bayesian classifiers that we applied to predict the success or not a determined embryo-batch transfer procedure. Section 3 shows experimental results and their interpretation applying these classifiers, and analyses the classification performance of our method. Finally in Section 4, conclusions and some trends for future work are presented.

2. Characterization of the most suitable embryos on IVF treatment

Human assisted reproduction methods like in vitro fertilization (IVF) through insemination of oocyte and sperm, and the process called intra cytoplasmic sperm injection (ICSI) in which sperm is injected into the oocyte, are widely applied to treat infertility. In the fertilization treatment oocyte and sperm are obtained separately. In order to obtain a sufficient number of oocyte ovulatory stimulants are used. These stimulants make pituitary to increase secretion of follicle stimulating hormones. Later, embryos are formed outside the body and are developed in a controlled atmosphere. A few embryos, the ones deemed best by the clinician and embryologist regarding the likelihood of bringing forth a child, are chosen and transferred to the woman's uterus within 48–72 h from their formation.

One of the most relevant aspects in human assisted reproduction is the characterization of the most suitable embryos to transfer in a patient. Embryologists select the embryos by non-invasive embryo examination based on simple observation focused on morphology and dynamics of embryo development. The analysis is performed under contrast-phase microscope with Hoffmann modulation contrast (HMC) or difference-interference contrast (DIC), enabling more precise assessment without fixing and staining. When the embryo is at pronuclear stage, the embryo examination is usually performed visually and the evaluation is totally subjective. Therefore, the experience and expertise of the embryologist is of high importance for the final success rate, since the appropriateness of the selected embryos is critical for a prospective good implantation.

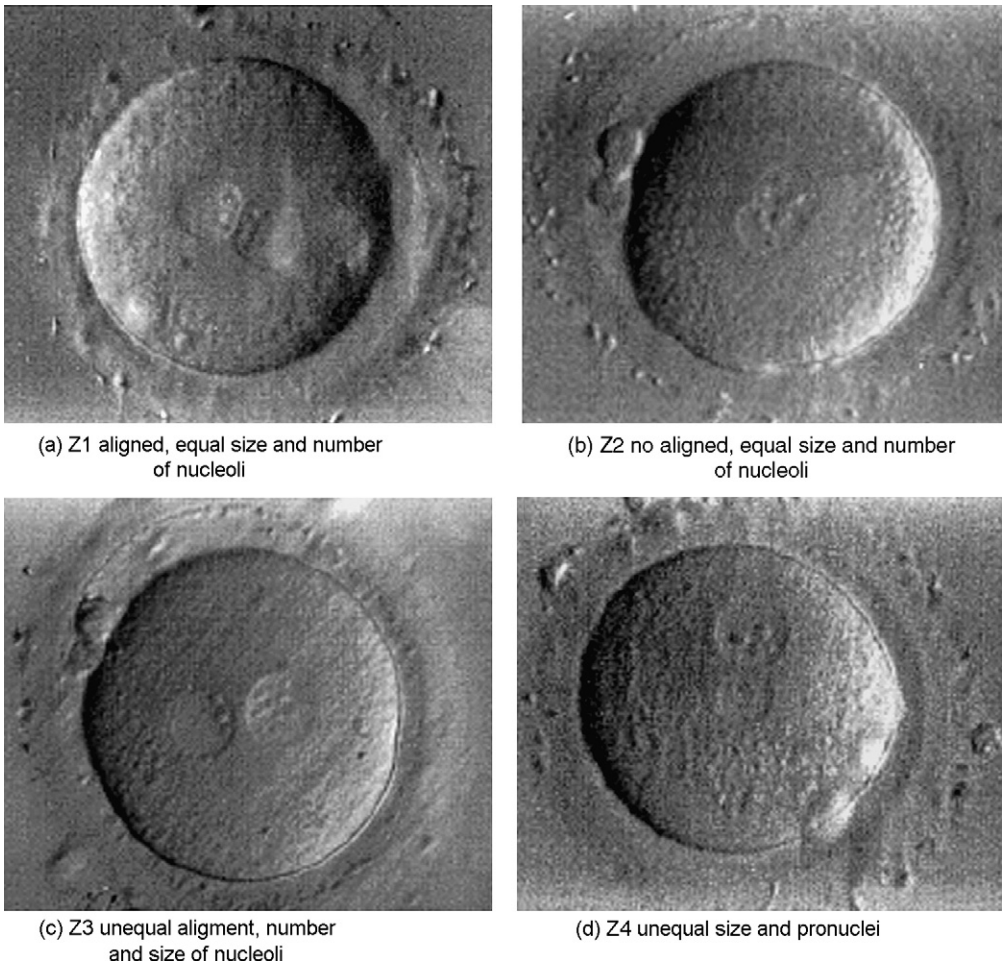


Fig. 1 – (a–d) Real zygote images of the database of Clinica del Pilar (San Sebastián, Spain), catalogued following the Scott score [7] which takes into account parameters such as nuclear size and alignment, nucleoli number, and their distribution. From all these categories, Z1 and Z2 are considered to be the most promising types, specially the first of them.

Routine embryo evaluation commences 16–18 h after oocyte fertilization (either in IVF or ICSI) with a zygote score and concludes at 48–72 h after fertilization with the characterization of each embryo using a concrete score. Fig. 1 shows examples of real images obtained in this clinic, which were catalogued by expert embryologists according to the pronuclear Scott's score [7] such as a that is nowadays widely accepted as a standard and with plenty of reports confirming its usefulness in selecting good quality zygotes as Z1 to Z2 usually yield prospectively better quality embryos per transfer, which results in higher implantation rates. However, there are also in the literature many other zygote scores that catalogue the embryo taking into account many other different criteria rather than uniquely morphological characteristics [8–11].

The selection of embryos protocol in Clínica del Pilar is a procedure that consists of several steps. The initial observation of embryos is performed 16–19 h after fertilization to check the existence of fertilization. Subsequent embryo inspections are routinely performed in daily intervals 40–44 h and 64–68 h, 48 h after fertilization being the most indicated time to proceed for catalogue their quality. Following these evaluations, embryos are divided into five different categories following the standard described in [12], which categorize them according to several morphological parameters such as the number of cells, percentage of fragmentation, good cell–cell contact and the existence or not of multinucleated blastomeres. This categorization leads to the following embryo cleavages:

- Type I: no fragmentation with equal size homogeneous blastomeres;
- Type II: <10% fragmentation with equal size homogeneous blastomeres;
- Type III: this type is usually divided in two subtypes: Type III-A: 10–25% fragmentation with equal size blastomeres; Type III-B: 25–50% fragmentation with unequal size blastomeres;
- Type IV: >50% fragmentation with equal or unequal size blastomeres;
- Type V: 100% fragmentation.

2.1. *Characteristics of samples and the database*

Data for this paper was obtained from 63 clinical files of the IVF programme in Clínica del Pilar in San Sebastián, during the period from July 2003 through December 2005.

Table 1 – Predictor features of the clinical database, with their respective values and codes

Feature	Range of values
Number of actual cycle	Numerical
Number of previous cycles of treatment	Numerical
Age	Numerical
Donate	{Yes, No}
Sperm quality	{Good, Medium, Poor}
Cause female	{Yes, No}
Cause male	{Yes, No}
Primary infertility	{Yes, No}
Secondary infertility	{Yes, No}
Cell number	Numerical
Zygote score	{Z1, Z2, Z3, Z4}
Embryo categorization	{Type I, Type II, Type III, Type IV, Type V}
Blastomeres size	{Equal, Unequal, Defects of cytoplasm}
Fragmentation blastomeres	{0%, (0–10]%, (10–25]%, (25–50]%, (50–99]%, 100%}
Thickness zona pellucida	None, Thick, Very Thick
Multinuclear	{Yes, No}
Number of transferred embryos	Numerical
Number of frizzed embryos	Numerical
Quality of transference of embryos	Good, Medium, Poor
Day of transference	Numerical

These 63 cases were chosen from a total of 89 cycles of in vitro treatment from patients aged 27–46, by discarding treatments that were not performed under comparable conditions. These 63 treatment cases have information of a total of 189 embryos that were transferred in batches of three embryos per patient aged 30–40. In all these cases, the morphological characteristics of embryos at the 4–8 cell stage were taken 40–50 h after fertilization and before transfer during the second or third day. Fresh and frozen embryos were considered in this retrospective study which 18 three embryo-batches were implanted.

The protocol applied in this clinic for the selection of embryos is based primarily on zygote score [7] and cleaving embryo morphology [12]. In order to reduce the incidence of higher-order multiple pregnancies, only two or three embryos are transferred between the second or third day after fertilization. We did not include in our study 26 cases treatments consisting of two embryo-batches cases, of women aged 27–29 and 41–46.

Apart from morphologic data obtained from each of the embryos, our database also registers other variables related with clinical parameters and features such as the age, number of previous cycles of treatment, actual cycle, sperm quality, cause of female or male infertility, primary or secondary infertility, number of embryos transferred, whether embryos were frozen or not, quality of transference and transference day. Primary infertility is defined as a couple that has never been able to conceive a pregnancy, after at least 1 year of unprotected intercourse. Secondary infertility is presented when a couple has previously been pregnant at least once, but has not been able to achieve another pregnancy. Table 1 shows the whole list of variables that contains the database with their respective values and codes.

In our study, we had to make some assumptions due to the nature of the data and the limitations to obtain it. The restriction of having three embryos transferred at a time leads embryologists experts to assume that when implantation occurs. There is no means to be more exhaustive, since the only solution would be to perform a biopsy to all embryos before transfer and this procedure is also known to affect directly the stress of embryos. Since there is no means to identify which of the original embryos was the one that was implanted, we must emphasize once again the assumption taken in our study that the embryos that reached implantation are always the ones that obtained a higher score from the three of each transferred batch. Since there is no means to be sure that this is always the case. This forced us to design the classifiers focusing on the transfer batch and not on the embryo itself, which adds further uncertainty to the Bayesian models applied. Furthermore, the fact of transferring embryos in batches or individually is known to provide different conditions for viable embryos. It must be noted that there exist studies in the literature confirming the fact that when transferring more than one single embryo at a time, even if not all of them manage to be implanted, the ones succeeding appear to be somehow helped by the others [13].

2.2. *Implantation success rate and transfer procedure*

The information on the proportion of embryos implanting is reflected in our case in the class variable, although we distinguish only between implantation obtained or not. We do this since the very clear outcome of lack of success is when no implantation is obtained, whereas the cases of having two or three implantation are regarded in clinical practise as a success too. Even if we have the information available and we initially planned to distinguish this information in case of success, we needed to estimate the cost of misclassifying cases between one, two or three implantations. However, when speaking to the clinical staff, we concluded that such a cost is very subjective to estimate, and therefore we decided to concentrate purely on implantation/not implantation this being considered the most important outcome from the clinical point of view. Reproduction unit of Clínica del Pilar, define a pregnancy (i.e. success rate) with an implantation verified by ultrasound.

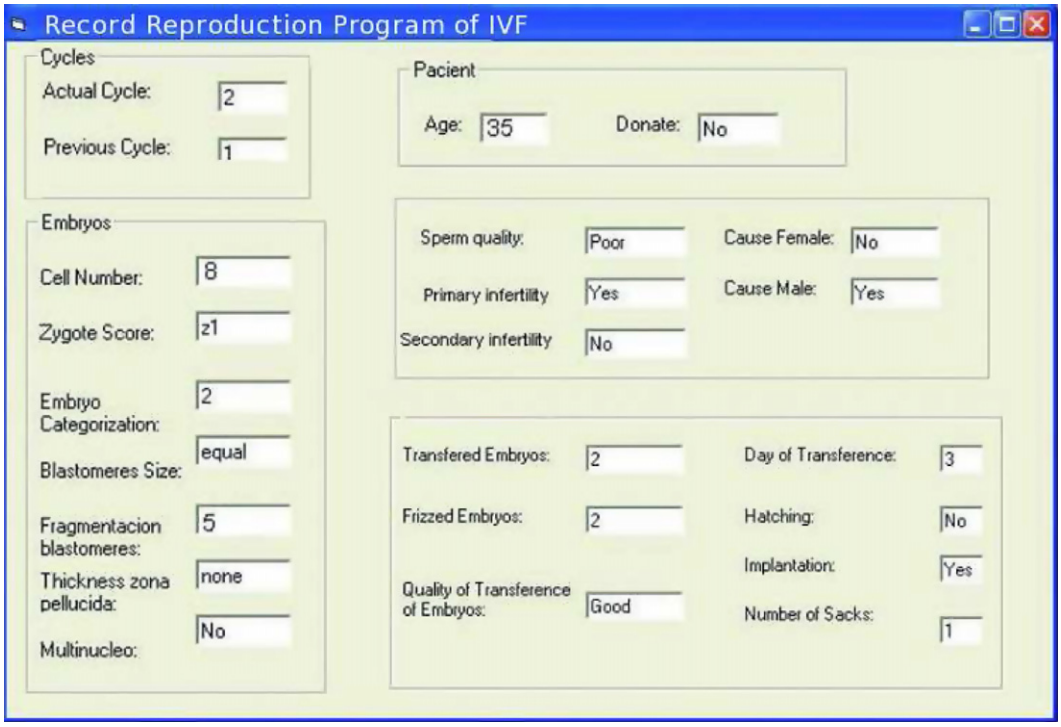


Fig. 2 – Screen capture of the software used to collect IVF patient clinical data as well as information from the embryos. The data entering screen is designed in a way that the predictor variables are organized into categories.

Regarding the transfer procedure, the normal practice in Clínica del Pilar's IVF unit is to select a batch of three embryos in each transfer, obtaining until September 2007 a take home baby rate of 35%. The transfer success is assumed when at least one of the embryos results in implantation (i.e. pregnancy by ultrasound study). However, since transfer is performed in groups, it is not possible to identify in our sample which of the transferred embryos is the one that resulted in pregnancy. This could only be done when all the transferred embryos are implanted, although this is very rare. Furthermore, the decision to transfer two or three embryos is motivated by the need to avoid multifetal pregnancies. That is why unless the cases in which the pregnancy expectation is lower the tendency is to transfer two embryos.

At this stage, and in order to obtain all the relevant data required for supplying the database with enough information for allowing a later classification model construction, we initially created a software to collect the data from real treatments that took place in this hospital. Fig. 2 shows the parameters recorded for this study per each of the IVF treatments, in which we can identify parameters related to morphology analysis of batches with three embryos, as well as clinical data of patients. This application allowed us to record new data of IVF treatments on the database in order to collect it in a procedural manner that fed our model-building paradigms with comparable data.

2.3. *Selection of the best embryos as a supervised classification problem*

In the present work, we approach the problem of human embryo selection for transfer using supervised classification techniques. The main goal of any supervised classification algorithm is to build a classification model using a concrete data set. Our particular data set is the database with all the variables described in the previous sections. All these variables of the database are considered predictor features, and each of the transfer is considered as a concrete case of the database, recording each of them as successful (i.e. pregnancy) or not, and in case of obtaining pregnancy, how many of the transferred embryos were implanted.

For our purpose, we took into account the cases of batches consisting of a total of three embryos that are transferred in each treatment. Therefore, each treatment batch is considered as a case to be classified as successful or not, and the predictor variables that are taken into account are the ones related to each of the embryos selected as well as clinical ones related to the infertile couple.

More formally, we denote by $\mathbf{x} = (x_1, \ldots, x_n)$ the vector of predictor variables that represents a concrete transfer batch in the database. A class variable C is defined, assigning each of the cases of the database to a concrete class. The value of the class to which a case belongs to is denoted by c and it takes values in the domain set $\{0, 1\}$: each pattern vector represents a batch of embryos transferred, assigned to class 1 if at least an implantation occurred (regardless of being just one or up to three implantations), and to class 0 otherwise. Note that we cannot know which embryo of the batch is implanted in each case to allow a study at embryo-level, which is the reason why the analysis of cases is performed at batch-level. However, since embryologists assign a score to each of the embryos of each transfer, these embryos are ordered from the most promising (highest scored) to the less, and therefore we respect this order priority in the batch pattern vector reflect-

Table 2 – Example of data in the final feature vector batch selection of implantation classification by Bayesian classifiers with three embryos per batch

Treatment	Embryo-related data			Clinical data patient	Class
	1	2	3		
Batch 1	Emb1	Emb2	Emb3	···	1
Batch 2	Emb1	Emb2	Emb3	···	0
Batch 3	Emb1	Emb2	Emb3	···	1
⋮	⋮	⋮	⋮	⋮	⋮
Batch *n*	Emb1	Emb2	Emb3	···	0

The class-variable C represents: 0 = 'no implantation'; 1 = 'embryo implanted'.

ing the higher probability of the most promising embryos to be the ones that reached implantation.

Following this, the embryo selection problem is therefore transformed in a supervised classification problem consisting on creating a model that assigns for any (new) batch of IVF treatment $\boldsymbol{x} = (x_1, \ldots, x_n)$ into one of the classes of variable C. Statistical classifiers such as Bayesian classifiers provide an estimate of $p(c|\boldsymbol{x})$, the probability that an IVF case with predictor vector $\boldsymbol{x} = (x_1, \ldots, x_n)$ belongs to class $c \in C$.

The cases to be classified as well as the training set are defined by their features. We can regard the classifier as a function $\gamma : (x_1, \ldots, x_n) \rightarrow \{0, 1\}$ that assigns labels to observations. In supervised classification the objective is to build a classifier that minimizes the total error cost by taking into account the joint probability distribution $p(x_1, \ldots, x_n, c)$ that is unknown *a priori*. According to [14], for the particular case of a symmetric function cost the total error cost is minimized by assigning the case $\boldsymbol{x} = (x_1, \ldots, x_n)$ to the class with the highest *a posteriori* probability:

$$\gamma(\boldsymbol{x}) = arg\max_c p(c|x_1, \ldots, x_n) \quad (1)$$

As a result of the conditions and limitations of the transfer procedure described in Section 2.2, a transfer is assumed to be successful if at least one implantation has been obtained.

As described previously, the number of embryos that each batch contains is three. Table 2 shows an example of the information structure per each batch (each case) in databases.

2.4. *Bayesian classifiers*

This paper investigates the ability of Bayesian classifiers to predict the success or not of a determined embryo-batch transfer procedure, by indicating for each IVF treatment the probability of implantation.[1] Obviously, this prediction is fully related on the quality of the embryos that have been decided to be included in this batch, as well as from the clinical context. Therefore, this decision is to be made based on the database of clinical and morphological data from embryos described in the previous section. The aim of intelligent methods is to support the selection of the most promising available embryos in order to choose the few of them which have good quality and the greatest potential for implantation. This selection problem is proposed as a supervised classification one which is based in the feature vector of clinical variables of IVF treatment, embryos' morphology, and their class (outcome). Bayesian classifiers have already demonstrated a good precision in complex medical problems [15]. Moreover, this models are transparent and comprehensive for medical practitioners. These reasons motivate their choice for this domain.

We present next some of the classifiers in the form of Bayesian networks that have been proposed in the literature [16]. Several paradigms in the form of Bayesian networks such as naive Bayes [17], selective naive Bayes [18], semi naive Bayes [19], tree augmented naive Bayes (TAN) [20] and k-dependence Bayesian classifier (kDB) [21] are thought specifically for supervised classification problems. The main characteristic that distinguishes them is the number of dependencies between predictor variables that each Bayesian classifier can take into account, which also determines its structural complexity.

A Bayesian classifier is usually constructed using a score to decide which configuration is better than another, except from the naive Bayes classifier which always has the same structure. The graphical representation of Bayesian classifiers makes it possible to understand the underlying probabilistic classification process and to provide a set of properties that can be directly interpreted by medical staff. At the same time, the conditional (in)dependence relationships between the features as well as the conditional and marginal probabilities of the model can be of interest to embryologist who want to better understand the uncertainty of the studied medical domain.

A model hierarchy of increased structural complexity can be established for different types of Bayesian classifiers, where the naive Bayes is at the bottom and a general Bayesian network is at the top of this hierarchy. Fig. 3 illustrates examples of some Bayesian classifier models grouped by the number of the dependencies between predictor variables that they are able to take into account.

2.4.1. *Naive Bayes*

The *naive Bayes* classifier [17,22,23] is an example of the simplest Bayesian supervised classification algorithm. In the *pattern recognition* community [14], the naive Bayes classifier is proposed for the first time in 1987 [24]. Gradually, the machine learning community realized on its potential and robustness for supervised classification problems.

The *naive Bayes* classifier assumes that all predictor variables are conditionally independent given the class C. This

[1] Please, note that in IVF, the final aim is to obtain an implantation, not a birth, since the causes for pregnancy interruption are not considered under this field.

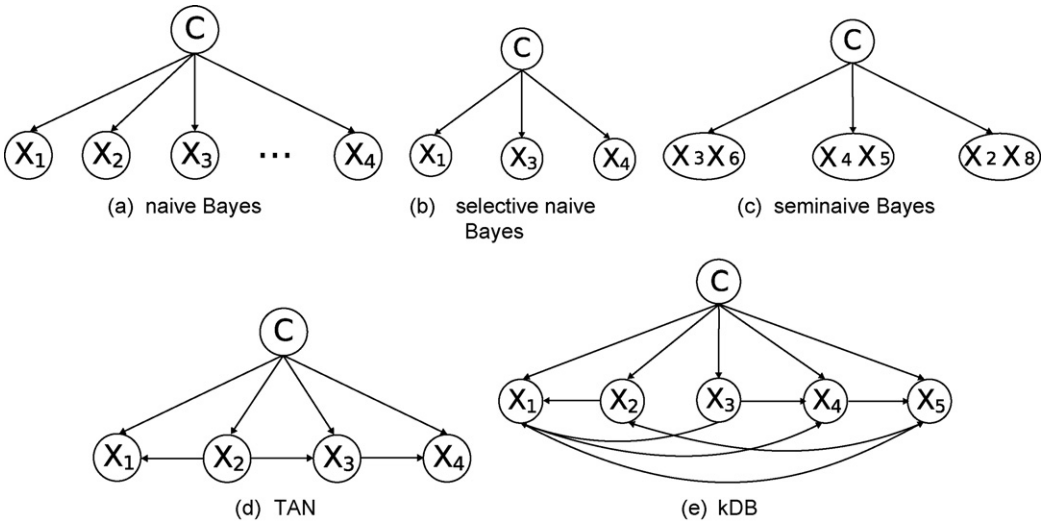


Fig. 3 – (a–e) Structure for Bayesian classifiers.

paradigm has always the same structure: all the predictor variables $X_1, \ldots, X_n$ are included in the model. The Fig. 3(a) shows the classifier structure in a problem with four predictor variables.

The naive Bayes classifier applies the Bayes theorem to predict for each unseen instance $\mathbf{x}$, the class $c \in C$ for which it has a higher a posteriori probability. This a posteriori probability is computed as

$$p(c|\mathbf{x}) \propto p(c, \mathbf{x}) = p(c)\prod_{i=1}^{n} p(x_i|c) \quad (2)$$

where $p(x_i|c)$ represents the conditional probability of x_i given that $C = c$ when all variables have discrete values. As a result, the naive Bayes classifier follows the following approach:

$$c^* = arg\max_c p(c)\prod_{i=1}^{n} p(x_i|c) \quad (3)$$

2.4.2. *Selective naive Bayes*

Despite the success of the naive Bayes classifier in some problems, in many real problem domains the predictive accuracy of learning algorithms is degraded by irrelevant predictor variables, where the information contribution is overlapped or repeated.

The naive Bayes classifier is robust with respect to irrelevant variables (i.e. variables that always have the same values in all classes), due to the assumption of the independence of the variables given the class C. On the other hand, it is very sensitive to correlated variables [25,26]. As a result, redundant variables (i.e. those in which all values appear similarly in the different classes and therefore do not reflect any difference between the features of the classes) decrease the accuracy of Bayesian classifier [18]. For this reason, a feature selection process (FSS) is required to remove those variables in order to obtain a new subset of predictor variables to induce the most efficient classifier.

The *selective naive Bayes* algorithm [18,27] is a combination of FSS and the naive Bayes classifier. The main difference between the *selective naive Bayes* approach and naive Bayes is that the former in its final model some of the predictive variables can be discarded and not been present. Furthermore, the need to build a structure for the Bayesian classifier constitutes an additional step that was not present in the naive Bayes classifier, and that in the case of the selective naive Bayes one is performed in the classical literature in two standard ways: forwardly starting with an empty set of variables and adding them one by one, or backwardly by removing in each iteration one of the variables that will be discarded. The forward sequential selection wrapper algorithm is one of the former possibilities, which stars with an empty set of variables. At each step the model adds the most accurate variable calculated by estimated accuracy [28] and stops when no improvement is obtained.

As an example of applying the selective naive Bayes classifier, if we consider the selective naive Bayes classifier illustrated in Fig. 3(b) as the representation of an instance $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5, x_6)$, this would be assigned to the class

$$c^* = arg\max_c p(c)p(x_1|c)p(x_3|c)p(x_4|c) \quad (4)$$

2.4.3. *Semi naive Bayes*

The selective naive Bayes algorithm is able to detect irrelevant and redundant variables, although no dependency between the variables present in the structure are taken into account. However, in most of real problems relationships between variables exist and need to be considered for a good classification performance. For this reason, other Bayesian classifiers overcome the assumption of conditional independence have been developed. The *semi naive Bayes* model [29] is an example of those.

The *semi naive Bayes* classifier is able to take into account the relationships between the variables $X_1, \ldots, X_n$ conditioned to the class variable C using a new type of variable: a joint

variable that is built as the result of the Cartesian product of a subset of variables. Since such a variable is represented as a single node in the Bayesian network, its consideration allows to surpass the assumption of conditional independence [19]. Each joint node represents a new variable that considers all the dependencies between the original variables that form it.

Once again, as in every Bayesian classifier but naive Bayes one, the learning of such a classifier requires an algorithm to build the Bayesian network structure. Ref. [19] proposes the induction of a semi naive Bayes classifier starting with an empty structure to which iteratively new nodes are added or new variables fused in existing nodes until non-improvement of the performance in terms of estimated accuracy is reached. As an example, Fig. 3(c) shows a possible semi naive Bayes model that could have been induced using this approach. Under this classifier, the pattern $\boldsymbol{x} = (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8)$ will be assigned to the following class:

$$c^* = arg\max_c p(c)p(x_3, x_6|c)p(x_4, x_5|c)p(x_2, x_8|c) \quad (5)$$

2.4.4. Tree augmented naive Bayes

The *tree augmented naive Bayes* (TAN) [20] is another Bayesian network classifier that allows dependencies between variables. The main restriction on the dependencies that can be considered by this Bayesian classifier is that each predictive variable can have a maximum of two parents: the class variable C and one of the other predictive variables $X_1, \ldots, X_n$.

In order to create the structure that will apply the TAN classifier, Friedman et al. [20] propose to follow the general outline of Chow and Liu's procedure [30]. Following this algorithm, the tree augmented naive Bayes structure is build in a two-phase procedure. Firstly, the dependencies between the different variables $X_1, \ldots, X_n$ are learned. This algorithm applies a score based on the information theory, and the weight of an arc (X_i, X_j) is defined by the mutual information measure conditioned to the class variable as

$$I(X_i, X_j|C) = \sum_c \sum_{x_i} \sum_{x_j} p(x_i, x_j, c) \log \frac{p(x_i, x_j|c)}{p(x_i|c)p(x_j|c)} \quad (6)$$

Using the mutual information of each predictive variable and the class $I(X_i, C)$, and the conditional mutual information of each pair of domain variables given the class $I(X_i, X_j|C)$, the algorithm builds a tree structure by adding at each step the arc between two variables X_i and X_j which has the highest $I(X_i, X_j|C)$ without forming a loop. This procedure adds a total of $n - 1$ arcs to the structure, forming a tree.

In a second phase the structure is augmented to the naive Bayes paradigm. Fig. 3(d) shows an example of a TAN classifier structure induced using this approach, where an instance $\boldsymbol{x} = (x_1, x_2, x_3, x_4)$ will be assigned to the class

$$c^* = arg\max_c p(c)p(x_1|c, x_2)p(x_2|c)p(x_3|c, x_2)p(x_4|c, x_3) \quad (7)$$

2.4.5. k-Dependence Bayesian classifier

The *k-dependence Bayesian classifier* (kDB) [21] tries to avoid the restriction of TAN structure where a predictive variable can have a maximum of two parents (the class and another predictive variable) allowing every predictive variable to have up to k parents besides the class. The main characteristic of a kDB structure is the fact that it is the user who fixes the restrictive condition of the value of k which represents the maximum number of parents per variable.

The kDB structure is built using $I(X_i|C)$ for each feature X_i, $i = 1, \ldots n$ and $I(X_i, X_j|C)$ with $i < j$, $j = 2, \ldots n$. The procedure starts with uniquely the class-node C in the structure. Each iteration, the algorithm selects the node not included in the structure with highest $I(X_i|C)$, the arc from C to X_i is added, and the value $I(X_i, X_j|C)$ is computed for all the possible new arcs from the X_j nodes already inserted in the structure. All these arcs are ordered from the highest to lowest and the highest k nodes are added to the structure (or all of them if the structure contains so far less or equal than k nodes excluding C).

Fig. 3(e) shows an example of kDB structure with five predictive variables and $k = 2$.

2.5. Filter approaches to induce structures for Bayesian classifiers

The filter approach to feature subset selection (FSS) consists of applying a function independent of the characteristics of the specific classifier for learning the Bayesian network, which has the advantage of a considerable reduction on the computation time required for the learning of the classifier. The mutual information is a commonly used measure for feature selection [31–33] and Bayesian classifier induction [20,21,34].

It is known that under the hypothesis of independence between X_i and C, $2NI(X_i, C)$ – where N denotes the size of the database – asymptotically follows a χ^2 distribution with $(r_i - 1)(r_0 - 1)$ degrees of freedom [35,36]. Thus, given the mutual information of a predictive variable and the class value, the χ^2-based test adapted by [15] can be performed to check the significance of the mutual information. However, we cannot apply a similar χ^2-based test for $2NI(X_i, X_j|C)$ since this distribution does not follow a χ^2 distribution.

We provide next a brief description of the filter approaches to FSS induction for the classification models selective naive Bayes, TAN and k DB Bayesian proposed in [15], which will later be applied to the supervised classification problem of the selection of best human embryo for transfer in IVF treatment.

The filter approach to *selective naive Bayes* (FSNB), is induced with only the subset of variables whose $2NI(X_i, C)$ surpasses the χ^2-test. Irrelevant variables disappear in the final selective naive Bayes classifier but FSNB cannot detect the dependencies between the different domain variables.

The *filter tree augmented naive Bayes* (FTAN) [15] is another novel Bayesian network classifier, in which a filter scheme is adopted to construct a TAN structure as proposed by [20]. The FTAN classifier is built over the subset of variables whose $2NI(X_i, C)$ surpasses the $\chi^2_{(r_i-1)(r_0-1);1-\alpha}$ value. At the same time, the forest structure is built following the Chow–Liu [30] algorithm.

The filter approach to a kDB classifier (FkDB), similarly as in the FTAN approach, applies mutual information and conditional mutual information regarding the class variable. The structure of the filter approach of Bayesian classifiers is build over the subset of variables whose $2NI(X_i, C)$ surpasses the χ^2-test. Every time an edge is included, the former property is checked, and only edges whose corresponding $2N_cI(X_i, X_j|C = c)$ surpasses for all $c \in C$ the value are finally added, where

Table 3 – Classifiers from three-embryo batch transfers in IVF treatment, with a total of 63 instances in databases

	Naive Bayes		Semi NB		Selective NB		TAN		k DB		FSNB		FTAN		lFKBB	
	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
Predicted class																
0	36	11	41	14	42	17	39	17	37	17	41	16	41	16	40	18
1	9	7	4	4	3	1	6	1	8	1	4	2	4	2	5	0
Accuracy	0.6825		0.7143		0.6825		0.6349		0.6032		0.6825		0.6825		0.6349	

The table below shows the results obtained by the filter Bayesian classifiers. The columns 0 and 1 in each classifier represent the true classes while the 0 and 1 lines are the class estimated by the classifier (e.g. in the case of Naive Bayes, there are 9 false positives and 11 false negatives when predicting pregnancy, which gives an accuracy of 68.25%).

N_c represents the number of cases of the database which are classified to the class $c \in C$.

As a final step in both FTAN and F*k* DB, the arcs between the class-variable *C* and all the variables present in the final structure are also included in the model—i.e. the structure is augmented with naive Bayes. Moreover, in the particular case of F*k* DB unconnected components are allowed in its structure, and *k* becomes the upper bound of the number of parents of a predictive variable.

3. Experimental results

We focus our empirical study on the accuracy of the proposed Bayesian classification models and their potential application to a support system for selecting the best combination of human embryos for transfer, and hence the best embryos to transfer.

The database used in this study contains continuous and categorical (discrete) variables, and each case corresponds to a treatment batch containing a total of three embryos. The Bayesian classifiers described in the previous section are implemented to manage uniquely discrete data. Therefore, as a pre-process step continuous variables of the dataset are discretized by the equal frequency algorithm [37] into two intervals. The Elvira software [38] is used in the implementation of the previously presented Bayesian classification models. In order to validate the Bayesian classification models a leave-one-out cross-validation method is performed [28], and the *estimated accuracy* for each classifier is computed using this method.

Table 3 shows the estimated accuracy and true positive, false negative, true negative and false positives values obtained for every Bayesian classifier applying the database of three-embryo batch that contains 63 instances. These are formed by a total of 18 and 45 cases of successful and unsuccessful treatments, respectively.

Table 4 shows the estimated accuracy, sensibility, specificity, implantation predictive value and no-implantation predictive value. It must be noted that the semi naive Bayes classifier obtained the best accuracy with a correct classification of the 71.43%, with 22.22% of sensibility and 91.11% of specificity while the there are four other Bayesian classifiers in the 68.25% of accuracy (naive Bayes, selective naive Bayes, FSNB, and FTAN). Naive Bayes classifier had 38.89% of sensibility and 80.00% of specificity and 43.75% of true predictive value. The worst accuracy obtained by Bayesian classifiers corresponded to *k* DB with 60.32% of accuracy although the most important finding is that it was not able to identify any single of the positive pregnancy cases of the database. The semi naive Bayes classifier which returned the best result is the one that has a structure with only two nodes: the first is formed by the predictor variables node formed by *Embryo1 – BlastomeresSize* × *Embryo1 – FragmentationBlastomeres* × *SpermQuality* and the second by *Embryo1 – Multinuclear* × *Embryo2 – ThicknessZonaPellucida* × *Embryo3 – BlastomeresSize*. This result is according to the experts' assumption on the need to take into account for the prediction clinical data as well as morphological data from the embryos [39–41].

McNemar's-test [42] was applied in order to evaluate the statistical significance of the difference in the performance between pairs of Bayesian classifiers. The McNemar's-test did not show any significant difference between pairs of Bayesian classifiers.

It must be noted that this accuracy considers equally false positives and false negatives. However, the error-cost effect might be considered differently if we take into account the

Table 4 – Results of three embryo-batch classification by Bayesian classifiers expressed as a percentage

Classifier	Accuracy	Sensibility	Specificity	Implantation predictive value	No-implantation predictive value
Naive Bayes	68.25	38.89	80.00	43.75	76.59
Semi naive Bayes	71.43	22.22	91.11	50.00	74.54
Selective naive Bayes	68.25	5.55	93.33	25.00	71.19
TAN	63.49	5.55	86.67	14.29	69.64
kDB	60.32	5.55	82.22	11.11	68.52
FSNB	68.25	11.11	91.11	33.33	71.92
FTAN	68.25	11.11	91.11	33.33	71.92
FkDB	63.49	0.00	88.89	0.00	68.96

Table 5 – List of variables included in the Bayesian classifiers from the databases of batches with three embryos

List of variables	Semi NB	Selective NB	FSNB	FTAN	FkDB
Embryol blastomere size	×		×	×	×
Embryol fragmentation	×				
Embryol thickness of zona pellucida			×	×	
Embryol multinuclear	×		×	×	×
Embryo2 thickness of zona pellucida	×		×	×	×
Embryo2 multinuclear			×	×	×
Embryo3 blastomere size	×		×	×	×
Embryo3 fragmentation		×			
Embryo3 thickness of zona pellucida			×	×	×
Embryo3 multinuclear			×	×	×
Sperm quality	×				
Primary infertility			×	×	×
Secondary infertility			×	×	×
Number of transferred embryos			×	×	×

psychological consequences of a false positive (i.e. a non promising embryo is classified wrongly leaded to a lack of pregnancy) are much more important than the false negative (i.e. a promising embryo is discarded although it was indeed promising one) due to patient-related impact. Nevertheless, there is no easy way of estimating such a cost difference as embryologist experts admit.

3.1. *Feature subset selection*

Another important aspect on our experiment was to measure the performance for examination and selection of subset features for semi naive Bayes, selective naive Bayes, FSNB, FTAN and F*k* DB Bayesian classifiers.

All the filter algorithms perform a complex inter-features analysis to consider initially all the variables, and the decision of removing one or another node in filter classifier-learning methods is automatically taken in terms of discarding in the classifier some redundant or irrelevant variables regarding others already present (always according to the data provided in our particular database). It is important to note that regarding the clinical usefulness, the fact of having removed a variable in a classifier does not imply that it is meaningless: all variables provide information, although according to our Bayesian classifier learning algorithms the model fits better the database when some of them are removed. The fact that the best classifiers are the ones that have some variables removed implies that some variables are appear to be more determinant to obtain implantation or not in an IVF treatment. Table 5 shows the subset of predictor variables that were found to be most relevant ones.

Our results of the subset of variables correspond to the opinion of embryologists, since they concentrate mainly on identifying two main features in order to judge the potentiality of an embryo for implantation, namely the *blastomeres size* and *grade of fragmentation*. These morphological embryo features have been observed during the embryo selection procedure by the embryologists in the Clínica del Pilar IVF unit. Furthermore, the blastomeres size and degree of fragmentation characteristics also play an important role on the ability of an embryo to progress to a clinical pregnancy [40,43]. The *Blastomere size*, is a feature related with rate of embryo growth (cleavage). Observations suggest that a high number of blastomeres corresponds to higher implantation rates.

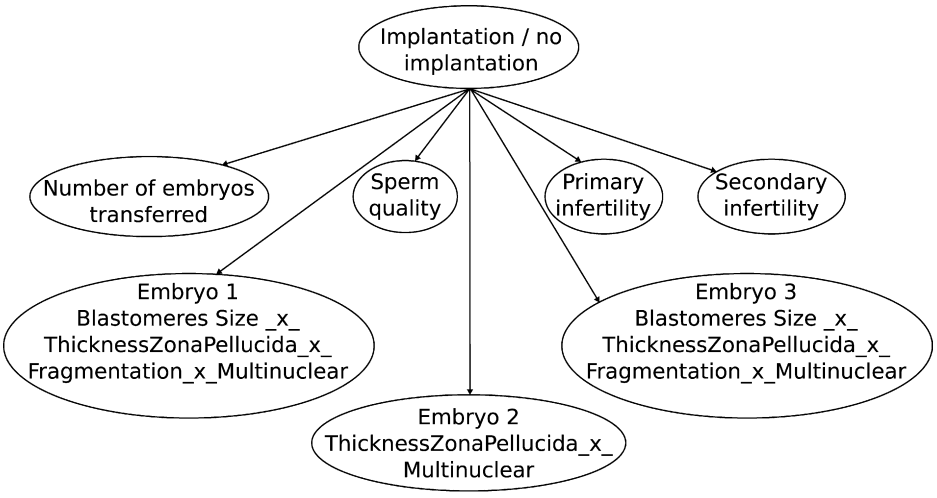
The feature subset of predictor variables that we propose in our work for the learning of Bayesian classifiers includes the two variables of blastomeres size and grade of fragmentation, apart from other such as the fact of the embryo being multinucleated or not, which is also fully according to the literature [39,40,44]. Another predictor variable, that we found, was the *thickness of zona pellucida*. This result is also coherent with the research of Gabrielsen et al. [41] which proves that zona pellucida thickness has a relationship with pregnancy on IVF outcome.

It must be noted that apart from variables fully related to the quality of each of the embryos, other clinical features have also been included on the Bayesian classifiers, namely the number embryos transferred, infertility being primary or secondary, and the sperm quality. These characteristics appear to be of importance by classifiers in order to discern between treatments leading to implant or not.

3.2. *The Bayesian classifier build from expert embryologist's experience*

Since in the problem of embryo selection for IVF transfer the choice of the embryos appears to be one of the most relevant features that determine the success of the treatment, and since this evaluation is fully subjective and very dependant on the experience and expertise of the embryologist, we decided to create manually a Bayesian classifier fully considering the experience transmitted by the embryologists, and using the database of cases only for the estimation of the parameters of the model. Our aim is to compare this classifier with the different algorithms to infer the Bayesian classifier's structure regarding the real cases.

All our preliminary tests showed that including the rest of predictor variables does not improve the performance (rather, it could even worsen it). Fig. 4 illustrates the structure of the Bayesian network as well as the table describing the classification performance showing that this manually built Bayesian classifier returned 18 false negative and 2 false positive, with a general accuracy of 68.25% with 95.55% of specificity and no-implantation predictive value of 70.49%. However, it must



		Real class 0	1
Predicted class	0	43	18
	1	0	0

Fig. 4 – Bayesian network with FSS variables suggested by the embryologists experts' experience. The table shows the classification results obtained, which means a general accuracy of 68.25%.

be noted that none of the pregnancy cases was identified, which advices us to consider this classifier worse than the Bayesian classifiers induced from data and lowers considerably the applicability of this classifier in real practice.

Even if we cannot forget that the accuracy and performance of the classifiers is fully dependent on the data of our particular database, the most direct applicability of this results is that some of the variables that are considered are indeed much more relevant to be able to predict the outcome of an IVF treatment of three batches than other ones. This being a field in which the subjectivity is key and therefore the training period of embryologists requires of as much help as possible to be able to increase the percentage of successful treatments. Our experimental results encourage embryologists to concentrate their efforts mainly on the variables that are described in the filtered classifiers that obtained the best score, although it is also important to perform further studies on the rest of variables too in order to measure their relative importance for the choice of the best embryos to transfer.

4. Conclusions and future work

This paper introduces a novel intelligent support system for embryologists in order to provide an estimation of the success of a concrete IVF treatment, which at the same time allows to choose the most promising embryos for a better outcome. This system applies Bayesian classifiers as a decision support system to this concrete problem, this being a paradigm that has been proved accurate and appropriate for this task.

The performance of Bayesian classifiers which take into account a subset of predictor variables such as the thickness of zona pellucida, degree of fragmentation, multinucleate and blastomere size obtain also quite good classification results. On the other hand, this subset of features is the same that expert embryologists take into account in normal practice and that literature on the field have draw attention upon.

Focusing on the potential applicability and impact of this work, we consider that in the near future this methodology could be applied as a decision support system by embryologist experts in the decision making on clinical practice, in order to aid on determining which information is relevant for the final success and to improve the process in the standardizing the selection of embryos for transfer. Such a system could also be able to be applied for training novel embryologists, this being an important contributing factor to the overall success rate of their treatments.

One of future work trends is on concentrating on other countries than Spain whose legislation does not allow transferring more than an embryo, since under this conditions it would be possible to proceed to a redefinition of the supervised classification problem fully embryo-centered and with a higher accuracy rates expected. That is why one of our future works concentrates on acquiring data of transfers of a single embryo, which would result in a more representative database for our classifiers and the possibility to obtain better classification rates.

Other future work trends include the acquisition of new data that includes other parameters not considered so far due to procedure-driven limitations, the possibility to include other Bayesian classifiers, as well as the consideration of applying a cost matrix in which false positives and false negatives have different error costs and their effect in the overall performance of the classifiers. In order to estimate

this cost ROC curves could be applied due to their cost-sensitiveness.

Acknowledgment

First, thanks to the helpful comments of the anonymous reviewers which have improved the quality of the paper. This work was supported by the projects *SAIOTEK* and *ETORTEK* of the Basque Government, and to the project *TIN2005-03824* of the Spanish Ministry of Education and Science.

REFERENCES

[1] T. Baczkowski, R. Kurzawa, W. Glabowski, Methods of embryo scoring in in vitro fertilization, Reprod. Biol. 4 (2004) 5–22.
[2] R. Saith, A. Srinivasan, D. Michie, I. Sargent, Relationships between the developmental potential of human in-vitro fertilization embryos and features describing the embryo, oocyte and follicle, Human Reprod. Update 4 (1998) 121–134.
[3] I. Jurisica, J. Mylopoulos, J. Glasgow, H. Shapiro, R. Casper, Case-based reasoning in IVF: prediction and knowledge mining, Artif. Intel. Med. 12 (1998) 1–24.
[4] J.R. Trimarchi, J. Goodside, L. Passmore, T. Silberstein, L. Hamel, L. Gonzalez, Comparing data mining and logistic regression for predicting IVF outcome, Fertil. Steril. 80 (2003), 100–100.
[5] G. Patrizi, C. Manna, C. Moscatelli, L. Nieddu, Pattern recognition methods in human-assisted reproduction, Int. Trans. Oper. Res. 11 (2004) 365–379.
[6] C. Manna, G. Patrizi, A. Rahman, H. Sallam, Experimental results on the recognition of embryos in human assisted reproduction, Reprod. BioMed. Online 8 (2004) 460–469.
[7] L. Scott, R. Alvero, M. Leondires, The morphology of human pronuclear embryos is positively related to blastocyst development and implantation, Hum. Reprod. 15 (2000) 2394–2403.
[8] J. Tesarik, E. Greco, The probability of abnormal preimplantation development can be predicted by a single static observation on pronuclear stage morphology, Hum. Reprod. 14 (1999) 1318–1323.
[9] J.D. Fisch, H. Rodriguez, R. Ross, G. Overby, G. Sher, The graduated embryo score (GES) predicts blastocyst formation and pregnancy rate from cleavage-stage embryos, Hum. Reprod. 16 (2001) 1970–1975.
[10] P. Terriou, C. Sapin, C. Giorgetti, E. Hans, J.L. Spach, R. Roulier, Embryo score is a better predictor of pregnancy than the number of transferred embryos or female age, Fertil. Steril. 75 (2001) 525–531.
[11] U. Zollner, K.P. Zollner, G. Hartl, J. Dietl, T. Steck, The use of a detailed zygote score after IVF/ICSI to obtain good quality blastocysts: the German experience, Hum. Reprod. 17 (2002) 1327–1333.
[12] L.L. Veek, An Atlas of Human Gametes and Conceptuses: An Illustrated Reference for Assisted Reproductive Technology, Parthenon, 1999.
[13] R. Matorras, F. Matorras, R. Mendoza, M. Rodríguez, J. Remohí, F.J. Rodríguez-Escudero, C. Simón, The implantation of every embryo facilitates the chances of the remaining embryos to implant in an IVF programme: a mathematical model to predict pregnancy and multiple pregnancy rates, Hum. Reprod. 20 (2005) 2923–2931.
[14] R. Duda, P. Hart, Pattern Classification and Scene Analysis, John Wiley and Sons, New York, 1973.
[15] R. Blanco, I. Inza, M. Merino, J. Quiroga, P. Larrañaga, Feature selection in Bayesian classifiers for the prognosis of survival of cirrhotic patients treated with TIPS, Biomed. Inform. 38 (2005) 376–388.
[16] J. Pearl, Probabilistic Reasoning in Intelligence Systems, Morgan Kaufmann, Los Altos, CA, 1988.
[17] M. Minsky, Steps toward artificial intelligence, Trans. Inst. Radio Eng. 49 (1961) 8–30.
[18] P. Langley, S. Sage. Induction of selective Bayesian classifiers, in: Proceedings of the 10th Conference on Uncertainty in Artificial Intelligence, Seattle, WA, 1994, pp. 399–406.
[19] M. Pazzani, Searching for dependencies in Bayesian classifiers, in: D. Fisher, H.-J. Lenz (Eds.), Learning from Data: Artificial Intelligence and Statistics V, Springer-Verlag, New York, NY, 1997, pp. 239–248.
[20] N. Friedman, D. Geiger, M. Goldsmidt, Bayesian network classifiers, Machine Learn. 29 (1997) 131–163.
[21] M. Sahami, Learning limited dependence Bayesian classifiers, in: Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining, 1996, pp. 335–338.
[22] C. Ohmann, Q. Yang, M. Kunneke, H. Stolzing, K. Thon, W. Lorenz, Bayes theorem and conditional dependence of symptoms: different models applied to data of upper gastrointestinal bleeding, Meth. Inform. Med. 27 (1988) 73–83.
[23] I. Kononenko, Comparison of inductive and naïve Bayesian learning approaches to automatic knowledge acquisition, in: B. Wielinga, J. Boose, B. Gaines, G. Shereiber, M. van Someren (Eds.), Current Trends in Knowledge Acquisition, IOS Press, Amsterdam, 1990, pp. 190–197.
[24] B. Cestnik, I. Kononenko, I. Bratko, ASSISTANT-86: a knowledge elicitation tool for sophisticated users, in: I. Bratko, N. Lavrac (Eds.), Progress in Machine Learning, Sigma Press, Wilmslow, U.K., 1987, pp. 31–45.
[25] H. Liu, H. Motoda, Feature Selection for Knowledge Discovery and Data Mining, Kluwer Academic Publishers, 1998.
[26] I. Inza, P. Larrañaga, R. Etxeberria, B. Sierra, Feature subset selection by Bayesian network-based optimization, Artif. Intel. 123 (2000) 157–184.
[27] R. Kohavi, G. John, Wrappers for feature subset selection, Artif. Intel. 97 (1997) 273–324.
[28] M. Stone, Cross-validatory choice and assessment of statistical predictions, J. Roy. Stat. Soc. Ser. B 36 (1974) 111–147.
[29] I. Kononenko, Semi-naïve Bayesian classifiers, in: Proceedings of the 6th European Working Session on Learning, 1991, pp. 206–219.
[30] C. Chow, C. Liu, Approximating discrete probability distributions with dependence trees, IEEE Trans. Inform. Theory 14 (1968) 462–467.
[31] D. Lewis, Naive (Bayes) at forty: the independence assumption in information retrieval, in: Proceedings of the 10th European Conference on Machine Learning, 1998.
[32] A.L. Blum, P. Langley, Selection of relevant features and examples in machine learning, Artif. Intel. 97 (1997) 245–271.
[33] M. Zaffalon, The naive credal classifier, J. Stat. Plan. Inf. 105 (2002) 5–21.
[34] R. Duda, P. Hart, D.G. Stork, Pattern Classification, Wiley Interscience, 2001.
[35] D. McKay, Information Theory, Inference, and Learning Algorithms, Cambridge University Press, 2003.
[36] T.M. Cover, J.A. Thomas, Elements of Information Theory, John Wiley and Sons, 1991.
[37] K. Dougherty, R. Kohavi, M. Sahami, Supervised and unsupervised discretization of continuous features, in: Proceedings of the Twelfth International Conference on Machine Learning, 1995, pp. 194–202.

[38] Elvira Consortium. Elvira: an environment for creating and using probabilistic graphical models, in: Proceedings of the 1st European Workshop on Probabilistic Graphical Models, Cuenca (Spain), 2002, pp. 222–230.
[39] S. Ziebe, K. Petersen, S. Lindenberg, A.G. Andersen, A. Gabrielsen, A.N. Andersen, Embryo morphology or cleavage stage: how to select the best embryos for transfer after in-vitro fertilization, Hum. Reprod. 12 (1997) 1545–1549.
[40] M. Alikanin, J. Cohen, G. Tomkin, J. Garrisi, C. Mack, R.T. Scott, Human embryo fragmentation in vitro and its implications for pregnancy and implantation, Fertil. Steril. 71 (1999) 836–842.
[41] A. Gabrielsen, S. Lindenberg, K. Petersen, The impact of the zona pellucida thickness variation of human embryos on pregnancy outcome in relation to suboptimal embryo development. A prospective randomized controlled study, Hum. Reprod. 16 (2001) 2166–2170.
[42] T.G. Dietterich, Approximate statistical tests for comparing supervised classification learning algorithms, Neural Comput. 10 (1998) 1895–1924.
[43] D. Sakkas, G. Percival, Y. DÁrcy, K. Sharif, M. Afnan, Assessment of early cleaving in vitro fertilized human embryos at the 2-cell stage before transfer improves embryo selection, Fertil. Steril. 76 (2001) 1150–1156.
[44] S. Shen, C. Wong, K. Ho, T.L. Telles, V.Y. Fujimoto, M.I. Cedars, The morphology of 2 pronuclear (2PN) embryos is related to the quality of day 3 embryos, Fertil. Steril. 78 (2002) S52–S53.