

1 **Supporting first FSH dosage for ovarian stimulation with Machine Learning.**

2

3 **Authors:** N. Correa^{a,b,c}, J. Cerquides^b, J.L. Arcos^b, R. Vassena^a

4 **Affiliations:**

5 ^a Clínica Eugin-Eugin Group, Barcelona, Spain

6 ^bInstituto de Investigación en Inteligencia Artificial – Consejo Superior de
7 Investigaciones Científicas (IIIA-CSIC), Barcelona, Spain

8 ^c Universitat Autònoma de Barcelona (UAB), Spain

9 **Corresponding autor:** R. Vassena

10 rvassena@eugin.es

11 Carrer de Balmes 236

12 08006 Barcelona

13 +34 93 322 11 22 Ext. 97223

14

15 **Abstract:**

16 **Research question:** Is it possible to identify accurately the optimal first dose of FSH in
17 controlled ovarian stimulation (COS) by means of a machine learning (ML) model?

18 **Design:** Observational study (2011 to 2021) including first In Vitro Fertilization (IVF)
19 cycles with own oocytes. 2713 patients from five private reproductive centers were
20 included in the development phase of the model (2011 to 2019), and 774 in the validation
21 phase (2020 to 2021). Predictor variables included: age, Body Mass Index (BMI),
22 Antimullerian Hormone (AMH), Antral Follicle Count (AFC), and previous live births.
23 Performance of the developed model was measured with a proposed score based on the
24 number of MII retrieved and the dose received and/or recommended.

25 **Results:** The cycles included were from women 37.2 ± 4.9 years old [18-45], with a BMI
26 of 23.7 ± 4.2 , AMH of 2.4 ± 2.3 , AFC of 11.8 ± 7.7 ; and an average number of MII obtained
27 7.2 ± 5.3 . The model reached a mean performance score of 0.87 (95% CI 0.86 to 0.88) in
28 the development phase; this value was significantly better than the one for the doses
29 prescribed by the clinicians for the same patients (0.83 [0.82, 0.84]; p-value= 2.44×10^{-10}).
30 The mean performance score of the model recommendations was 0.89 (95% CI 0.88 to
31 0.90) in the validation phase, also significantly better than clinicians (0.84 [0.82, 0.86];
32 p-value = 3.81×10^{-5}). With these results the model was shown to surpass the performance
33 of the standard practice.

34 **Conclusion(s):** The ML model developed could be deployed as a training and learning
35 tool for new clinicians and serve as quality control for experienced ones; further, it could
36 be used as second opinion, for instance by providing information in peer-to-peer case
37 discussions.

38 **Key Message:** A Machine Learning model was trained to recommend first FSH doses
39 for ovarian stimulation. When compared to clinicians the model developed had
40 consistently better performance scores. The model could be used as a second opinion and
41 as learning tool for new clinicians; to avoid as many non-optimal outcomes as possible.

42 **Key words:** Ovarian Stimulation, Artificial Intelligence, Machine Learning, Prediction.

43

44

45 **Introduction:**

46 Although significant strides have been made in the last 40 years, the mean pregnancy rate
47 after an In Vitro Fertilization (IVF) cycle still hovers around 30%, with a 20% chance
48 of delivery (De Geyter et al., 2018). An important requisite to the success of an IVF cycle
49 is the availability of a certain number of mature oocytes (MII); usually obtained after
50 Controlled Ovarian Stimulation (COS).

51 COS represents therefore a key step for IVF success, as failing to ensure an optimal
52 number of MII oocytes will likely hinder the rest of the procedure. As the number of MII
53 oocytes retrieved increases, so does the chance of producing some embryos with high
54 pregnancy potential (Drakopoulos et al., 2016; Esteves et al., 2019), but stimulating a
55 patient too much leads to an increased risk of Ovarian Hyperstimulation Syndrome
56 (OHSS). As such, a compromise must be reached to strive to retrieve a number inside of
57 a range considered as optimal by literature, that does not increase chances of OHSS but
58 maintains good pregnancy potential. The definition of an optimal range of oocytes we
59 consider during this study goes from 10 to 15 oocytes and considers 4 to 9 oocytes as
60 suboptimal (Polyzos & Sunkara, 2015). Anything outside these values is considered too
61 much or too few. Whenever a patient falls outside the defined range, there is either the
62 risk that the cycle is not successful or even cancelled, or the risk of OHSS, which also
63 implies the need to freeze all the embryos in case they were generated, increasing the
64 costs and delays in treatment.

65 Essential, then, to all COS protocols is the starting dose of exogenous Follicle Stimulating
66 Hormone (FSH). This dose should be enough to recruit all FSH responsive follicles, but
67 should not be any higher, to avoid unsafe effects such as OHSS or decreased oocyte
68 quality. After approximately 8 days of stimulation, changing the FSH dose does not allow
69 for a significant further recruitment of follicles (Fleming et al., 2006). In other words, if

70 the starting dose of exogenous FSH is not adequate, little can be done to fix its effects on
71 MII yield.

72 The choice of the FSH starting dose is mostly based on the patient's characteristics such
73 as age, Body Mass Index (BMI), or ovarian reserve and clinical characteristics such as
74 past gravidity and parity. Sometimes, COS leads to unexpected and widely different
75 results even among apparently similar patients, resulting in either too many or too few
76 oocytes collected. Further, where the MII retrieved are in the expected number range,
77 they may still be of insufficient quality to achieve success, as only 30-40% of
78 microinjected oocytes develop to blastocyst (Maggiulli et al., 2020; Vaiarelli et al., 2020),
79 and around 11% to an euploid blastocyst (Chamayou et al., 2017).

80 Clinicians use their knowledge and experience to prescribe a starting FSH dose to reach
81 the appropriate range of follicular stimulation. So far, some Machine Learning (ML)
82 models have been developed to encapsulate that medical experience reflected in historical
83 data to try to automate that decision. Two separate nomograms based on patient age,
84 Antimullerian Hormone (AMH) or Antral Follicle Count (AFC), and basal FSH levels
85 have been developed for this task (Ebid et al., 2021; La Marca et al., 2012). One of them
86 was tested prospectively (Allegra et al., 2017), reporting an increase in the number of
87 patients with an optimal range of MII retrieved, and a decrease in patients with lower
88 response in those using the nomogram. These two nomograms did not include patients
89 older than 40 or those with irregular cycles, including Polycystic Ovary Syndrome
90 (PCOS) patients. In a Randomized Controlled Trial (RCT) study for another model
91 developed specifically for individualized dosage of FSH delta (Nyboe Andersen et al.,
92 2017) no differences in pregnancy rate were observed.

93

94 Additionally, the CONSORT model, based on multivariate regression (Howles, Saunders,
95 Alam, & Engrand, 2006) predicted overall lower starting doses compared to those
96 prescribed by clinicians in normo-ovulatory patients (Naether et al., 2015; Pouly et al.,
97 2015). CONSORT was also tested by RCT (Olivennes et al., 2015), showing that the
98 model was able to reduce OHSS risk in patients while maintaining comparable pregnancy
99 rates compared to the clinician chosen dose, in spite of a reduction in the number of
100 retrieved oocytes.

101 The goal of this study is to develop and validate an ML based model designed to identify
102 the optimal starting dose for all variants of FSH except delta (as it is not quantified in
103 IU/ml), allowing to collect a number of MII as close as possible to 12 (a middle point in
104 the optimal range considered in this study), for all types of patients.

105

106 **Materials and Methods:**

107 **Patient population and ethical approval**

108 Data from a total of 2713 first IVF cycles registered in five private centers located in
109 Spain and Italy, from 2011 to 2019, were used to develop the model. Natural cycles and
110 cycles where gonadotrophins doses were not expressed in IU/ml were excluded. The
111 inclusion of first cycles aimed to prevent bias due to unrecorded clinician knowledge
112 (such as previous cycles' FSH dosage and results). An additional 774 cycles from 2020
113 to 2021 were used for prospective validation of the model. Three categories of data were
114 collected as variables. First, the input data, composed by age, BMI, proven fertility (Y/N)
115 and reserve markers AMH and AFC. Second, the intervention, namely the first dose of
116 FSH prescribed by the clinician. And third, result data expressed as number of metaphase
117 II oocytes (MII) collected after stimulation (see Table I for details). During the whole
118 study only cases with complete data on all the variables were included.

	Development database (n=2713)	Validation database (n=774)	p-value
Age (years)	37.9 ± 4.6	38.3 ± 4.4	<0.05*
AMH (ng/ml)	2.4 ± 2.3	2.2 ± 2.2	<0.05*
AFC (n)	11.1 ± 7.3	11.3 ± 8.5	0.7
BMI (kg/m²)	23.6 ± 4.2	23.2 ± 3.9	<0.05*
Number of MII retrieved	6.8 ± 5.0	6.7 ± 6.5	<0.05*
Proven female fertility	13%	10.1%	0.067

119 Table I: Patient characteristics in the two databases used in the study. Values are expressed as average and SD.
120 Variables were compared using Mann-Whitney U test. For proportions a 2-sample z-test was conducted. A p-value of
121 <0.05 was decided as significant.
122

123 Permission to conduct this study was requested and obtained from the Ethical Committee
124 for Research of Eugin (approval code: 20/10/2020 ALGO2).
125

126 **Predictive model construction**

127 In agreement with the current literature (Polyzos & Sunkara, 2015) we aimed at predicting
128 the initial dose of FSH to achieve a number of MII as close as possible to 12, with the
129 range 10-15 considered desirable, the range 4-9 suboptimal, and MII <4 or >15 not
130 desirable. Given patient characteristics and limitations in the maximum dose of FSH
131 administered, not every patient is considered able to reach the desired goal.

132 A predictive model was constructed to predict the patient's capability to react to the first
133 dose of FSH. This capability can be described by the slope of a simplified linear dose-
134 response function. For any patient, during a natural cycle (0 UI/ml of exogenous FSH)
135 the outcome in number of MII collected would stay mainly between 0 to 1. Given that in
136 the database there are the results of a specific dose of FSH, the value of individual slopes

137 is easily computable. To avoid negative slope values, it was assumed that all patients
138 would get 0 MII if given 0 exogenous FSH.

139 The slope of a linear function is defined as follows:

140
$$m = \frac{y_2 - y_1}{x_2 - x_1}$$

141 As the first data point (x_1, y_1) is set at the origin (0, 0), the slope value for every patient
142 is computed by dividing the outcome or y_2 (MII) by x_2 (the first dose of FSH).

143 A Linear Regression algorithm was trained to predict the slope for every case (defined by
144 its values at the start of the stimulation in age, BMI, AFC, AMH and proven fertility).

145 Training was conducted on a random 80% of the development database. The remaining
146 20% was reserved for testing purposes. The training process was cross validated 5 times
147 with 5 randomly selected training datasets, with their corresponding 5 test sets.

148 The final model used was composed by an intercept of 0.067, and coefficients for age,
149 BMI, AFC, AMH an proven fertility of -0.00135, -0.00371, 0.00107, 0.00445 and -
150 0.00038 correspondingly.

151

152 **Dose recommendation by the model**

153 For dose recommending purposes, the predicted slope for each test patient was used to
154 compute the necessary FSH to get an outcome of 12 MII using the following linear
155 function:

156
$$y = m * x$$

157 Where y is the number of MII, m the value of the slope, and x the FSH quantity.

158 As prescribing more than 300 UI/ml of FSH is described to give little to no advantages
159 (Bastu et al., 2013; Harrison et al., 2001), recommended doses were capped at 300 UI/ml.

160

161 **Development of a performance score**

162 We designed a score function to compare the model's recommendations to the clinician
163 prescriptions; given any FSH prescription with its resulting MII outcome, the score
164 function assigns a score for a hypothetical recommended dose from -1 (the recommended
165 dose was too low) to 1 (too high), being 0 the best possible value (the dose recommended
166 was appropriate). FSH doses were categorized in 4 ordinal ranks (100 to 150, 151 to 200,
167 201 to 250, and 251 to 300) to create the score function.

168 The score function also allows assessing clinical prescriptions by setting the
169 recommended dose equal to the clinician prescribed one. In doing so, the function
170 evaluates how close the MII outcome is from the optimal range (10 to 15) and if there is
171 any room for improvement dose-wise (see details in Appendix).

172

173 **Evaluation of performance of the model**

174 Performance of model-based recommendations was evaluated using the proposed score
175 in the reserved 20% for test of the development database and in the prospective validation
176 database. In both cases, two scores were computed for each patient. One score for the
177 dose prescribed by the clinician, and another score for the model recommended one.
178 Absolute values of both scores were compared across all cases, in order to identify which
179 group (clinical or model recommended) had more scores closer to 0, being here of no
180 importance if the dose was too high or too low. The Wilcoxon signed-rank test was used
181 for this purpose, as distribution of the scores was not normal. For an easier interpretation
182 of the results, mean score values were expressed as $1 - |score|$, where a value close to
183 1 is best. All computations were done using Scikit-learn 0.24 in Python 3.7.6.

184

185 **Results:**

186 **Predictive and recommendation performance**

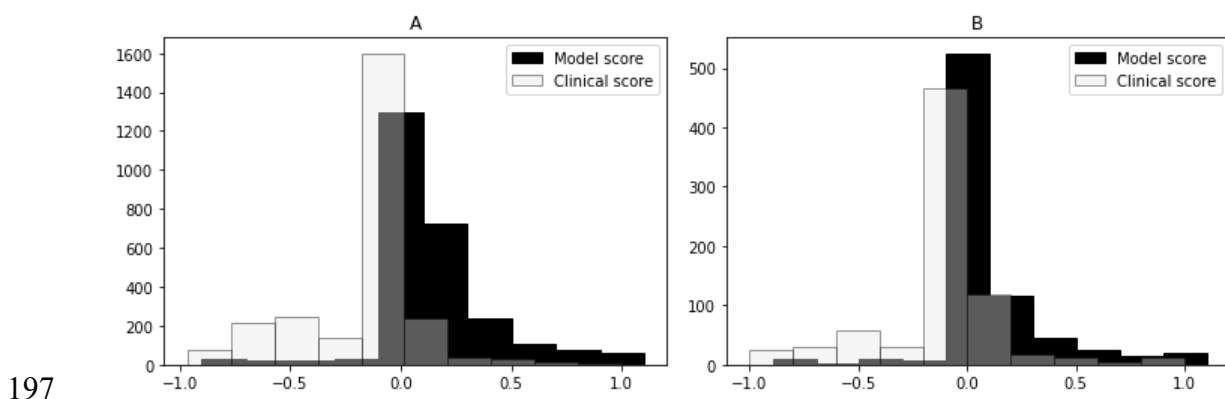
187 During the development phase of our model, the mean performance score for clinical
188 doses was 0.83 (95% CI 0.82 to 0.84), and for model recommendations was 0.87 [0.86,
189 0.88] (p-value= 2.44 e-10).

190 During validation, the mean score for prescriptions was calculated to be 0.84 [0.82, 0.86],
191 and for the model's recommendations 0.89 [0.88, 0.90] (p-value = 3.81 e-05).

192

193 Score and dosage analysis

194 To further understand the performance of the model and of the clinical prescriptions, we
195 compared graphically the model, and the clinicians score distributions in the test set of
196 the development database and in the validation one (Figure 1).

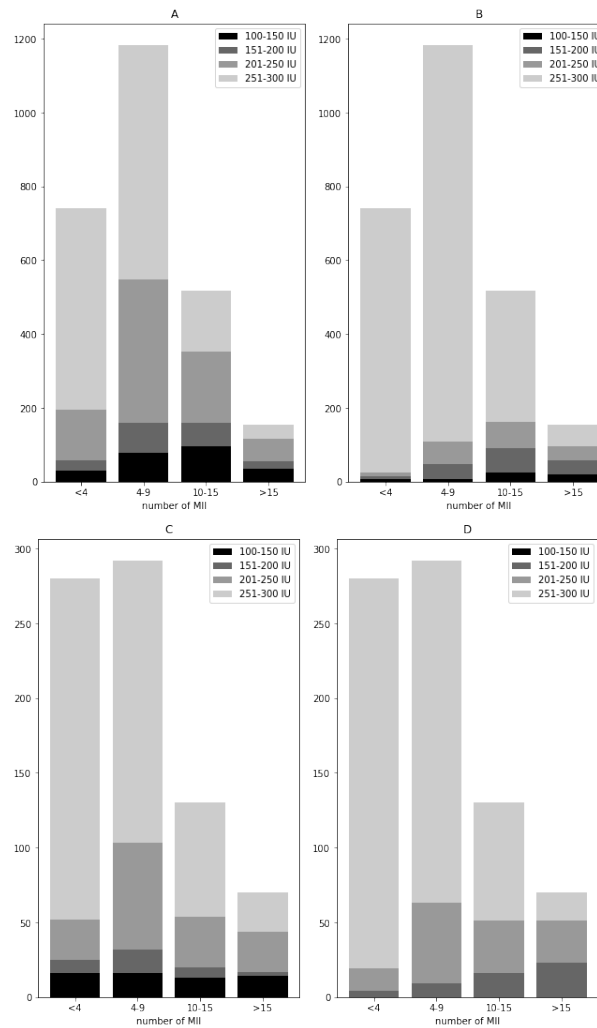


198 Figure 1: Performance during development and validation. Both clinical and model scores are plotted during
199 development in panel A, and in panel B during validation. In black the model scores, and in light grey the clinical
200 scores.

201

202 The model's score approaches 0 (the best possible dose) more times than the clinician
203 does and tends to suggest a dose higher than the one favored by the clinicians when not
204 approaching 0. In 57.4% of cases in the test set and in 68.8% in the validation database
205 the dose rank was not modified in relation to the clinician prescribed dose.

206 Further, we analyzed how the dosage was changed from clinician prescription to model
207 recommendation, in relation to the real outcome in number of MII (Figure 2).



208

209 Figure 2: Dose ranks prescribed per bracket of MII retrieved. Panel A: by the clinicians during development. Panel B:
210 by the model during development. Panel C: by the clinicians during validation. Panel D: by the model during validation.
211 Dose ranks are represented lower to higher from black to light grey.

212

213 The model tends to increase the dose for patients with low and sub-optimal oocyte
214 retrievals, but also increases dosage for some of the hyper-responders.

215

216 **Discussion:**

217 Current state of the art of FSH dosage models includes several recommender models that
218 have provided very optimistic results. Yet, some of them have not been tested by an RCT
219 study (Ebid et al., 2021) or the ones that do (Allegra et al., 2017; Nyboe Andersen et al.,
220 2017; Olivennes et al., 2015) have not been developed to be used on all types of patients.

221 Including only normo-ovulatory patients (Howles et al., 2006), or younger than 40
222 patients with regular cycles (La Marca et al., 2012; Nyboe Andersen et al., 2017), is
223 restricting this new personalization of the first FSH dose to a small subset of patients.
224 And a subset, also, where this personalized dose finding is not as critical as with the
225 patients excluded. The model presented in this study, improves state of the art in this
226 aspect by including every type of patient to enhance results for all of them.

227 Including age, AFC, AMH, BMI and presence of previous successful pregnancies as
228 variables in the core model has shown to allow for a good prediction of the dose-response
229 function slope. This value has already been used as ovarian sensitivity (oocytes recovered
230 per unit of starting FSH) in the development of an RCT tested nomogram (La Marca et
231 al., 2012). Its use as objective variable of the core model mitigates the confounding effect
232 that appears in any non-randomized treatment database, and that could lead a direct model
233 (oocyte number as objective variable) to learn for example that higher doses, which are
234 often prescribed for low-responders, lead to smaller oocyte yields. As in this setting this
235 is always true (the treatment is tailored for the patient by the clinician), it is especially
236 important to account for it. Removal of this effect, also allows for the extension of the
237 recommender model for all types of patients, as they do not confuse the model, on the
238 contrary, the core model learns that extreme patients (low-responders and hyper-
239 responders) have extreme ovarian potential values.

240 As a separate contribution to an inclusive recommender model, we have developed a way
241 to test in-silico whether the model would improve results versus the historical data, as a
242 step preceding an RCT. To this end, the performance score was designed to encode and
243 automate faithfully an expert clinical assessment of treatment-recommendation-outcome
244 combinations. In other words, it enables us to assert if a recommended dose could fare
245 better than the one already prescribed, given the real result in MII retrieved. In this way,

246 it is possible to estimate reliably whether the model has the possibility to improve current
247 clinical practice. With this information, the investment in a well-designed RCT can be
248 made more confidently. Additionally, results of the in-silico performance of the model
249 are more informative than the sole prediction scores of the core recommender model.

250 We show that, the scores of the model were consistently better than those of the clinical
251 practice, both in the development database and the validation one. This is of interest as
252 the model holds its value even though the population of the validation database is
253 significantly older. Therefore, it means that the core model learnt the important aspects
254 of the relationship between the patient's characteristics and her ovarian potential or slope
255 in the dose-response function. It is worth of note that the most significant predicted
256 improvement was predicted for the patients whose oocyte yield was low or sub-optimal,
257 where doses get increased on average. Upon implementation, the system's
258 recommendations may improve the average results and most probably avoid some cycle
259 cancellations due to lack of embryos for transference.

260 Detailed analysis of the behavior of the model revealed its tendency, when wrong, to
261 overdose some patients. This contrasts with clinical practice, where the tendency is to
262 underdose when the prescription is not adequate. These instances of overestimations by
263 the models correspond mainly to hyper-responder patient profiles, which are
264 underrepresented on our databases. As such, the algorithm couldn't learn appropriately
265 due to the lack of a sufficient sample size. Importantly, while the model does tend to
266 overdose these patients, it is of note that it still recommends the same or lower doses than
267 the clinician in most of these cases, i.e. the clinician also tend to overdose. Nonetheless,
268 we cannot dismiss the possibility that this could lead to a small increase in risk of OHSS.

269 This contrasts with previous results in the literature, where RCT tested models reduced
270 OHSS risk of incidence (Allegra et al., 2017; Nyboe Andersen et al., 2017; Olivennes et

271 al., 2015). Secondary results of these studies, though, failed to show an increase either in
272 retrieved oocytes or pregnancy results, with one reporting a reduction in oocyte yield
273 (Olivennes et al., 2015). While OHSS risk must be taken very seriously, it is also true that
274 nowadays it is manageable cycle-wise with proper prevention such as GnRH agonist
275 trigger. All things considered, perhaps a manageable risk for a small portion of patients
276 could be a fair trade-off to avoid a lack of embryos suitable for transference for other
277 patients.

278 Further analysis of the instances where the model made a suboptimal suggestion led to
279 another conclusion. The cases where the model gets negative error scores seem to
280 coincide several times with also negative error scores for the clinician's prescription.
281 Looking into the specific cases where this happens gave us a profile of patients with very
282 good markers and an unexplained low retrieval of oocytes. This could possibly be related
283 to not diagnosed genetic polymorphisms in the FSHR or LHB genes (Lledo et al., 2014)
284 that, obviously, nor the clinicians or the model could detect.

285 Despite the already described possible limitations of the system, it is encouraging that the
286 preliminary results show in the vast majority of cases a similar or better performance
287 score of the model's recommendation versus the dose prescribed by the clinician.

288

289 **Conclusions:**

290 Clinicians prescribe the first FSH dose for each patient based in their characteristics,
291 reserve markers and their own experience with similar cases. Although most of the time
292 they prescribe the dose necessary for an optimal result, sometimes the outcome can
293 unexpectedly vary and fall in suboptimal or extreme ranges. Our model could avoid most
294 of these deviations by analyzing the patient's profile and making a suggestion for the
295 medical professional to assess.

296 Once tested and its performance confirmed via RCT, the ML model developed could be
297 deployed as a training and learning tool for new clinicians and serve as quality control
298 for experienced ones; further, it could be used as an second opinion, for instance by
299 providing information in peer-to-peer case discussions.

300

301 **Authors' roles:**

302 N.C. involved in study design, data collection and curation, analysis, model construction,
303 interpretation and manuscript preparation. J.C. involved in analysis, model construction,
304 interpretation and manuscript preparation. J.A. involved in analysis, model construction,
305 interpretation and manuscript preparation. R.V. involved study supervision, expert
306 knowledge, and manuscript preparation.

307

308 **Funding:**

309 This work was supported by Doctorat Industrial funded by Generalitat de Catalunya [DI-
310 2019-24], by project CI-SUSTAIN funded by the Spanish Ministry of Science and
311 Innovation [PID2019-104156GB-I00], and by intramural funding by Clínica Eugin-
312 Eugin Group.

313

314 **Conflicts of interest:**

315 None.

316

317 **Acknowledgements:**

318 We would like to thank dr. Maria Jesús López for kindly lending us her time and her
319 expertise on Controlled Ovarian Stimulation protocols.

320

321 **Reference list:**

- 322 Allegra, A., Marino, A., Volpes, A., Coffaro, F., Scaglione, P., Gullo, S., & La Marca,
323 A. (2017). A randomized controlled trial investigating the use of a predictive
324 nomogram for the selection of the FSH starting dose in IVF/ICSI cycles.
325 *Reproductive BioMedicine Online*, 34(4), 429–438.
326 <https://doi.org/10.1016/j.rbmo.2017.01.012>
- 327 Bastu, E., Celik, C., Keskin, G., & Buyru, F. (2013). Evaluation of embryo transfer time
328 (day 2 vs day 3) after imposed single embryo transfer legislation: When to
329 transfer? *Journal of Obstetrics and Gynaecology*, 33(4), 387–390.
330 <https://doi.org/10.3109/01443615.2012.761186>
- 331 Chamayou, S., Sicali, M., Alecci, C., Ragolia, C., Liprino, A., Nibali, D., ...
332 Guglielmino, A. (2017). The accumulation of vitrified oocytes is a strategy to
333 increase the number of euploid available blastocysts for transfer after
334 preimplantation genetic testing. *Journal of Assisted Reproduction and Genetics*,
335 34(4), 479–486. <https://doi.org/10.1007/s10815-016-0868-0>
- 336 De Geyter, C., Calhaz-Jorge, C., Kupka, M. S., Wyns, C., Mocanu, E., Motrenko, T., ...
337 Baranowski, R. (2018). ART in Europe, 2014: Results generated from European
338 registries by ESHRE. *Human Reproduction*, 33(9), 1586–1601.
339 <https://doi.org/10.1093/humrep/dey242>
- 340 Drakopoulos, P., Blockeel, C., Stoop, D., Camus, M., De Vos, M., Tournaye, H., &
341 Polyzos, N. P. (2016). Conventional ovarian stimulation and single embryo
342 transfer for IVF/ICSI. How many oocytes do we need to maximize cumulative live
343 birth rates after utilization of all fresh and frozen embryos? *Human Reproduction*,
344 31(2), 370–376. <https://doi.org/10.1093/humrep/dev316>
- 345 Ebid, A. H. I. M., Motaleb, S. M. A., Mostafa, M. I., & Soliman, M. M. A. (2021).

- 346 Novel nomogram-based integrated gonadotropin therapy individualization in in
347 vitro fertilization/ intracytoplasmic sperm injection: A modeling approach. *Clinical*
348 *and Experimental Reproductive Medicine*, 48(2), 163–173.
349 <https://doi.org/10.5653/cerm.2020.03909>
- 350 Esteves, S. C., Carvalho, J. F., Bento, F. C., & Santos, J. (2019). A novel predictive
351 model to estimate the number of mature oocytes required for obtaining at least one
352 euploid blastocyst for transfer in couples undergoing in vitro
353 fertilization/intracytoplasmic sperm injection: The ART calculator. *Frontiers in*
354 *Endocrinology*, 10(FEB), 1–14. <https://doi.org/10.3389/fendo.2019.00099>
- 355 Fleming, R., Deshpande, N., Traynor, I., & Yates, R. W. S. (2006). Dynamics of FSH-
356 induced follicular growth in subfertile women: Relationship with age, insulin
357 resistance, oocyte yield and anti-Mullerian hormone. *Human Reproduction*, 21(6),
358 1436–1441. <https://doi.org/10.1093/humrep/dei499>
- 359 Harrison, R. F., Jacob, S., Spillane, H., Mallon, E., & Hennelly, B. (2001). A
360 prospective randomized clinical trial of differing starter doses of recombinant
361 follicle-stimulating hormone (follitropin- β) for first time in vitro fertilization and
362 intracytoplasmic sperm injection treatment cycles. *Fertility and Sterility*, 75(1),
363 23–31. [https://doi.org/10.1016/S0015-0282\(00\)01643-5](https://doi.org/10.1016/S0015-0282(00)01643-5)
- 364 Howles, C. M., Saunders, H., Alam, V., & Engrand, P. (2006). Predictive factors and a
365 corresponding treatment algorithm for controlled ovarian stimulation in patients
366 treated with recombinant human follicle stimulating hormone (follitropin alfa)
367 during assisted reproduction technology (ART) procedures. An analysis . *Current*
368 *Medical Research and Opinion*, 22(5), 907–918.
369 <https://doi.org/10.1185/030079906X104678>
- 370 La Marca, A., Papaleo, E., Grisendi, V., Argento, C., Giulini, S., & Volpe, A. (2012).

- 371 Development of a nomogram based on markers of ovarian reserve for the
372 individualisation of the follicle-stimulating hormone starting dose in in vitro
373 fertilisation cycles. *BJOG: An International Journal of Obstetrics and*
374 *Gynaecology*, 119(10), 1171–1179. [https://doi.org/10.1111/j.1471-](https://doi.org/10.1111/j.1471-0528.2012.03412.x)
375 [0528.2012.03412.x](https://doi.org/10.1111/j.1471-0528.2012.03412.x)
- 376 Lledo, B., Ortiz, J. A., Llacer, J., & Bernabeu, R. (2014). Pharmacogenetics of ovarian
377 response. *Pharmacogenomics*, 15(6), 885–893. <https://doi.org/10.2217/pgs.14.49>
- 378 Maggiulli, R., Cimadomo, D., Fabozzi, G., Papini, L., Dovere, L., Ubaldi, F. M., &
379 Rienzi, L. (2020). The effect of ICSI-related procedural timings and operators on
380 the outcome. *Human Reproduction*, 35(1), 32–43.
381 <https://doi.org/10.1093/humrep/dez234>
- 382 Naether, O. G. J., Tandler-Schneider, A., & Bilger, W. (2015). Individualized
383 recombinant human follicle-stimulating hormone dosing using the CONSORT
384 calculator in assisted reproductive technology: A large, multicenter, observational
385 study of routine clinical practice. *Drug, Healthcare and Patient Safety*, 7, 69–76.
386 <https://doi.org/10.2147/DHPS.S77320>
- 387 Nyboe Andersen, A., Nelson, S. M., Fauser, B. C. J. M., García-Velasco, J. A., Klein,
388 B. M., Arce, J. C., ... Arce, J. C. (2017). Individualized versus conventional
389 ovarian stimulation for in vitro fertilization: a multicenter, randomized, controlled,
390 assessor-blinded, phase 3 noninferiority trial. *Fertility and Sterility*, 107(2), 387-
391 396.e4. <https://doi.org/10.1016/j.fertnstert.2016.10.033>
- 392 Olivennes, F., Trew, G., Borini, A., Broekmans, F., Arriagada, P., Warne, D. W., &
393 Howles, C. M. (2015). Randomized, controlled, open-label, non-inferiority study
394 of the CONSORT algorithm for individualized dosing of follitropin alfa.
395 *Reproductive BioMedicine Online*, 30(3), 248–257.

- 396 <https://doi.org/10.1016/j.rbmo.2014.11.013>
- 397 Polyzos, N. P., & Sunkara, S. K. (2015). Sub-optimal responders following controlled
398 ovarian stimulation: An overlooked group? *Human Reproduction*, 30(9), 2005–
399 2008. <https://doi.org/10.1093/humrep/dev149>
- 400 Pouly, J. L., Olivennes, F., Massin, N., Celle, M., Caizergues, N., & Contard, F. (2015).
401 Usability and utility of the CONSORT calculator for FSH starting doses: A
402 prospective observational study. *Reproductive BioMedicine Online*.
403 <https://doi.org/10.1016/j.rbmo.2015.06.001>
- 404 Vaiarelli, A., Cimadomo, D., Conforti, A., Schimberni, M., Giuliani, M., D’Alessandro,
405 P., ... Ubaldi, F. M. (2020). Luteal phase after conventional stimulation in the
406 same ovarian cycle might improve the management of poor responder patients
407 fulfilling the Bologna criteria: a case series. *Fertility and Sterility*, 113(1), 121–
408 130. <https://doi.org/10.1016/j.fertnstert.2019.09.012>
409