ISSN: 2641-6247
DOI: 10.33552/WJGWH.2022.05.000624
World Journal of Gynecology & Women's Health
Iris Publishers

Research Article

Evaluating the Discriminatory Performance of Embryonic Day 14 Serum β-Human Chorionic Gonadotropin Levels in Women Undergoing *In Vitro* Fertilization Using the Volume Under a Three-Class Receiver Operating Characteristic Surface

Enrique Rodriguez-Villasana[1], Lucia Muradas[1], Leonardo M Porchia[2], M Elba González-Mejia[3] and Esther López-Bayghen[4]*

[1]*Instituto de Fertilidad y Genética Monterrey, México*

[2]*Instituto de Infertilidad y Genética, Ingenes México*

[3]*Departamen to de Genética, Facultad de Medicina, Benemérita Universidad Autónoma de Puebla, Puebla, México*

[4]*Departamento de Toxicología, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional (CINVESTAV-IPN), México*

***Corresponding author:** : Esther López-Bayghen Ph.D., Departamento de Toxicología, Cinvestav-IPN

Av. Instituto Politécnico Nacional 2508, Col. San Pedro Zacatenco, CP 07360, México City, Distrito Federal, México.

Received Date: July 28, 2022

Published Date: August 18, 2022

Abstract

Purpose: Human chorionic gonadotropin (hCG) correlates with the number of chorionic sacs and the pregnancy prognosis, wherein researchers have speculated it could distinguish between a viable pregnancy (singleton, twin, triplet) and a non-viable pregnancy (aborted, biochemical, ectopic). This study investigated the predictability of serum β-hCG on Embryonic Day 14 for pregnancy outcomes.

Methods: Patients charts were retrospectively assessed for patients that underwent 2-3 embryo transfers. Area and volume under a receiver-operating characteristic curve was used to determine the predictability between 2-classes (AUC) and 3-classes (VUS).

Results: β-hCG ranged between 10 and 799 mIU/mL for 663 IVF cycles. Non-viable results clustered towards lower values, whereas viable pregnancies clustered towards higher levels, which significant differences between the singleton, twin, triplet, and non-viable outcomes were observed ($p<0.05$). There was no effect on the results when stratified by ova source, transfer protocol, or number of embryos transferred. β-hCG was highly predictable for viable pregnancies (AUC=0.939, 95%CI: 0.916-0.961, $p<0.001$) with a cutoff of 126 mIU/mL. Using data from 2-embryo transfers, β-hCG's predictability for 3 classes (non-viable, singleton, twin pregnancies) was highly selective (VUS=0.644, 95%CI: 0.559-0.690, $p<0.001$). Cutoffs were 74 and 245 mIU/mL using VUS, 121 and 265 mIU/mL using the Youden Index, respectively. Using all viable pregnancies, β-hCG was highly predictive for viable outcomes (VUS=0.479, 95%CI: 0.391-0.576, $p<0.001$). Thresholds were 171 and 427 mIU/mL using VUS and 266 and 399 mIU/mL using the Youden Index, for singleton-twin and twin-triplet pregnancies, respectively.

Conclusion: Embryonic Day 14 serum β-hCG sufficiently distinguish between non-viable, singleton, twin, and triplet pregnancies.

Keywords: β-hCG; Assisted reproductive technology; Biochemical predictors; Biochemical pregnancy; Ectopic pregnancy; Live birth; Multiple pregnancies; Reproductive endocrinology

Introduction

Numerous advancements have been made to improve the outcomes of assisted reproductive technologies (ART), such as *in-vitro* fertilization (IVF), the use of donor embryos, assisted hatching, vitrification of the embryo, increasing the number of embryos transferred, endometrial preparation, and pre-implantation genetic testing [1-4]. During this stressful process, many patients and clinicians opt to assess the viability of the pregnancy and the need for potential interventions, such as prenatal care or embryo reduction for high-risk multifetal pregnancies. One of the initial time points for this assessment is within 10 days of embryo implantation. Typically, human chorionic gonadotropin (hCG) is measured.

hCG, secreted by syncytiotrophoblast cells, can initially be detected in maternal blood 6 to 8 days after fertilization [5]; however, most ART clinics evaluate the viability of the pregnancy from either Embryonic Day 11 to 16. Interestingly, the prognosis of the pregnancy can be determined using serum β-hCG, which is correlated with the number and size of the gestational sacs [6,7]. For example, Lawler et al. demonstrated that serum β-hCG values >80 mIU/ml can predict an ongoing pregnancy [8]. However, when considering certain factors—the number of amniotic sacs per chorionic sac, the patients' age, weight, ethnicity, and lifestyle (diet, exercise, and drug use) [7,9,10] as well as polycystic ovary syndrome (PCOS) [11], diabetes [12], and certain cancers [13,14]—the value for β-hCG can range for a specific pregnancy. A study by Qui et al. recently demonstrated an age-dependent inverse relationship with β-hCG values measured on Day 10 after blastocyst transfer for clinical pregnancy [15]. In comparison, Yuan et al. demonstrated on Day 7 the potential of distinguishing between a singleton and twin pregnancy using serum β-hCG [16]. However, the inconsistencies between the day for measuring β-hCG and the differences in critical components of the cohorts (the number of embryos transferred, frozen/thaw versus fresh embryos) allow different interpretations of the association between β-hCG values and IVF outcomes. Moreover, due to the overlap between β-hCG values for each IVF outcome, comparing β-hCG values with IVF outcomes and determining proposed cutoffs leads to potential flaws.

The purpose of the study was to evaluate the effect of 3 key factors [ova source (donor versus patient), type of embryo transfer (fresh versus frozen), and the number of embryos transferred (2 versus 3)] on β-hCG values in determining viable and non-viable pregnancies. Moreover, cutoffs between non-viable, singleton, twin, and triplet pregnancies were determined using volume under the receiver operating characteristic (ROC) curve (VUS).

Materials and Methods

Study design and participants

Potential patients' charts were reviewed from January 2015 to December 2016. The cohort pool consisted of patients who attended the Ingenes Institute in Mexico. They had to fulfill the following inclusion criteria: 1) age >18 years, 2) had history of male factor, 3) had a normal endometrial cavity as determined by hysteroscopy or hysterosonography, 4) good quality embryo available for embryo transfer, 5) only 2 or 3 embryos were transferred, and 6) singletons, twins, and triplet pregnancies were only mono-, di- and trichorionic, respectively. In addition, the patients were excluded if it was documented or suspected that 1) the patient suffered from an autoimmune disease, thrombophilia, hematologic disorders, uncontrolled endocrine disorders or other medical conditions, 2) congenital and untreated acquired uterine abnormalities, 3) couple with genetic and chromosomal abnormalities, or 4) only produced poor-quality embryos. Once the potential records were identified, data were appropriately anonymized, and informed consent was obtained at the time of original data collection. Ethics Committee approval was not needed. Moreover, the study was conducted according to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines [17].

IVF and Embryo Transfer

All patients underwent controlled ovarian stimulation for 10 to 14 days with gonadotrophin-releasing hormone agonists and antagonists. First, the ovarian response was assessed by measuring serum estradiol levels, and ultrasound examination evaluated follicular development. Then, oocyte retrieval was conducted 36 hours after hCG administration. At the same time, the partner's semen was prepared by density gradient centrifugation. The oocytes were all inseminated by intracytoplasmic sperm injection, and fertilization was judged by forming two pronuclei 19 hours after insemination. Embryos were cultured in Global Total for Fertilization media (Cat # LGGT-30, Life Global) and incubated at 37 °C in 8% CO_2, 5% O_2, and 87% N_2. An embryologist monitored and recorded all information about the antral follicle count, fertilization rates, embryo development, and embryo morphology for each oocyte.

Embryos were either transferred during a natural (fresh transfer) or the following cycle (frozen embryo transfer). Fresh transfers took place on Day 5. The resulting blastocysts were cryopreserved by vitrification for frozen embryo transfer and thawed later. The uterine transfer occurred during a controlled endometrial development cycle for frozen embryos, free of gonadotropin stimulation. Embryo implantation was confirmed on Day 14 by serum β-hCG concentrations (>10 mUI/ml). All the patient's demographics, IVF cycle information, implantation rate, and IVF outcomes (pregnancies and miscarriages) were recorded by the Physician. Clinical decisions about which cycle and the number of embryos to transfer were determined by the Physician with the patient's approval.

Statistical analysis

All analyses were carried out with the Statistical Package for the Social Sciences software (SPSS, v. 22.0, Chicago, IL USA) or with the R software [18]. The cohort was stratified based on the patient's cycle outcome as either: 1) biochemical (β-hCG >10 mUI/ml on Day 14 but an absence of a fetal heartbeat at 6 weeks); 2) aborted (β-hCG >10 mUI/ml on Day 14 and the presence of a fetal heartbeat at 6 weeks, but the pregnancy terminated before 30 weeks or the delivery resulted in a non-viable offspring); 3) ectopic (β-hCG >10 mUI/ml on Day 14 but the gestational sac located outside of the womb); 4) singleton (β-hCG >10 mUI/ml on Day 14 and the pregnancy resulted in the delivery of 1 baby); 5) twin (β-hCG >10 mUI/ml on Day 14 and the pregnancy resulted in the delivery of 2 babies); and 6) triplet (β-hCG >10 mUI/ml on Day 14 and the pregnancy resulted in the delivery of 3 babies). The normality of continuous variables was determined using the Kolmogorov-Smirnov test. Levene's test was used to examine the difference in the variances of continuous vari-

ables. For parametric data, differences between groups were determined by the ANOVA with a post hoc Dunnett T3 or Bonferroni test. For non-parametric data, differences between groups were determined with the Kruskal-Wallis test with a post hoc Dunn's test. Associations were determined by calculating Spearman coefficient rho (ρ) [19]. ROC analysis was used to determine specificity and sensitivity between β-hCG and pregnancy outcomes. The area under the ROC curve (AUC) was calculated using the method described by Hanley and McNeil [20]. Youden's index (sensitivity + specificity – 1) was calculated using the sensitivity and specificity parameters, and the highest score was considered the optimal cutoff value. Analogous to AUC for a 2-class system, where a value of 0.5 indicates classification by chance, VUS and extended Youden index were used for a 3-class system. Here, the VUS value can vary from 1/6 (classifying by chance) to 1 (perfect classification) [21]. The DiagTest3grp package and the bcROC surface package for R were used to calculate VUS and the extended Youden Index as well to determine the optimal cutoff points [22,23]. Data are represented as the mean ± standard deviation unless stated otherwise. P-values <0.05 (two-tailed) were considered statistically significant.

Results

During the period in which potential patient charts were reviewed, 8918 charts were available. However, only 30% were seen for specifically male factor. Of these, 695 cycles were considered for review. Ten had β-hCG values which were outliers; 1 had a pregnancy with unexplained chronicity, 13 had single embryo transfer, 6 had more than 3 embryos transferred, whereas 2 had the β-hCG value not recorded. Thus, 663 IVF cycles from 635 patients were included. The characteristics of the cohort are shown in Table 1. Most of the cohort was normal weight or overweight, with most women's age ranging between 30 and 40 years. To note, the patient's body-mass index nor their age had any effect on the results. The most common reason for attending IVF was advanced age, low response, tubal factor, polycystic ovary syndrome, and endometriosis.

Table 1: Characteristics of included patients.

Category	Value
Number of patients (n)	635
Number of cycles (n)	663
Age (years)	38.0 ± 5.9
18-30 (%)	8.8
30-40 (%)	47.3
≥40 (%)	43.9
Body-mass index (kg/m^2)	25.9 ± 4.0
Underweight (%)	0.8
Normal weight (%)	47
Overweight (%)	37.8
Obese (%)	13.8
Morbid obese (%)	0.6
Antral follicle count (n)	12.1 ± 8.6
Ova collected (n)	13.2 ± 7.6
Fertilization rate (%)	77.2 ± 17.9
Causes of infertility (%)	
Advanced age	31.2
Anovulation	2.1
Endometriosis	8.4
Recurrent pregnancy lost	2
No fertilization	0.9
Insemination failure	1.4
Low-quality oocytes	2.1
Low response	14.2
Polycystic ovary syndrome	9.8
Unknown primary infertility	3.9
Unknown secondary infertility	2
Tubal factor	11.8

Values are either frequency, percent, or mean ± standard deviation.

For the cohort, the β-hCG values ranged between 10 and 799 mIU/mL; however, it was observed that non-viable results, such as biochemical (min: 10 mIU/mL; max: 176 mIU/mL) and ectopic pregnancy (min: 19 mIU/mL; max: 178 mIU/mL), clustered towards lower values of β-hCG, whereas viable pregnancies, such as singletons (min: 22 mIU/mL; max: 603 mIU/mL), twins (min: 63 mIU/mL; max: 784 mIU/mL), and triplets (min: 53 mIU/mL; max: 799 mIU/mL), clustered towards higher levels of β-hCG (Figure 1A). Interestingly, the aborted group (min: 17 mIU/mL; max: 506 mIU/mL) had similar values as the singleton group. When the groups were compared, it was confirmed that there were significant differences between the singleton, twin, triplet, and non-viable outcomes, except for the aborted group (Table 2). When the cohort was stratified by the source of the ova (patient versus donor), the transfer protocol (fresh versus frozen), and the number of embryos transferred (2 versus 3), there was no effect on the result, as indicated by the interaction p-value. Since the reason for miscarriages/abortions are numerous and are associated with external factors, this group was removed for future analyses. The density map shows that each group's distributions are spaced significantly (Figure 1B).

Figure 1: Serum human chorionic gonadotropin (hCG) concentrations by pregnancy outcome. A) A histogram was constructed using data from 663 in vitro fertilization cycles that was plotted against the outcome of the cycle as non-viable (pink; biochemical and ectopic), singleton (green), twin (blue), and triplet (purple). B) Based on the histogram results, a density plot was constructed in which the vertical dotted lines correspond to the mean value for non-viable (pink; biochemical and ectopic), singleton (green), twin (blue), and triplet (purple) pregnancies.

Table 2: The effects of IVF outcomes on β-hCG levels.

Non-viable outcomes				Viable outcomes			
Group	Biochemical	Aborted	Ectopic	Singleton	Twin	Triplet	p- value[g]
sample	81	74	9	265	202	32	N/A
Overall	48.4±39.0[b,d,e,f]	166.9±109.5[a,c,e,f]	79.3±49.3[b,d,e,f]	171.5±104.4[a,c,e,f]	303.8±153.0[a,b,c,d,f]	425.9±182.6[a,b,c,d,e]	
Ova source Patient							
Patient	52.4±38.4[b,d,e,f]	164.7±104.5[a,e,f]	95.5±59.9[e,f]	159.2±99.4[a,e,f]	291.1±160.2[a,b,c,d]	469.8±179.1[a,b,c,d]	0.361
Donor	42.0±39.9[b,d,e,f]	170.4±118.7[a,c,e,f]	59.0±26.6[b,d,e,f]	183.0±107.9[a,c,e,f]	313.0±147.5[a,b,c,d]	399.6±184.1[a,b,c,d]	
Transfer							
Fresh	47.1±41.8[b,d,e,f]	157.0±106.1[a,e,f]	78.1±56.5[d,e,f]	168.8±96.4[a,c,e,f]	296.9±149.0[a,b,c,d]	393.7±190.5[a,b,c,d]	0.302
Frozen	51.9±31.0[b,d,e,f]	195.7±117.0[a,c,e,f]	83.5±16.1[b,d,e,f]	180.8±128.8[a,c,e,f]	336.8±168.9[a,b,c,d,f]	522.5±118.9[a,b,c,d,e]	
# Embryos							
2 embryos	56.7±43.2[b,d,e]	152.0±102.9[a,e]	89.2±35.0[e]	179.0±105.8[a,e]	306.8±140.3[a,b,d]	Not applicable	0.779
3 embryos	40.7±33.4[b,d,e,f]	177.2±113.9[a,e,f]	74.3±57.5[e,f]	162.5±102.4[a,e,f]	301.0±164.5[a,b,c,d,f]	425.9±182.6[a,b,c,d,e]	

Values are mean ± standard deviation. Differences between groups were determined using 1-way ANOVA.

[a] vs. the biochemical group ($p<0.05$).

[b] vs. the aborted group ($p<0.05$).

[c] vs. the ectopic pregnancy group ($p<0.05$).

[d] vs. the singleton group ($p<0.05$).

[e] vs. the twin group ($p<0.05$).

[f] vs. the triplet group ($p<0.05$).

[g] Two-way ANOVA was used to determine if the ova source, the transfer method, or the number of embryos transferred affected the results. The p-value reported is for the interaction between the results and the variables mentioned above.

A ROC curve was constructed to assess the predictability of β-hCG for viable pregnancies. β-hCG was highly predictive of a viable pregnancy (AUC=0.939, 95%CI: 0.916-0.961, $p<0.001$, Figure 2A). This result was minimally affected if the aborted pregnancy group was included (AUC=0.807, 95%CI: 0.767-0.848, $p<0.001$, data not shown). Using the highest Youden index, the optimal cutoff of β -hCG was determined to be 126 mIU/mL (Youden index=0.483; sensitivity=0.752; specificity=0.732). Next, using 2-embryo transferred data, the predictability of β-hCG for a non-viable, singleton, and twin pregnancies was assessed; however, instead of using traditional ROC analysis, either one-versus-all or one-versus-one, VUS was utilized. For distinguishing between these 3 classes, β-hCG was highly selective (VUS=0.644, 95%CI: 0.559-0.690, $p<0.001$, Figure 2B). Using VUS, a cutoff of 74 and 245 mIU/mL was observed for determining between non-viable, singleton, and twins, respectively (Table 3). However, using the Youden index, the cutoffs were 121 (similar to the ROC results) and 265 mIU/mL, respectively. Nevertheless, the correct classification probability of singletons for VUS (0.68) was superior to the Youden index (0.50). To note, the sample size used (n=284) does exceed the minimal sample size required for VUS (n=172) and the Youden Index (n=109).

Table 3: Using data from 2-embryo transfers (n=284) to determine β-hCG cutoff values for non-viable, singleton, and twin births based on VUS and Youden index

Category	VUS	Youden Index
Effect size (95% confidence intervals)	0.644 (0.559-0.690)	0.521 (0.466-0.576)
Cutoff between non-viable and singleton	74	121
Cutoff between singleton and twins	245	265
Specificity	0.7143	0.9258
Sensitivity	0.6392	0.6138
Correct classification probability of singletons	0.6828	0.4978
Minimally sample size	172	109

Abbreviations: β-hCG: β-human chorionic gonadotropin, VUS: volume under the three-class receiver-operating characteristic surface.

To predict viable pregnancy outcomes (singleton, twin, and triplets), all data from viable pregnancies (2- and 3-embryo transfers) were used. Here, β-hCG was highly predictive for viable pregnancy outcome (VUS=0.479, 95%CI: 0.391-0.576, p<0.001, Figure 2C). Cutoffs, which are shown in Table 4, are highly variable; however, the correct classification probability of twins suggests that the results for VUS (0.65) are more accurate than the Youden index (0.33). Again, the sample size used (n=499) does exceed the minimal sample size required for VUS (n=167) and the Youden Index (n=117).

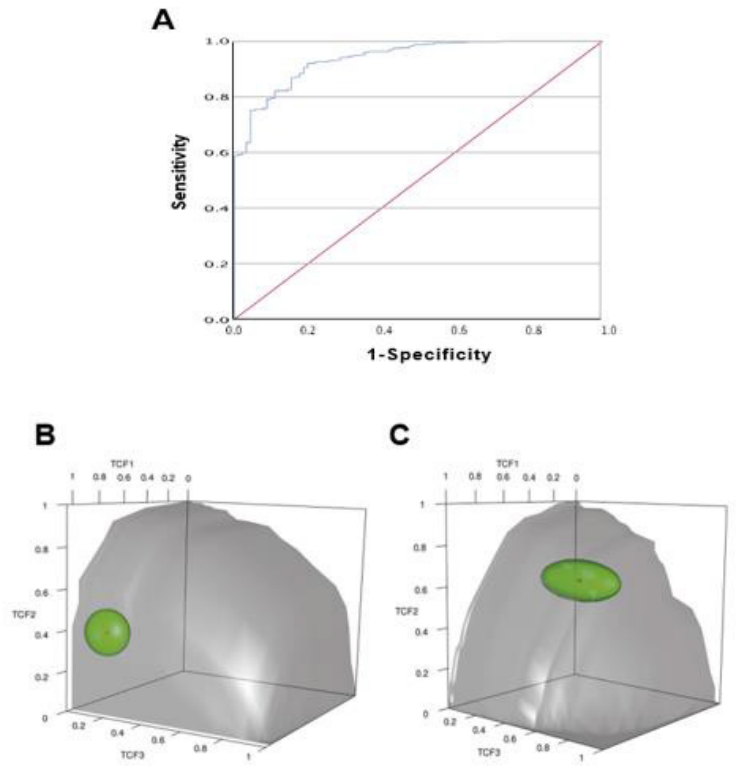

Figure 2: Serum human chorionic gonadotropin (hCG) concentrations can distinguish between pregnancy outcomes. A) Using all data, a receiver operating characteristic (ROC) curve was constructed to determine the predictability between serum β-hCG and pregnancy viability as measure by the area under a ROC curve (AUC=0.939, 95%CI: 0.916-0.961, p<0.001). The diagonal line represents AUC=0.500 (no predictability). B) Using only 2-embryo transfers, a ROC surface was constructed to assess the predictability of serum β-hCG for non-viable, singleton and twin pregnancies, where the 3-axes represent the True Class Fractions (TCFs) for each class. Predictability assessed using volumes under a ROC surface (VUS=0.644, 95%CI: 0.559-0.690, p<0.001). A VUS<0.167 indicates no predictability. The ellipsoid corresponds to the confidence region for TCFs. C) Using all viable data (2- and 3-embyro transfers), a ROC surface was constructed to assess the predictability of serum β-hCG for singleton, twin, and triplet pregnancies (VUS=0.479, 95%CI: 0.391-0.576, p<0.001).

Table 4: Using data from 2-embryo and 3-embryo transfers (n=499) to determine β-hCG cutoff values for singleton, twins, and triplet births based on VUS and Youden index.

Category	VUS	Youden Index
Effect size (95% confidence intervals)	0.479 (0.391-0.576)	0.353 (0.281-0.426)
Cutoff between singleton and twin	170.5	266.3
Cutoff between twin and triplet	427	398
Specificity	0.6038	0.8181
Sensitivity	0.5625	0.5612
Correct classification probability of twins	0.6485	0.3274
Minimally sample size	167	117

Abbreviations: β-hCG: β-human chorionic gonadotropin, VUS: volume under the three-class receiver-operating characteristic surface.

Discussion

This study aimed to determine the predictability of β-hCG for a viable pregnancy and determine if Embryonic Day 14 serum β-hCG concentrations could distinguish between a singleton, twin, or triplet pregnancy. Indeed, β-hCG was able to determine a viable pregnancy with specificity for the singleton, twin, and triplet pregnancies.

β-hCG is secreted by the trophoblastic cells, suggesting that its detection could be attained within 1 week after fertilization. Yuan et al. were able to confirm β-hCG's predictability between a biochemical pregnancy versus an ongoing pregnancy (4.34 U/L threshold, AUC=0.852) and an ongoing twin pregnancy with a single ongoing pregnancy (17.95 U/L threshold, AUC=0.903) [16]; however, the potential procedural differences could significantly affect the variability of serum β-hCG. In a study conducted by Kathiresan et al. they demonstrated that β-hCG's predictability is dependent on the embryo transfer day (D3 versus D5), where the threshold value to determine an ongoing pregnancy was 78 IU/L and 160 IU/L, respectively [24]. In addition, β-hCG's serum concentration depends on the number and the activity of the trophoblastic cells [25]. Therefore, it is more likely that procedural characteristics could affect serum β-hCG concentrations. Here, serum β-hCG was assessed on Embryonic Day 14, and, independently, if 2 or 3 embryos were transferred, serum β-hCG concentrations were similar for each pregnancy outcome. This is most likely due to the outcome of the pregnancy and not the initial number of embryos transferred.

Only high-quality embryos were transferred at our facilities as determined by the criteria established by the Istanbul Consensus Workshop on Embryo Assessment [26]. The embryo quality was similar between the donor and patient ova groups, which ranged between AA and BC, and the ova source did not affect the results. Interestingly, this result was expected. Women who typically produced poor-quality embryos in previous cycles could have access to high-quality embryos using donor ova. The effect embryo quality has on serum β-hCG has minimally been investigated. Kuspinar et al. demonstrated a correlation between serum β-hCG and blastocyst quality scores [27]. However, Steiner, et al. stated that there is no association between the blastocyst morphology and serum β-hCG levels; even though a very weak association was observed between Gardner's blastocyst grade and serum β-hCG levels on Day 16 [28]. Nevertheless, due to the causes of variability of serum β-hCG, this factor must be considered in a future study when assessing the predictability of β-hCG for a viable pregnancy, especially when discriminating between a singleton, twin, and triplet pregnancy.

When considering the transfer of frozen/thawed embryos versus fresh embryos, the transfer protocol did not affect serum β-hCG concentrations for each pregnancy outcome. This disagrees with the studies conducted by Zhao et al. and Keane et al. Zhao et al. indicated a significantly higher level of β-hCG, measured 12 days after transfer, for clinical pregnancy, early abortions, and live births for frozen embryo transfers; however, fresh embryo transfers were better predictors for clinical pregnancies and live births [29]. Similar results were observed by Keane et al. when β-hCG was measured 14 to 15 days after transfer for Day 3 embryos and 12 to 13 days after transfer for Day 5 embryos [30]. However, Reljic et al. determined that using fresh or frozen embryos did not affect serum β-hCG when measured 13 days after blastocyst transfer [31], which is similar to our results. Therefore, the effect of the type of embryo used during IVF remains inconclusive.

During the early stages of pregnancy, β-hCG linear increases, peaking between weeks 8 to 10 [25]. Since the serum concentration of β-hCG is dependent on the number of trophoblastic cells, it would be reasonable to assume serum β-hCG could predict the outcome of IVF and any timepoint after 7 days could be assessed. However, the ideal time point would need to occur when non-viable, singleton, twin, and triplet pregnancies could be determined with sufficient separation of serum β-hCG concentrations. We decided on using Embryonic Day 14, as this is a commonly used timepoint during IVF. Indeed, β-hCG was highly predictive of these IVF outcomes, in which the thresholds are, using VUS, 74, 171-245, and 427 mIU/mL to determine a non-viable, singleton, twin, and triplet pregnancy, or using the Youden Index, 121, 265, and 398 mIU/mL, respectively. The cutoffs reported here are similar to other reported studies, in which the cutoffs ranged between 111 to 213 mIU/mL and up to 400 mIU/ml for frozen/thawed embryo transfers [15,6,32,29,33,34]. Concerning viable pregnancy outcomes, there is a paucity of information. Nevertheless, Olgan et al. reported a cutoff between singleton and twin pregnancies of 175 mIU/ml measured 14 days after oocyte retrieval [35], which is close to our estimate. However, Wang et al. report much higher β-hCG cutoffs-values, measured 14 days after embryo transfer, for the singleton-twin threshold (987 mIU/mL) and for the twin-triplet threshold (2207 mIU/mL), which is significantly higher than our estimates. Furthermore, they also reported a cutoff value of 213 mIU/mL for clinical pregnancies, which exceeded our estimate. This points out the importance of comparing data to an established event during IVF. To improve the comparability of these studies, we suggest using the number of days post-fertilization. Nevertheless, these studies support the sequential increasing values proposed here as cutoff values to determine the IVF outcomes.

One strength of our study is using VUS. Many reports have demonstrated that β-hCG can be used to evaluate a pregnancy; however, most of these studies that used ROCs for their analysis fail to address the complexity of a multi-class outcome [36-38,21,39]. One concern is that the predictability for a biomarker is overestimating the effect due to the one-versus-one comparison. Here, we demonstrate that there is much overlap between each pregnancy category (Figure 1), and as shown above, the 2-class comparisons do give superior AUCs. However, using VUS for 3-class analysis for 2- embryo transfers, we see β-hCG's predictability is not as ideal (VUS=0.644).

Our study has a few limitations. First, the retrospective design of the study. A good prospective study should be designed to improve the quality of the results in which the same patients are assayed from multiple rounds of IVF to minimize certain biases. Second, even though the sample size of the study was large, we did not perform any adjustments to the effect sizes due to age, weight, or drug treatment. These factors are shown to affect β-hCG levels

[40-43]. Third, we exclude pregnancies that resulted in early or late abortions from determining cutoff values. The causes of an early pregnancy loss are numerous and can range from factors associated with the embryos, which could be determined with β-hCG, to factors associated with the mother, which could include insulin resistance, autoimmune disorders, and metabolic disorders, to name a few. Nevertheless, removing this group had minimal effect on the results. Fourth, the IVF cycles were categorized based on the outcome, not the number of gestational sacs. These factors could lead to some of the β-hCG's value-changing groups based on the vanishing twin syndrome. The rate of occurrence of this syndrome can be as high as 10% in ART [44].

Conclusion

Here, we confirm the predictability of serum β-hCG for determining a viable pregnancy on Embryonic Day 14 with an optimal cutoff of 126 mIU/mL. These results were independent of the ova source, if frozen/thawed or fresh embryos were used, or the number of embryos transferred. Moreover, serum β-hCG levels could selectively predict singleton, twin, and triplet pregnancies with cutoff values of 171-245 and 427 mIU/mL as determined using VUS or 265 and 398 mIU/mL as determined using the Youden Index, respectively.

Acknowledgement

We want to express our gratitude to this study's participants and the IVF and medical staff at Ingenes Monterrey. We thank Ing. Lucero Cervantes for their editorial assistance.

Statements and Declarations

Funding: The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Competing interests: The authors have no relevant financial or non- financial interests to disclose.

Authors' contributions: ERV conceived the project. ERV and LM performed clinical data acquisition, handled the cases as clinicians, and collected information regarding the parental history and IVF data. ERV, LMP, MEGM, and ELB analyzed and interpreted data. ELB and LMP drafted the article. MEGM critically revised the manuscript. All authors have approved the final version of the manuscript.

Ethics approval and consent to participate: This study was performed in line with the principles of the Declaration of Helsinki. Data were appropriately anonymized, and informed consent was obtained at the original data collection. Due to the retrospective nature of the study design, no Ethics Committee approval was needed.

References

1. Thornton KL (2000) Advances in assisted reproductive technologies. Obstet Gynecol Clin North Am 27(3): 517-527.
2. Brezina PR, Ning N, Mitchell E, Zacur HA, Baramki TA, Zhao Y (2012) Recent Advances in Assisted Reproductive Technology. Current Obstetrics and Gynecology Reports 1 (4): 166-173.
3. Fasouliotis SJ, Schenker JG (2000) Ethics and assisted reproduction. Eur J Obstet Gynecol Reprod Biol 90(2): 171-180.
4. Huang MZ, Sun YC, Gau ML, Puthussery S, Kao CH (2021) First-time mothers' experiences of foetal reduction in pregnancy following assisted reproductive technology treatment in Taiwan: a qualitative study. J Health Popul Nutr 40(1): 47.
5. Handschuh K, Guibourdenche J, Tsatsaris V, Guesnon M, Laurendeau I, et al. (2007) Human chorionic gonadotropin expression in human trophoblasts from early placenta: comparative study between villous and extravillous trophoblastic cells. Placenta 28(2-3): 175-184.
6. Wang Z, Gao Y, Zhang D, Li Y, Luo L, et al. (2020) Predictive value of serum β-human chorionic gonadotropin for early pregnancy outcomes. Arch Gynecol Obstet 301(1): 295-302.
7. Bree RL, Edwards M, Böhm-Vélez M, Beyler S, Roberts J, et al. (1989) Transvaginal sonography in the evaluation of normal early pregnancy: correlation with HCG level. AJR Am J Roentgenol 153(1): 75-79.
8. Lawler CC, Budrys NM, Rodgers AK, Holden A, Brzyski RG, et al. (2011) Serum beta human chorionic gonadotropin levels can inform outcome counseling after *in vitro* fertilization. Fertil Steril 96(2): 505-507.
9. Korevaar TI, Steegers EA, de Rijke YB, Schalekamp-Timmermans S, Visser WE, et al. (2015) Reference ranges and determinants of total hCG levels during pregnancy: the Generation R Study. Eur J Epidemiol 30(9): 1057-1066.
10. Namlı Kalem M, Kalem Z, GÜRgan T (2017) Factors Affecting Initial Beta-HCG Values in Pregnancies Achieved by Assisted Reproductive Techniques. Türk Üreme Tıbbı ve Cerrahisi Dergisi 1:67-74.
11. Hacivelioglu S, Uysal A, Gungor AN, Gencer M, Cakir DU, et al. (2015) The effect of maternal polycystic ovary morphology on first-trimester maternal serum biochemical markers of aneuploidy and fetal nuchal translucency thickness. Clin Exp Obstet Gynecol 42(1): 32-35.
12. Braunstein GD, Mills JL, Reed GF, Jovanovic LG, Holmes LB, et al. (1989) Comparison of serum placental protein hormone levels in diabetic and normal pregnancy. J Clin Endocrinol Metab 68(1): 3-8.
13. Goldstein J, Pandey P, Fleming N, Westin S, Piha-Paul S (2016) A non-pregnant woman with elevated beta-HCG: A case of para-neoplastic syndrome in ovarian cancer. Gynecol Oncol Rep 17: 49-52.
14. Kölbl AC, Schlenk K, Behrendt N, Andergassen U (2018) The importance of hCG in human endometrial adenocarcinoma and breast cancer. Int J Biol Markers 33(1): 33-39.
15. Qiu P, Wang Y, Ji H, Wang L, Lin J, et al. (2021) Predictive value of serum HCG concentrations for outcomes of vitrified-warmed blastocyst transfers in women of different ages. Reprod Biomed Online 43(5): 962-969.
16. Yuan L, Yu L, Sun Z, Song J, Xiao J, et al. (2020) Association between 7-day serum β-hCG levels after frozen-thawed embryo transfer and pregnancy outcomes: a single-centre retrospective study from China. BMJ Open 10(10): e035332.
17. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, et al. (2007) The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Bull World Health Organ 85(11): 867-872.
18. Ripley BD (2001) The R project in statistical computing. MSOR Connections. The newsletter of the LTSN Maths, Stats & OR Network, vol 1.
19. Mukaka MM (2012) Statistics corner: A guide to appropriate use of correlation coefficient in medical research. Malawi Med J 24(3): 69-71.
20. Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143(1): 29-36.
21. Nakas CT, Yiannoutsos CT (2004) Ordered multiple-class ROC analysis with continuous measurements. Stat Med 23(22): 3437-3449.

22. Luo J, Xiong C (2012) DiagTest3Grp: An R Package for Analyzing Diagnostic Tests with Three Ordinal Groups. J Stat Softw 51(3): 1-24.

23. To Duc K (2017) bcROCsurface: an R package for correcting verification bias in estimation of the ROC surface and its volume for continuous diagnostic tests. BMC Bioinformatics 18(1): 503.

24. Kathiresan AS, Cruz-Almeida Y, Barrionuevo MJ, Maxson WS, Hoffman DI, et al. (2011) Prognostic value of beta-human chorionic gonadotropin is dependent on day of embryo transfer during *in vitro* fertilization. Fertil Steril 96(6): 1362-1366.

25. Fournier T (2016) Human chorionic gonadotropin: Different glycoforms and biological activity depending on its source of production. Ann Endocrinol (Paris) 77(2): 75-81.

26. (2011) The Istanbul consensus workshop on embryo assessment: proceedings of an expert meeting. In: Hum Reprod 6: 1270-1283.

27. Kuspinar G, Kasapoglu I, Cakır C, Ata B, Uncu G, et al. (2019) What is the effect of embryo morphology on serum β-hCG levels? Eur J Obstet Gynecol Reprod Biol 233: 107-113.

28. Steiner N, Al Mamari N, Rotshenker-Olshinka K, Khayat S, Alzawawi N, et al. (2021) Blastocyst morphology has no relationship with serum β-hCG levels and live birth rates once pregnant. Eur J Obstet Gynecol Reprod Biol 258: 98-102.

29. Zhao WE, Li YJ, Ou JP, Sun P, Chen WQ, et al. (2017) Predictive value of initial serum human chorionic gonadotropin levels for pregnancies after single fresh and frozen blastocyst transfer. J Huazhong Univ Sci Technolog Med Sci 37(3): 395-400.

30. Keane KN, Mustafa KB, Hinchliffe P, Conceicao J, Yovich JL (2016) Higher β-HCG concentrations and higher birthweights ensue from single vitrified embryo transfers. Reprod Biomed Online 33(2): 149-160.

31. Reljič M, Knez J, Vlaisavljević V (2013) Human chorionic gonadotropin levels are equally predictive for pregnancy outcome after fresh and vitrified- warmed blastocyst transfer. J Assist Reprod Genet 30(11): 1459-1463.

32. Xiong F, Li G, Sun Q, Chen P, Wang Z, et al. (2019) Obstetric and perinatal outcomes of pregnancies according to initial maternal serum HCG concentrations after vitrified-warmed single blastocyst transfer. Reprod Biomed Online 38(3): 455-464.

33. Oron G, Shavit T, Esh-Broder E, Weon-Young S, Tulandi T, et al. (2017) Predictive value of serum HCG concentrations in pregnancies achieved after single fresh or vitrified-warmed blastocyst transfer. Reprod Biomed Online 35(3): 272-278.

34. Løssl K, Oldenburg A, Toftager M, Bogstad J, Praetorius L, et al. (2017) Predictive value of plasma human chorionic gonadotropin measured 14 days after Day-2 single embryo transfer. Acta Obstet Gynecol Scand 96 (8): 960-967.

35. Olgan S, Bozdag G, Sokmensuer LK, Mumusoglu S, Gunalp S (2016) Revisiting serum beta-human chorionic gonadotropin concentrations as a predictor for dizygotic twinning after *in vitro* fertilization. Clin Exp Obstet Gynecol 43(4):597-601.

36. Ferri C, Hernández-Orallo J, Salido MA (2003) Volume under the ROC Surface for Multi-class Problems. In: Lavrač N, Gamberger D, Blockeel H, Todorovski L (eds) Machine Learning: ECML 2003, Berlin, Heidelberg, Springer Berlin Heidelberg, pp 108-120.

37. He X, Frey EC (2008) The meaning and use of the volume under a three-class ROC surface (VUS). IEEE Trans Med Imaging 27(5): 577-588.

38. Sampat MP, Patel AC, Wang Y, Gupta S, Kan CW, et al. (2009) Indexes for three-class classification performance assessment--an empirical comparison. IEEE Trans Inf Technol Biomed 13(3): 300-312.

39. Liu S, Xu W, Sun X, Zhang Y (2019) Fast and Unbiased Estimation of Volume Under Ordered Three-Class ROC Surface (VUS) Based on Dynamic Programming. IEEE Access 7: 63972-63982.

40. Ueno S, Ezoe K, Abe T, Yabuuchi A, Uchiyama K, et al. (2014) Maternal age and initial β-hCG levels predict pregnancy outcome after single vitrified-warmed blastocyst transfer. J Assist Reprod Genet 31(9):1175-1181.

41. Haavaldsen C, Fedorcsak P, Tanbo T, Eskild A (2014) Maternal age and serum concentration of human chorionic gonadotropin in early pregnancy. Acta Obstet Gynecol Scand 93(12):1290-1294.

42. Choux C, Barberet J, Ginod P, Cottenet J, Bruno C, et al. (2017) Severe ovarian hyperstimulation syndrome modifies early maternal serum beta-human chorionic gonadotropin kinetics, but obstetrical and neonatal outcomes are not impacted. Fertil Steril 108(4): 650-658.

43. Liu S, Kuang Y, Wu Y, Feng Y, Lyu Q, et al. (2017) High oestradiol concentration after ovarian stimulation is associated with lower maternal serum beta-HCG concentration and neonatal birth weight. Reprod Biomed Online 35(2):189-196.

44. Poikkeus P, Gissler M, Unkila-Kallio L, Hyden-Granskog C, Tiitinen A (2007) Obstetric and neonatal outcome after single embryo transfer. Hum Reprod 22(4):1073-1079.