Validation on Real Data of an Extended Embryo-Uterine Probabilistic Graphical Model for Embryo Selection

Adrián Torres Martín

Supervisors: Jerónimo Hernández-González
Jesús Cerquides

Universitat de Barcelona, July 2021

Overview

ART. Embryo selection

Artificial Reproductive Techniques. Process

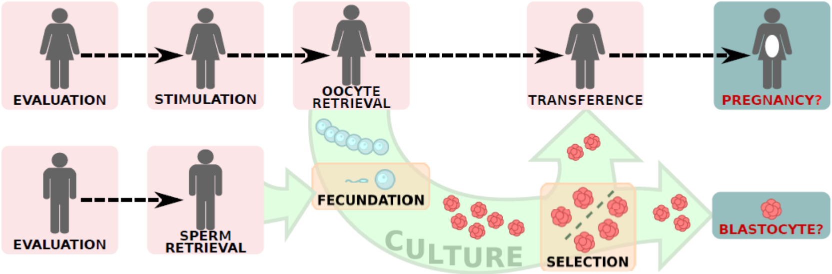


Figure: Process of an ART cycle. Source: J. Hernández-González

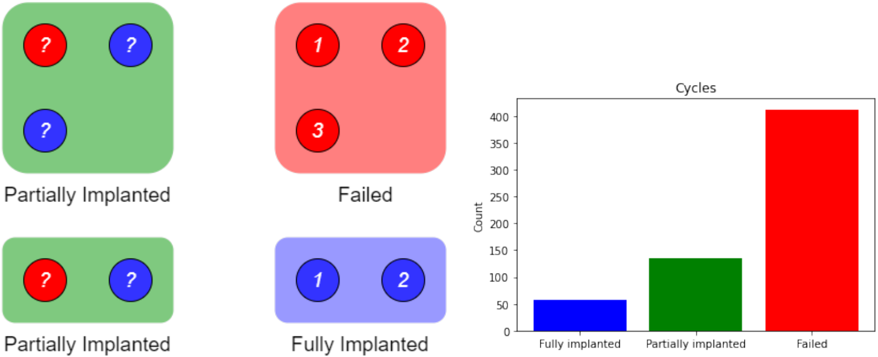


Figure: Different outcomes of a cycle. For partially implanted cycles, the identity of the implanted embryos is unknown.

Cycle characteristics

Cycle dataset with 25 features, related to:

- The female patient. Including previous undergone cycles, pregnancies and abortions. And current indicators: age, BMI, quantity of various hormones, etc.
- The male patient (quality of sperm).
- The stimulation procedure
- A summary of embryos (e.g., number of obtained embryos).

Embryo characteristics

Embryo dataset with 20 features, including:

- Fertilization technique (IVF or ICSI)
- Morphological characteristics
  - Related to oocytes
  - Embryo at D+1
  - Embryo at D+2
- ASEBIR score (A,B,C,D)

Figure: Embryos at D+2 with different number of cells.

Effect of ASEBIR score

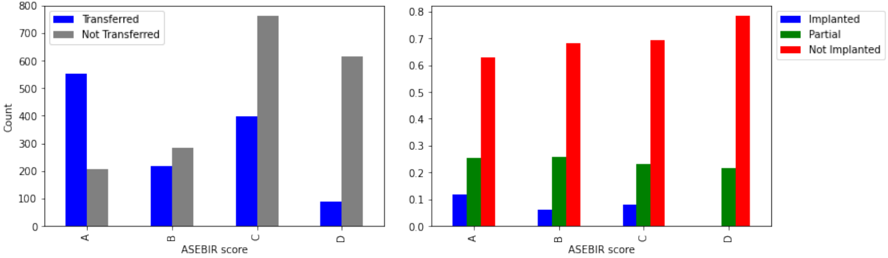


Figure: Effect of the ASEBIR score on the transfer selection (left) and on the implantation rate of those transferred (right).

Probablistic graphical model

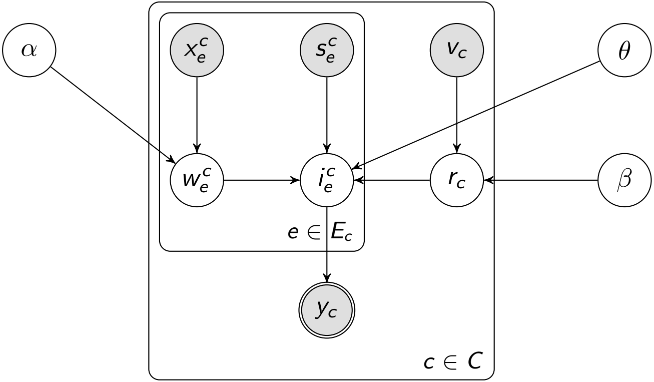
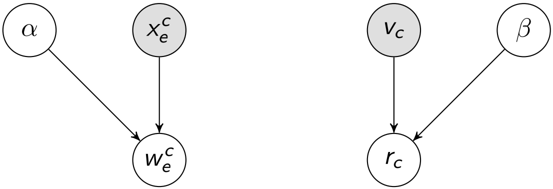


Figure: Graphical description of the proposed model. Shadowed nodes represent observed variables. Double line denotes a deterministic variable.

$$p(\mathbf{x}, \mathbf{w}, \mathbf{v}, \mathbf{r}, \mathbf{s}, \mathbf{i}, \mathbf{y}; \alpha, \beta, \theta) = p(\mathbf{w}|\mathbf{x}; \alpha)p(\mathbf{x})p(\mathbf{r}|\mathbf{v}; \beta)p(\mathbf{v})p(\mathbf{s})p(\mathbf{y}|\mathbf{i})p(\mathbf{i}|\mathbf{w}, \mathbf{r}, \mathbf{s}; \theta) \tag{1}$$

Probabilistic graphical model



$p(w_e^c|x_e^c, \alpha)$

$p(r_c|v_c, \beta)$

Probabilistic classifiers

- Logistic Regression
- Random Forest
- Extremely Randomized Trees
- Gradient Boosting

EM Algorithm. General setting.

Combines the completion (expectation) of the latent variables with the estimation of the hyperparameters.

E-step

$$Q(\theta;\theta_t) := \mathbb{E}_{Z\sim p(z|X;\theta_t)}[l(\theta;X,Z)] \tag{2}$$

M-step

$$\theta_{t+1} := \operatorname{argmax}_\theta Q(\theta;\theta_t) \tag{3}$$

In our case, $\theta = (\alpha, \beta, \theta)$ and $Z = (w_e^c, i_e^c, r_c)$

EM Algorithm. Implementation

Initialization: Random assignment of weights.

E-step Computation of weights: $q(w_e^c = w)$, $q(r_c = r)$ and $q(\mathbf{i}^c = \mathbf{i})$ (probability of obtaining each latent variable value, taking into account the whole model with current hyperparameters).

M-step Update hyperparameters with current weights.

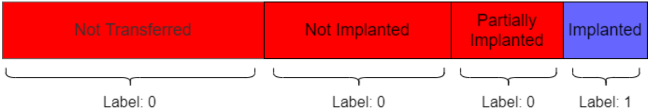
- Re-training probabilistic classifiers $\longrightarrow \hat{\alpha}, \hat{\beta}$.
- Maximizing conditional log-likelihood $\longrightarrow \hat{\theta}_1$.

Repeat E and M steps until convergence.

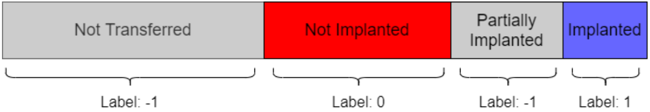
Experimental setup

Baseline methods

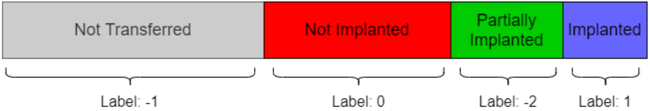
- Baseline 0 and Baseline cycles



- Naive EM



- EM with label proportions



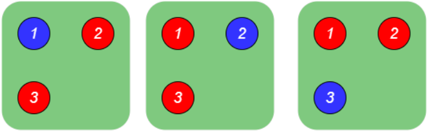
Baseline EM with label proportions

- Probability of a given configuration $\mathbf{i}^c$

$$p(\mathbf{i}^c|x;\alpha) = \prod_{e\in S_c} p(C_e = i_e^c|x_e;\alpha) \tag{4}$$

- Conditional probability given number of implanted embryos y_c

$$q(C_e = d) = \frac{\sum_{i^c\in \mathbb{I}_{y_c}} p(\mathbf{i}^c|x;\alpha)\mathbb{1}[i_e^c = d]}{\sum_{i^c\in \mathbb{I}_{y_c}} p(\mathbf{i}^c|x;\alpha)} \tag{5}$$



$\mathbf{i}^c = [1,0,0]$ $\mathbf{i}^c = [0,1,0]$ $\mathbf{i}^c = [0,0,1]$

Evaluation

- AUC-ROC
- LP-loss: Mean difference between true and predicted y_c (number of implantations in cycle c).
- Negative log-likelihood:

$$\mathcal{L}(\mathbf{Y}; \alpha, \beta, \theta) = -\frac{1}{B} \sum_{c=1}^{B} \sum_{j=0}^{N_c} \mathbb{1}[y_c = j] \log p(y_c), \tag{6}$$

where $p(y_c)$, the probability of cycle c having y_c implanted embryos, is,

$$p(y_c) = \sum_{i^c \in \mathbb{I}_{y_c}} \prod_{e} [i_e^c p(i_e^c = 1) + (1 - i_e^c) p(i_c^c = 0)] \tag{7}$$

Results and validation

Effect of ASEBIR score

Table: Metrics and control measures obtained using 5-fold cross validation

Model	Classifier	AUC	lp_loss	loglikelihood
Full Model	ETREES	0.64 ± 0.07	$\mathbf{0.54} \pm \mathbf{0.05}$	1.45 ± 1.59
	GBOOST	$\mathbf{0.71} \pm \mathbf{0.04}$	0.72 ± 0.03	0.45 ± 0.05
	LR	0.63 ± 0.08	0.60 ± 0.05	0.51 ± 0.10
	RF	$\mathbf{0.71} \pm \mathbf{0.05}$	0.80 ± 0.05	$\mathbf{0.42} \pm \mathbf{0.07}$
Full Model (Hidden quality)	ETREES	0.64 ± 0.05	$\mathbf{0.54} \pm \mathbf{0.05}$	1.27 ± 1.57
	GBOOST	$\mathbf{0.73} \pm \mathbf{0.07}$	0.73 ± 0.07	0.43 ± 0.06
	LR	0.62 ± 0.08	0.64 ± 0.07	0.52 ± 0.10
	RF	0.71 ± 0.05	0.80 ± 0.05	$\mathbf{0.42} \pm \mathbf{0.07}$

Effect of ASEBIR score

Table: Estimated parameter θ_1 for the four different classifiers.

Model	Classifier	θ_1
Full Model	ETREES	0.60 ± 0.04
	GBOOST	0.49 ± 0.00
	LR	0.52 ± 0.01
	RF	0.48 ± 0.01
Full Model (Hidden quality)	ETREES	0.58 ± 0.04
	GBOOST	0.49 ± 0.01
	LR	0.51 ± 0.00
	RF	0.48 ± 0.01

Effect of ASEBIR score

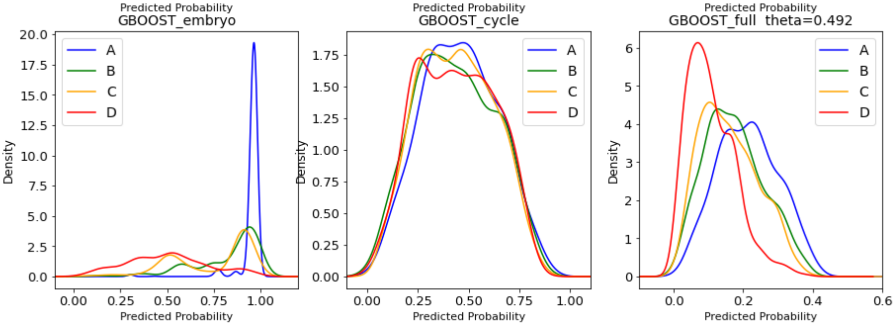


Figure: Predicted probability depending on the ASEBIR score.

Separation by true outcome.

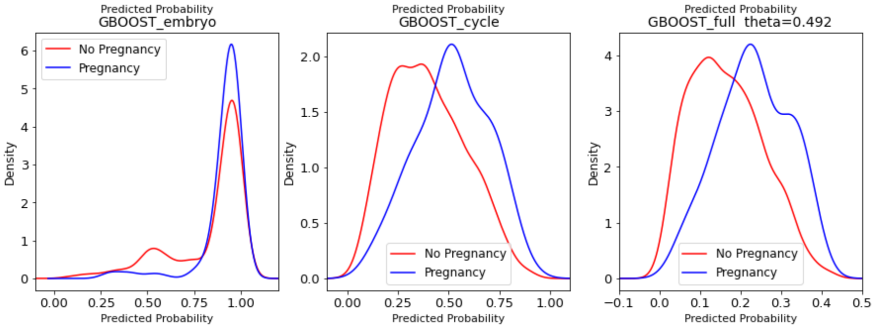
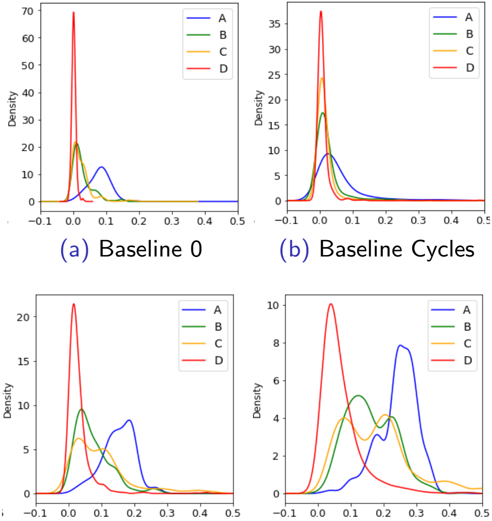


Figure: Predicted probability depending on the true outcome.

Comparison with baseline methods

Method	Classifier	AUC	lp_loss	loglikelihood
Baseline_0	ETREES	0.64 ± 0.07	$\mathbf{0.20 \pm 0.05}$	∞
	GBOOST	0.62 ± 0.05	0.21 ± 0.05	0.67 ± 0.26
	LR	0.58 ± 0.06	0.21 ± 0.05	0.63 ± 0.18
	RF	0.61 ± 0.06	$\mathbf{0.20 \pm 0.05}$	0.62 ± 0.18
Baseline_cycles	ETREES	0.62 ± 0.05	$\mathbf{0.20 \pm 0.05}$	∞
	GBOOST	0.72 ± 0.05	$\mathbf{0.20 \pm 0.05}$	0.64 ± 0.12
	LR	0.63 ± 0.07	$\mathbf{0.20 \pm 0.05}$	0.70 ± 0.25
	RF	$\mathbf{0.74 \pm 0.06}$	$\mathbf{0.20 \pm 0.05}$	0.58 ± 0.15
Naive EM	ETREES	0.50 ± 0.08	0.27 ± 0.04	∞
	GBOOST	0.61 ± 0.08	0.21 ± 0.05	0.51 ± 0.12
	LR	0.56 ± 0.06	$\mathbf{0.20 \pm 0.05}$	0.51 ± 0.11
	RF	0.55 ± 0.07	$\mathbf{0.20 \pm 0.05}$	0.46 ± 0.11
EM w LP	ETREES	0.50 ± 0.09	0.28 ± 0.04	∞
	GBOOST	0.60 ± 0.08	0.21 ± 0.05	0.44 ± 0.05
	LR	0.56 ± 0.06	$\mathbf{0.20 \pm 0.05}$	0.47 ± 0.05
	RF	0.58 ± 0.08	$\mathbf{0.20 \pm 0.05}$	$\mathbf{0.42 \pm 0.06}$
Full PGM	ETREES	0.64 ± 0.05	0.54 ± 0.05	1.27 ± 1.57
	GBOOST	0.73 ± 0.07	0.73 ± 0.07	0.43 ± 0.06
	LR	0.62 ± 0.08	0.64 ± 0.07	0.52 ± 0.10
	RF	0.71 ± 0.05	0.80 ± 0.05	$\mathbf{0.42 \pm 0.07}$

Separation by embryo quality for each baseline method.



(a) Baseline 0 (b) Baseline Cycles

(c) Naive EM (d) EM w LP

Figure: Probability densities separated by embryo quality. Classifier: GBOOST

Conclusions

Conclusions

- The model predicts with higher probability embryos with higher ASEBIR score. Once the selection process has been made, the model does not provide more information about individual embryos.
- The cycle features are an important factor to predict implantation. The EM strategy helps to predict partially implanted cycles and to separate by quality grade.
- The model gathers all these good properties while also providing extra value:
  - Isolation of embryo viability (good for selection).
  - Estimation of unknown factors.

Further research

- Better tuning of probabilistic classifiers
- Considering a simplified model
- Baseline with EM and cycle features.

Thank you!