Computers in Biology and Medicine ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

A Bayesian network model for predicting pregnancy after in vitro fertilization

Q1 G. Corani [a,*], C. Magli [b], A. Giusti [a], L. Gianaroli [b], L.M. Gambardella [a]

[a] Istituto Dalle Molle Intelligenza Artificiale (IDSIA), Manno, Switzerland
[b] International Institute for Reproductive Medicine (IIRM), Lugano, Switzerland

ARTICLE INFO

Article history:
Received 10 January 2013
Accepted 28 July 2013

Keywords:
In vitro fertilization (IVF)
Bayesian networks
EM algorithm
MAP estimation
Classification

ABSTRACT

We present a Bayesian network model for predicting the outcome of in vitro fertilization (IVF). The problem is characterized by a particular missingness process; we propose a simple but effective averaging approach which improves parameter estimates compared to the traditional MAP estimation. We present results with generated data and the analysis of a real data set. Moreover, we assess by means of a simulation study the effectiveness of the model in supporting the selection of the embryos to be transferred.

1. Introduction

According to the World Health Organization, infertility affects more than 80 million people worldwide; in vitro fertilization (IVF) is a treatment for addressing this problem. In IVF, a semen specimen is merged with a female egg in the laboratory to eventually generate an *embryo*. Whenever possible, multiple embryos are cultured for each woman. Embryos are cultured for 2–5 days, before being transferred to the woman. During the culture, the morphology of each embryo is monitored at fixed time intervals; embryos with certain morphologies have indeed high implantation potential [1–3] and are thus graded as of *top* quality. Despite the effort for designing effective scoring system for the embryos [2], predicting blastocyst development remains a challenging problem [4], although promising results have been recently obtained by analyzing time-lapse embryo images collected by automated image monitoring systems [5,6].

Reliably predicting the IVF outcome is thus still substantially an open problem [7–9]. A pioneering approach for estimating the probability of single and multiple pregnancy after an IVF treatment is the EU model [10], which assumes that, for pregnancy to happen, both a *receptive* uterus and a *viable* embryo are necessary. We represent *uterine receptivity*[1] as the binary variable U, with states $\{u, \neg u\}$ (u denoting receptivity, $\neg u$ non-receptivity); we represent *embryo viability* as the binary variable E, with states $\{e, \neg e\}$ (e denoting viability, $\neg e$ non-viability).

We denote by θ_e and θ_u respectively the probabilities of the embryo to be viable and of the uterus to be receptive, namely $\theta_e = P(E = e)$ and $\theta_u = P(U = u)$. The EU model estimates the probability of pregnancy after the transfer of a *single* embryo as $\theta_e\theta_u$, thus assuming the independence of viability and receptivity. When dealing with the transfer of *multiple* embryos, each embryo is assumed to implant independently from the others. For instance, if *two* embryos are transferred, the probability of a single pregnancy is $2\theta_u\theta_e(1-\theta_e)$, accounting for the fact that two embryos can give rise to pregnancy; the probability of double pregnancy is instead $\theta_e^2\theta_u$. The *EU assumption* is thus that if the uterus is not receptive, no pregnancy will follow; if the uterus is receptive, k babies will be born where k is the number of viable embryos among the transferred ones. The main limitation of the original EU model is the unrealistic assumption of θ_e and θ_u being identical for respectively all embryos and all women. Therefore, in [11] the model has been reworked (adopting a generalized linear model framework) by letting vary both θ_u and θ_e on external covariates; in particular, by letting θ_u depend on the age of the woman and θ_e on the number of cells present in the embryo at a given day (this is a marker of implantation capability). More recently it has been investigated [12] how to select the number and the types of covariates on which θ_u and θ_e should depend. In fact, quantifying how uterine receptivity and embryo viability vary as a function of respectively e.g. the age of the woman or the embryo score can provide important insights into domain experts.

However, analyzing the IVF data under the EU assumption implies a *partial observability* problem. For instance, if pregnancy does *not* occur, it cannot be ascertained whether (a) the uterus was

Q3 * Corresponding author. Tel.: +41 58 666 6509.
E-mail address: giorgio@idsia.ch (G. Corani).

[1] A more appropriate terminology could be *maternal receptivity*: this would highlight the broader meaning of receptivity, which includes factors associated with the patient characteristics and the treatment cycle. We keep however throughout this paper the term uterine receptivity for consistency with previous works in this area.

0010-4825/$ - see front matter © 2013 Published by Elsevier Ltd.
http://dx.doi.org/10.1016/j.compbiomed.2013.07.035

non-receptive, (b) *all* the transferred embryos were *non-viable* or (c) both. If pregnancy occurs, the uterus is known to be receptive, but it is still unknown which of the embryos gave rise to the pregnancy, unless the number of babies equals the number of transferred embryos. The missingness process is MAR (missing at random) and thus the parameters can be learned via the Expectation-Maximization (EM) algorithm [11,8].

In the Bayesian setting, EM is typically used to identify the parameter values which maximize (although only in a local fashion) the posterior probability of the data; this is the so-called MAP (most probable a posteriori) estimation. When dealing with incomplete samples, the fully Bayesian estimation of the parameters (which requires integrating over the posterior distribution of the parameters rather than finding its maximum) is not feasible. MAP estimation is also feasible with incomplete samples, but it "*does not offer the same benefits as a full Bayesian estimation. It does not attempt to represent the shape of the posterior and thus does not differentiate between a flat posterior and a sharply peaked one. As such, it does not give us a sense of our confidence in different aspects of the parameters, and the predictions do not average out our uncertainty*." [13, Section 17.4.4].

In a previous publication [14] we have introduced a novel probabilistic model of IVF transfers, which is a Bayesian network model based on the EU assumption. In [15] we have proposed a simple but effective averaging approach for estimating the parameters of the model from incomplete samples, which improves over the traditional MAP estimation.

In this paper we extend the analysis of [15], dealing with models which contain more variables than previously considered. Novel experiments confirm that the averaging methodology yields better parameter estimates than MAP estimation. Moreover, we compare the proposed model with state-of-the-art classification algorithms in the analysis of a data set containing IVF cycles performed at IIRM (International Institute for Reproductive Medicine) of Lugano. Eventually, we investigate via simulation the effectiveness of the model in supporting the decision of which embryos to transfer to the woman. Such a decision is typically difficult: it entails a trade-off between *maximizing* the probability of *single* pregnancy and *minimizing* the probability of *multiple* pregnancy (which is dangerous for the health of both mother and babies). In particular, we compare via simulation the outcome of the decisions taken on the basis of the model predictions and the outcome of the single-embryo transfer, which prevents multiple pregnancy but increases at the same time the probability of no-pregnancy [16,17].

2. The Bayesian network model

Given a generic variable X, we denote by θ_X the probability mass function which associates a marginal probability to each different value of X; we denote by $\theta_{X|Pa(X)}$ the probability mass function which associates a conditional probability to each different value of X, given each possible configuration of the parents of X, denoted as Pa(X). We moreover denote by $\boldsymbol{\theta}$ the set of all the parameters of the BN model.

As a first proposal, we represent the IVF transfer by the BN_1 *structure* shown in Fig. 1; a structure is a directed graph which connects the nodes representing the variables. The model manages IVF cycles with up to three embryos, as this is the maximum allowed under the Swiss law; however, it can be straightforwardly extended to manage a higher number of transferred embryos. The woman age is discretized as $\{<34, 34\text{–}40, 40+\}$.

We denote by $\mathcal{S}$ the set of nodes $\{S_1, S_2, S_3\}$; in the following, they are referred to as the $\mathcal{S}-$nodes. Such nodes take values in {*no-transfer*, *ntop*, *top*, *toph*} and thus represent the score of the embryos; ntop stands for non-top and toph for top-history. The

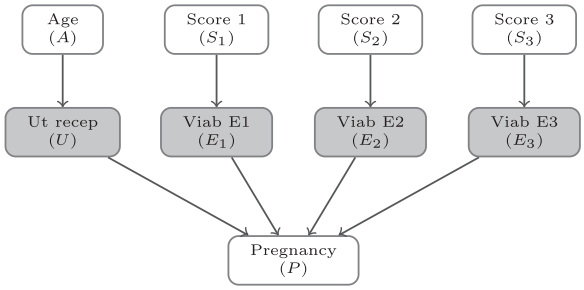

Fig. 1. The BN_1 structure: nodes affected by the missingness process are shown with a gray background. Each node contains its full name and, within parentheses, its abbreviation.

no-transfer state allows to model cycles with less than 3 transferred embryos: in most cycles only 1 or 2 embryos are transferred in order to reduce the danger of multiple pregnancy. Notice that the different positions (1,2,3) are randomly assigned to the embryos.

The $\mathcal{S}-$nodes are *tied*: they share the same mass function θ_S instead of having separate mass functions θ_{S_1}, θ_{S_2} and θ_{S_3}. This prevents the same embryo score (e.g., top) having a different marginal probability depending on whether one refers to node S_1, S_2 or S_3.

Node U represents uterine receptivity; it is therefore binary, with states $(u, \neg u)$.

We denote by $\mathcal{E}$ the set of nodes $\{E_1, E_3, E_3\}$, which are referred to in the following as $\mathcal{E}-$nodes. Each $\mathcal{E}-$node represents the viability of a different embryo; each $\mathcal{E}-$node is thus binary with states $(e, \neg e)$. The $\mathcal{E}-$nodes share the parameter set of the conditional mass function $\theta_{E|S}$, rather than having independent mass functions $\theta_{E|S_1}$, $\theta_{E|S_2}$ and $\theta_{E|S_3}$. Again, this prevents two embryos with the same score being given different probability of being viable just because they occupy a different position.

The pregnancy node P has four states $\{0, 1, 2, 3\}$, corresponding to the number of babies which might be born after having transferred up to three embryos. The CPT (conditional probability table) of P encodes the EU assumption; namely if the uterus is not receptive, no pregnancy will follow; if instead the uterus is receptive, k babies will be born where k is the number of viable embryos among the transferred ones. For instance, given a receptive uterus and two viable embryos out of three transferred, the CPT of node P assigns probability 1 to the outcome $P=2$ and probability 0 to all the remaining outcomes. In other words, the CPT of P assigns probability 1 to the pregnancy outcome whose value equals $(U = u) \cdot \sum_{i=1}^{3}(E_i = e)$.

2.1. The missingness process

In the following we describe the missingness process which affects receptivity and viabilities. The missingness process (MP) turns the complete data into incomplete according to a certain probability. The missingness process is MAR (*missing at random*) if the probability of a certain value to be turned into missing is independent of the value itself, although it can depend on other observed variables [18, Chapter 21]. As an example, consider a clinical practice in which test A is always observed while test B is performed only if test A is positive. Thus, B is missing whenever A is negative. Given the observed outcome of A, the probability of B to be missing does not depend on the value of B itself. The missingness process is instead MCAR (*missing completely at random*) if the distribution of the missingness process is independent of both the missing and the observed values. Thus, MCAR is a particular case of MAR.

Training stage: Let us consider an IVF cycle in which all the 3 embryos are transferred. At the training stage, the class variable P is always observed. In case of no-pregnancy ($P=0$), it is unknown

Please cite this article as: G. Corani, et al., A Bayesian network model for predicting pregnancy after in vitro fertilization, Comput. Biol. Med. (2013), http://dx.doi.org/10.1016/j.compbiomed.2013.07.035