ORIGINAL ARTICLE

EStradiol and PRogesterone in In vitro ferTilization (ESPRIT): a multicenter study evaluating third- versus second-generation estradiol and progesterone immunoassays

N. P. Polyzos[1,2,3] · E. Anckaert[4] · P. Drakopoulos[5] · H. Tournaye[5] · J. Schiettecatte[4] · H. Donner[6] · G. Bobba[7] · G. Miles[8] · W. D. J. Verhagen-Kamerbeek[7] · E. Bosch[9]

Abstract

Purpose To assess estradiol (E_2) and progesterone levels during ovarian stimulation determined by third-generation (Gen III) and second-generation (Gen II) Elecsys® immunoassays.

Methods E_2 and progesterone concentrations were measured using Elecsys® Gen III and Gen II immunoassays, and progesterone concentrations on the day of ovulation triggering were determined by LC–MS/MS. This was a retrospective, non-interventional study conducted at European tertiary referral infertility clinics in women aged 18–45 years, with a body mass index 18–35 kg/m^2, regular menses, and both ovaries.

Results Serum samples were obtained from 230 women classified by oocyte retrieval as poor (33.0%; 0–3 oocytes), normal (40.9%; 4–15 oocytes), or high (26.1%; > 15 oocytes) responders. E_2 and progesterone levels increased during ovarian stimulation, with greatest increases observed in high responders. Elecsys® Gen III and Gen II assay results were highly correlated for E_2 (Pearson's $r = 0.99$) and progesterone ($r = 0.89$); Gen III results were lower than Gen II for both E_2 and progesterone. On the day of triggering, Gen III E_2 and progesterone levels showed a difference of − 15.0% and − 27.9%, respectively. Progesterone levels (on day of triggering) measured by LC–MS/MS correlated better with Gen III (0.98) than Gen II (0.90). Mean relative differences for Gen III and Gen II assays versus LC–MS/MS were 14.6% and 62.8%, respectively.

Conclusion E_2 and progesterone levels determined with Elecsys® Gen II and III assays were highly correlated; results were lower for Gen III versus Gen II. Differences observed for progesterone on the day of triggering may be clinically relevant.

Keywords Estradiol · Progesterone · Immunoassay · Ovarian stimulation · In vitro fertilization

✉ N. P. Polyzos
nikpol@dexeus.com; n.polyzos@gmail.com

1 Department of Reproductive Medicine, Dexeus University Hospital, Gran Via Carles, 11171-75, 08028 Barcelona, Spain

2 Faculty of Medicine and Pharmacy, Vrije Universiteit Brussel, Laarbeeklaan 101, 1090 Brussels, Belgium

3 Department of Clinical Medicine, Faculty of Health, Aarhus University Incuba/Skejby, Building 2, Palle Juul-Jensens Boulevard 82, Aarhus N, 8200 Aarhus, Denmark

4 Laboratory of Hormonology and Tumour Markers, Universitair Ziekenhuis Brussel, Free University of Brussels, Brussels, Belgium

5 Centre for Reproductive Medicine, University Hospital (UZB), Laarbeeklaan 101, 1090 Jette, Brussels, Belgium

6 Medical and Scientific Affairs, Roche Diagnostics GmbH, Nonnenwald 2, 82377 Penzberg, Germany

7 Medical and Scientific Affairs, Roche Diagnostics International Ltd, Forrenstrasse 2, 6343 Rotkreuz, Switzerland

8 Biostatistics and Data Management, Roche Diagnostics Operations Inc., 9115 Hague Road, Building R, Indianapolis, IN 46256, USA

9 Human Reproduction Unit, IVI-RMA, Plaza de la Policia Local 3, 46015 Valencia, Spain

Introduction

The measurement of estradiol (E_2) and progesterone is an essential part of ovarian stimulation for in vitro fertilization (IVF) and intracytoplasmic sperm injection (ICSI). The accurate monitoring of E_2 levels, along with ultrasound, are key aspects of IVF that support dose adjustment and evaluation of ovarian hyperstimulation syndrome risk and potential cycle cancellation. Serum progesterone measurement during the follicular phase has become part of routine clinical practice in the last decade, mainly due to the inverse association between late follicular phase progesterone levels and pregnancy rates [1, 2].

As hormone measurement during IVF or ICSI cycles is important for optimizing ovarian response and treatment safety and efficacy, several different assays have been developed for the measurement of E_2 and progesterone serum levels. Automated analyzers and direct immunoassays are commonly used, but previous studies have shown that different assays demonstrate various degrees of bias and interlaboratory variability [1, 3–7]. Liquid chromatography (LC) coupled with mass spectrometry (MS) methods have the potential to provide greater specificity and sensitivity than immunologic methods, but doubts have been expressed over discrepancies between different assays [8]. Moreover, the "gold standard" method of isotope dilution (ID) with gas chromatography (GC)/MS (ID-GC/MS) is complex, and the throughput is too low for routine clinical use [8]. Modern assays need to fulfil a number of criteria: have good precision to characterize patient response to treatment; be reliable over a wide concentration range; have high specificity for E_2 and be able to exclude exogenous estrogens; and provide good interlaboratory reproducibility [8]. The Elecsys® second-generation (Gen II) Estradiol and Progesterone immunoassays have been widely used for monitoring serum E_2 and progesterone levels. The Elecsys® third-generation (Gen III) Estradiol and Progesterone immunoassays have demonstrated good correlation with Gen II assays during assay development, but IVF patient samples were not available for validation.

Given the lack of evidence, the ESPRIT (EStradiol and PRogesterone in In vitro ferTilization) study was conducted to determine E_2 and progesterone levels during ovarian stimulation in a population of poor, normal, and high responders [9] who were undergoing IVF using gonadotropin-releasing hormone (GnRH) agonist or antagonist protocols in a routine clinical setting. The ESPRIT study aimed to compare the recently developed Elecsys® Estradiol and Progesterone Gen III assays with the respective commonly used Gen II assays. The Elecsys® Gen III and Gen II Progesterone assays were also compared with a LC–MS/MS method to further assess assay performance.

Materials and methods

Study design

ESPRIT was an exploratory, retrospective, non-interventional, multicenter study. The study used serum samples that were originally collected during the stimulation cycle from patients on GnRH agonist or antagonist protocols at two tertiary referral infertility clinics in Europe (IVI-RMA Valencia [IVI], Spain; UZ Brussel [UZB], Belgium).

The study included stored samples from women aged 18–45 years, with a body mass index of 18–35 kg/m^2, and who had both ovaries present and regular ovulatory menses (every 25–35 days). Samples were excluded if women had current or past disease affecting the ovaries or gonadotropin/sex steroid hormone levels, polycystic ovary syndrome, or any known untreated endocrine abnormality, or were undergoing hormone therapy at the time of blood sampling.

Ethics approval

All participants provided written informed consent prior to enrolment. The protocol was approved by the Clinical Research Ethics Committee of the IVI-RMA Valencia (December 20, 2016) and the Medical Ethics Committee of the Free University of Brussels (January 25, 2017).

Immunoassays

The Elecsys® Estradiol Gen III and Gen II assays are used for in vitro quantitative determination of E_2 in human serum and plasma on **cobas e** analyzers (Roche Diagnostics GmbH, Mannheim, Germany). Measuring ranges for each assay (defined by the lower limit of detection and the maximum of the master curve) are as follows: Elecsys® Estradiol Gen III assay, 18.4–11,010 pmol/L (5.0–3000 pg/mL); Elecsys® Estradiol Gen II assay, 18.4–15,781 pmol/L (5.0–4300 pg/mL) (Roche, data on file). Sample dilution allows measurement of higher levels up to 30,000 pg/mL for Gen III (recommended dilution ratio of 1:10) and up to 21,500 pg/mL for Gen II (recommended dilution ratio of 1:5). Standardization is achieved in the E_2 assays using a panel of samples with ID-GC/MS-assigned target values [10]. The Gen II E_2 assay uses polyclonal antibodies, whereas the Gen III E_2 assay uses high-affinity monoclonal antibodies, which permit higher sensitivity. The Gen III assay also uses a lower concentration of antibodies than the Gen II assay, and thus, cross-reactivity is lower for Gen III. Clinical and Laboratory Standards Institute (CLSI) protocol EP05-A2 precision and repeatability profiles for the Gen III assay are similar to the Gen II assay on the **cobas e** 601 analyzer, and are improved at some concentrations (Roche, data on file).

The Elecsys® Progesterone Gen III and Gen II assays determine progesterone levels in human serum and plasma on **cobas e** analyzers. Measuring ranges for each assay are as follows: Elecsys® Progesterone Gen III, 0.159–191 nmol/L (0.050–60.0 ng/mL); Elecsys® Progesterone Gen II, 0.095–191 nmol/L (0.030–60.0 ng/mL) (Roche, data on file). The Elecsys® Progesterone Gen II assay employs mouse monoclonal antibodies, whereas the Elecsys® Progesterone Gen III assay uses sheep monoclonal antibodies, which provide higher specificity for progesterone (Roche, data on file) and reduced cross-reactivity against major steroids and metabolites.

LC–MS/MS

Progesterone concentrations determined by the Elecsys® Progesterone Gen II and Gen III assays were compared with those measured by LC–MS/MS in 148 samples at UZB (treatment groups: GnRH agonist, $n = 60$; GnRH antagonist, $n = 88$; response groups: poor, $n = 58$; normal, $n = 60$; high, $n = 30$). The determination of serum progesterone is based on ID–LC–MS/MS, subsequent to a sample preparation procedure involving protein precipitation and ultra-centrifugation. Dissolved deuterated progesterone was added as an internal standard and simultaneously as a protein precipitation reagent to samples and calibrators. After precipitation, the samples were centrifuged and the supernatant was transferred to centrifugation filters. The filtrate of unknown serum samples and calibrators was injected into the high-performance LC system to separate the analytes from other matrix components, using an analytical column in the reverse-phase mode. Each analyte was then detected in the MS (TSQ Quantum Ultra; Thermo Fisher Scientific) with atmospheric pressure ionization and in selected reaction monitoring mode. Quantification was based on the area ratio between analyte and internal standard. Further details are available in Online Resource 1.

Study procedures

Demographic and health assessment information was collected for all patients. Samples were collected according to routine clinical practice at each site (IVI: Visit 1, Day 5 ± 1 of ovarian stimulation; Visit 2, Day 7 ± 1; Visit 3, Day $9 + 3$; Visit 4, day of ovulation triggering; UZB: Visit 1, Day 0/1 of ovarian stimulation; Visit 2, Day 8; Visit 3, Day 10–11; Visit 4, day of ovulation triggering). All follicles ≥ 10 mm were measured in sagittal and transverse planes by 2D ultrasound, and the mean diameter of each follicle was recorded; small (< 12 mm), intermediate (12–15 mm), and mature (≥ 16 mm) follicle counts were recorded at each visit.

E_2 and progesterone concentrations in each sample were determined retrospectively using the Elecsys® Gen III and Gen II assays (per the manufacturer's instructions) on **cobas e** 601 (UZB) or **cobas e** 411 (IVI) analyzers. Progesterone concentrations in the samples collected on the day of ovulation triggering were also determined by LC–MS/MS (UZB only).

Statistical methods

A simulation was performed to estimate the precision of the 95% confidence interval (CI) for the difference between Elecsys® Estradiol Gen III and Gen II assays (E_2 has a greater measuring range and a larger bias between assay generations than progesterone). Based on these calculations and pooling across timepoints ($N = 120$), a sample size of ≥ 30 patients per subgroup would attain an average precision on the 95% CI of differences between assay generations of $\pm 2.7\%$ of the mean E_2 value. We assumed a bias between Gen III and Gen II assays, consistent with that reported from data on technical performance studies with other systems, which could be extended to the measurement range expected for this study: Passing–Bablok 0.86 (E_2 Gen II) + 8.27; E_2 range 1450–10,000 pmol/L [11]. The bias is constant over stimulation cycles and on day of ovulation triggering.

Patient baseline characteristics/demographics were summarized using descriptive statistics (mean, median, standard deviation [SD], minimum, maximum, and percentiles [5%, 25%, 75%, and 95%]). E_2 and progesterone concentrations were presented by site, GnRH therapy protocol (agonist/antagonist), and ovarian response (poor/normal/high). To account for potential effects of different gonadotropin use, E_2 levels were calculated per mature follicle on day of ovulation triggering. Responses to ovarian stimulation were predefined based on the number of oocytes retrieved following ovarian stimulation as poor (0–3), normal (4–15), or high (> 15). Mean relative difference (Gen III vs. Gen II) and associated SDs, range, and upper and lower limits (equal to ± 2SD) were calculated relative to Gen II for E_2 and progesterone. Method comparisons (Gen III vs. Gen II; Progesterone Gen III vs. LC–MS/MS; Progesterone Gen II vs. LC–MS/MS) were performed using Passing–Bablok regression analysis and Pearson's correlation coefficients were calculated; bias was examined using Bland–Altman analysis. Analyses were performed in subgroups at truncated ranges to assess variability at the proposed 1.5 ng/mL progesterone threshold (obtained previously using AxSYM system [Abbott Científica S.A., Madrid, Spain]) [1, 12, 13]. Statistical analyses were performed using SAS software (version 9.4; SAS System for

Windows 7) and R (version 3.4.0). Regression analyses were performed using the R package mcr (version 1.2.1).

Results

Study population

Patient characteristics were representative of a European population (Online Resource 2). In the overall cohort of 230 women, median age was 36 years (range, 25–44 years) and 67.8% of women were White/Caucasian. Patients were classified as poor responders (0–3 oocytes; $n = 76$ [33.0%]), normal responders (4–15 oocytes; $n = 94$ [40.9%]), or high responders (> 15 oocytes; $n = 60$ [26.1%]) based on the number of oocytes retrieved; as expected, response to ovarian stimulation decreased with age (Online Resource 2).

A total of 62 patients received a GnRH agonist protocol and 168 patients received a GnRH antagonist protocol. Median age was comparable in the overall population and in subgroups of patients who underwent GnRH agonist or antagonist protocols.

E$_2$ and progesterone levels and response status

E$_2$ and progesterone levels increased during ovarian stimulation, with greater increases in high responders versus poor or normal responders (Fig. 1). On the day of ovulation triggering, progesterone levels increased versus previous visits in both agonist- and antagonist-treated patients, regardless of responder status. The ratios of mean analyte concentration to mature follicle count on day of ovulation triggering were in the range 324–411 (E$_2$ Gen III) and 0.11–0.24 (progesterone Gen III) across protocols and response groups, based on follicle counts (Online Resource 3 and 4).

Method comparison of Gen III and Gen II assays

Method comparison based on Passing–Bablok regression demonstrated a good correlation between Elecsys® Gen III and Gen II assays for E$_2$ ($y = 0.86x - 4.35$; 0–12,500 pg/mL; Pearson's $r = 0.99$), but correlation was weaker for progesterone ($y = 0.86x - 0.20$; 0–5 ng/mL; Pearson's $r = 0.89$).

E$_2$ concentrations were lower when measured with the Gen III assay versus the Gen II assay (Figs. 1a–b, 2a; Table 1; Online Resource 5 and 6). The mean relative difference between Gen III and Gen II assays was − 15.8% (SD, 6.5%; $N = 684$ samples) across visits; on the day of ovulation triggering, the mean relative difference was − 15.0%. Bland–Altman analysis showed a constant bias up to ~ 3000 pg/mL; at concentrations > 3000 pg/mL, Gen III had a lower recovery than Gen II (Fig. 2a).

Progesterone concentrations were also lower when measured with the Gen III assay compared with the Gen II assay (Figs. 1c–d, 2b; Table 1; Online Resource 5 and 7). The mean relative difference between Gen III and Gen II assays was − 41.7% (SD, 25.8; $N = 674$ samples) across visits; on the day of ovulation triggering, the mean relative difference was − 27.9%. For 20/36 (56%) patients with Gen II progesterone levels > 1.5 ng/mL, results for Gen III were concordant; however, 16/36 (44%) patients had a progesterone level < 1.5 ng/mL with Gen III.

### Method comparison of progesterone Gen III and Gen II assays versus LC–MS/MS

Passing–Bablok analyses showed a higher correlation between the progesterone Gen III assay and LC–MS/MS ($y = 1.02*x + 0.04$; Pearson's $r = 0.98$; $N = 148$ samples; Fig. 3a) compared with the progesterone Gen II assay and LC–MS/MS ($y = 1.01*x + 0.28$; Pearson's $r = 0.90$; $N = 148$ samples; Online Resource 8).

Bland–Altman analyses showed that the mean (SD) relative difference versus LC–MS/MS was lower with Gen III (14.6% [23.8%]) than with Gen II (62.8% [69.2%]) (Fig. 3b; Online Resource 8 and 9); this trend was most pronounced for progesterone concentrations < 0.5 ng/mL. Relative to LC–MS/MS, a positive bias was observed with Gen II across the whole measuring range (0–5 ng/mL; Online Resource 8) compared with a more uniform bias (trending toward positive) with Gen III (Fig. 3b). There were fewer samples above the proposed 1.5 ng/mL threshold with Gen III (16/148 [10.8%]) compared with Gen II (28/148 [18.9%]) (Online Resource 9).

Discussion

Measurement of hormone levels during ovarian stimulation has an important role in optimizing the likelihood that IVF will result in pregnancy. Accurate and reliable methods to measure E$_2$ and progesterone are essential to assess treatment response, and support clinical decision-making relating to dose adjustment and the assessment of ovarian hyperstimulation syndrome risk. Our findings demonstrate that E$_2$ levels determined with the Elecsys® Estradiol III assay are highly correlated with results determined using the well-established Elecsys® Estradiol II assay ($r = 0.99$); the correlation between Elecsys® Progesterone III and Elecsys® Progesterone II assay results was slightly weaker ($r = 0.89$). We also show that although E$_2$ and progesterone levels determined by the Gen III assays were lower than with the Gen II assays, the differences are only of potential clinical relevance for progesterone.

Fig. 1 Change in estradiol concentration measured using the Elecsys® Estradiol Gen III assay by ovarian response group (poor, 0–3 oocytes; normal, 4–15 oocytes; high, > 15 oocytes) following gonadotropin-releasing hormone **a** agonist and **b** antagonist treatment. Change in progesterone concentration measured using the Elecsys® Progesterone Gen III assay by ovarian response group (poor, 0–3 oocytes; normal, 4–15 oocytes; high, > 15 oocytes) following gonadotropin-releasing hormone **c** agonist and **d** antagonist treatment
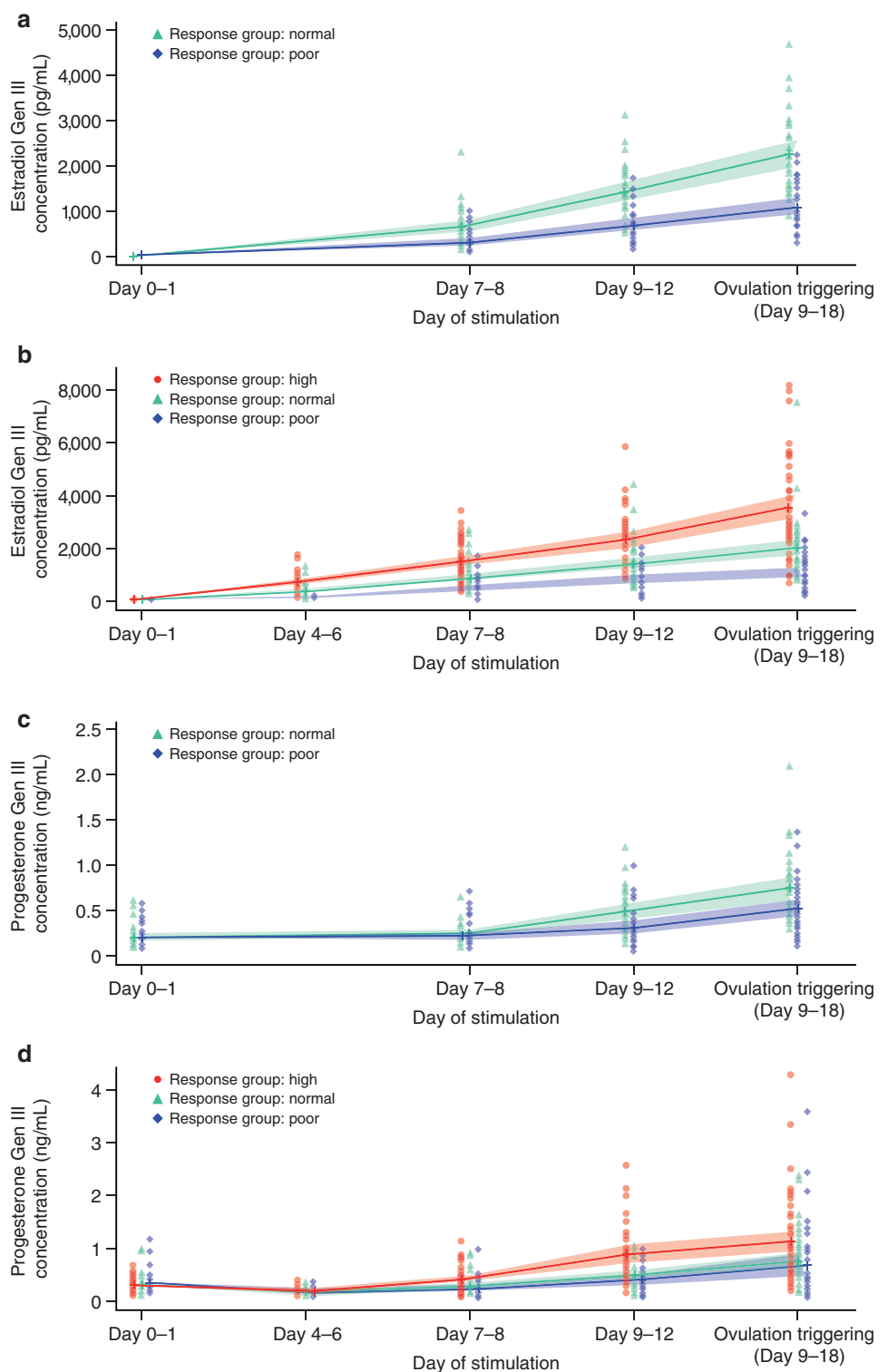
The lower progesterone levels measured with the Elecsys® Progesterone Gen III assay versus the Elecsys® Progesterone Gen II assay were particularly apparent in the lower concentration range (0–1 ng/mL). The magnitude of this difference was ~ 28% on day of ovulation triggering, which may have clinical implications. This may be particularly relevant if a "freeze-all" strategy is considered alongside segmentation of treatment [14], based on late follicular phase progesterone elevation to avoid a detrimental effect on pregnancy rates [15]. To further investigate the potential differences between the Elecsys® Gen III and Gen II Progesterone assays, we compared measurements obtained
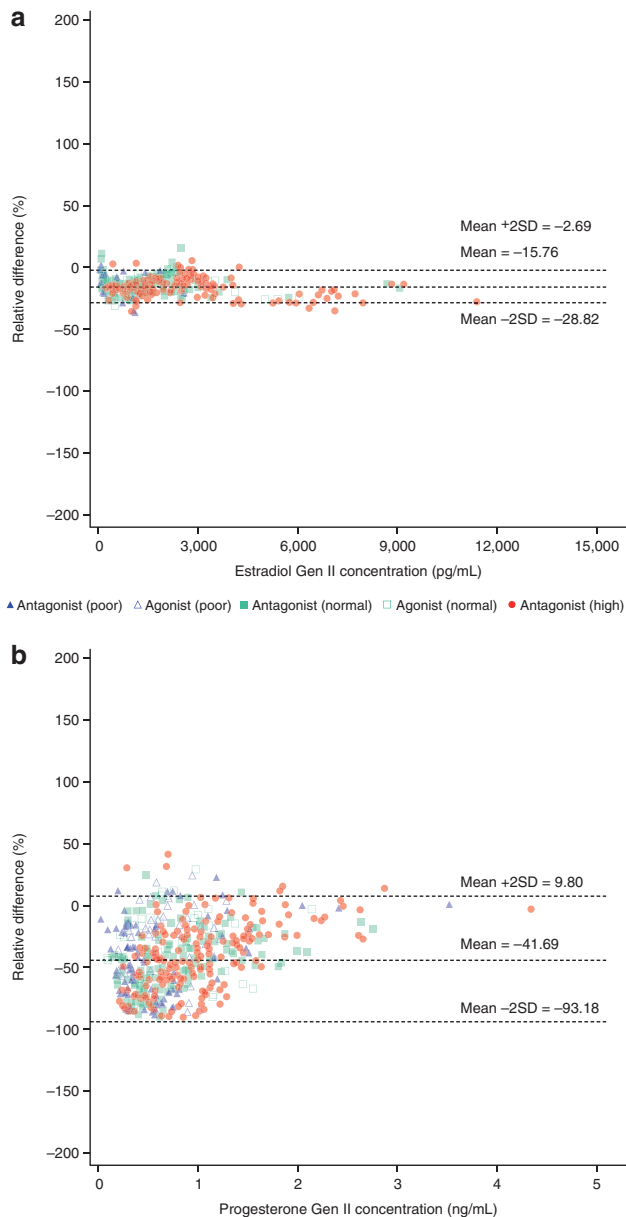
Fig. 2 Bland–Altman plots for the comparison of **a** estradiol and **b** progesterone concentrations measured using the Elecsys® Gen III and Gen II assays. *2SD* two standard deviations

using the assays and LC–MS/MS on the day of ovulation triggering. These data confirmed that Gen III results correlate better with LC–MS/MS than Gen II (Pearson's $r = 0.98$ vs. 0.90, respectively) and show improved accuracy; this is likely due to the higher specificity and reduced cross-reactivity of antibodies used in the Gen III assay (Roche, data on file).

Previous findings support the use of a threshold of 1.5 ng/mL to define progesterone levels associated with pregnancy rates [1, 12, 13]; however, this threshold was obtained using the AxSYM system, and is not applicable to all assays or

LC–MS/MS methods. Based on our findings, and if the threshold was applicable to the LC–MS/MS method used in the present study, the difference between progesterone results determined with each assay generation would translate into fewer misclassifications for Gen III versus Gen II. Clinicians changing from Gen II to Gen III progesterone assays should be aware of these differences when making clinical decisions, especially when measuring progesterone on the day of ovulation triggering. Our findings are generally consistent with earlier studies, where the Gen II assay demonstrated good correlation with LC–MS/MS (Pearson's $r = 0.987$), but delivered higher values than LC–MS/MS at low concentrations [3]. Moreover, the Gen II assay showed good intra- and inter-assay precision, with coefficients of variation < 10% at all timepoints and concentrations tested, the lowest of all four analyzers assessed [3].

Good correlation was observed between the Elecsys® Gen III and Gen II Estradiol assays. Although E_2 levels were lower when measured with the Gen III assay compared with the Gen II assay above ~3000 pg/mL, the magnitude of this difference was not considered to be clinically relevant (~15% on day of ovulation triggering). This small difference is potentially explained by the use of additional samples in the standardization of the Gen III assay, which are better distributed to cover the measuring range (Roche, data on file). For the Gen III assay, E_2 concentrations were similar for the poor-, normal-, and high-response groups when normalized to follicle count (e.g., 324–377 pg/mL for the GnRH antagonist patients); this was expected, as E_2 production per follicle was expected to be the same across response groups. Taieb and colleagues previously found mean E_2 concentrations per mature follicle of ~1200 pmol/L (327 pg/mL) for women undergoing GnRH agonism or antagonism in an assessment of another automated immunoassay [11]. Furthermore, in their retrospective review of 342 IVF cycles in patients undergoing a long GnRH agonist protocol in which the E_2/follicle ratio was used, Mittal and colleagues found a positive correlation between the ratios of E_2 to mature follicles (> 14 mm diameter) and E_2 to oocytes ($p = 0.0001$), and showed that pregnancy rates increased when the E_2 concentration per follicle was between 200 and 300 pg/mL [16].

Reliable assays that are specific to the measurement of E_2 are required to precisely quantify E_2 levels in women undergoing ovulation induction. These assays need to be able to measure both low baseline and high (i.e., 250–2000 pg/mL) E_2 levels to assess treatment efficacy, the timing of ovulation triggering, and with cut-offs designed to abort cycles that risk hyperstimulation [8]. In addition, long-standing evidence strongly suggests that progesterone levels on the day of ovulation triggering are related to IVF outcomes [1, 2, 4, 17–23]. Some data have been contradictory [24–26], but the balance of evidence and the most recent systematic reviews [2, 27] support an association between progesterone

Table 1 Relative difference in E_2 and progesterone levels as measured with the Elecsys® Gen III and Gen II assays

Test	Measurement range	N	Relative difference (%)			
			Mean (SD)	Min, max	Lower limit (mean relative difference − 2SD)	Upper limit (mean relative difference + 2SD)
E_2 (pg/mL)	Full	684	− 15.76 (6.53)	− 35.71, 14.29	− 28.82	− 2.69
Progesterone (ng/mL)	Full	674	− 41.69 (25.75)	− 90.91, 42.66	− 93.18	9.80
	Truncated: 0–1	493	− 46.32 (25.37)	− 90.91, 42.66	− 97.06	4.43
	Truncated: > 1–1.5	129	− 33.55 (22.16)	− 86.12, 25.00	− 77.87	10.78
	Truncated: > 1.5	52	− 18.04 (18.97)	− 67.32, 16.22	− 55.97	19.89

Data are combined by protocol, response group, and site (Days 0–1 excluded); differences are expressed as Elecsys® Gen III versus Gen II assays

2SD two standard deviations, E_2 estradiol, *SD* standard deviation

levels on the day of final oocyte maturation and successful pregnancy outcomes. The Gen III Elecsys® Estradiol and Progesterone assays will enable more accurate and reliable monitoring of E_2 and progesterone levels to support clinical decision-making, including the adjustment of hormone levels during ovarian stimulation to optimize clinical outcome and the likelihood of pregnancy.

An important limitation of this study is that it was retrospective, and thus required reanalysis of frozen/thawed samples. Nonetheless, this is unlikely to have influenced our results given that previous findings support the use of frozen samples [28]. Furthermore, the intra- and inter-sonographer variation in the measurement of follicular diameter could have been associated with a potential source, although the use of experienced sonographers was expected to minimize this limitation. However, ultrasound follicular measurement was not at all associated with primary outcome assessment, which was the correlation between the two different generation assays.

Considering all this, our findings have important implications for clinical practice. During cycle monitoring, clinicians should be aware that progesterone levels on the day of final oocyte maturation are ~30% lower when measured with the Elecsys® Gen III assay versus the Elecsys® Gen II assay, probably due to the higher specificity and reduced cross-reactivity of antibodies used in the Gen III assay. Consequently, treatment strategies should be adapted accordingly taking into account this difference.

Conclusions

The ESPRIT study expands the evidence surrounding E_2 and progesterone testing in the IVF setting. The Elecsys® Gen II and Gen III assays for the measurement of E_2 demonstrated a high level of correlation; the correlation for progesterone measurement was slightly weaker. Progesterone levels measured by the Gen III assay were lower than those measured by the Gen II assay, and more similar to levels measured by LC–MS/MS. Clinicians changing from Elecsys® Gen II to Gen III Elecsys® assays for measuring progesterone should be aware of these differences, which are potentially clinically relevant, especially when measuring progesterone levels on day of ovulation triggering.
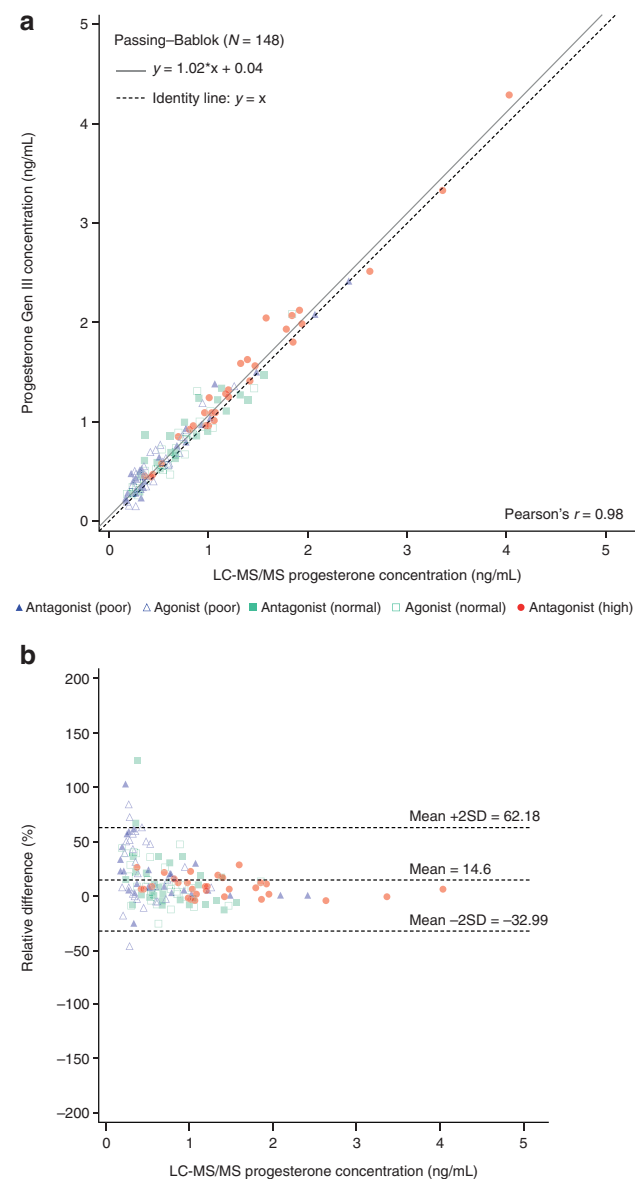
Fig. 3 **a** Passing–Blok regression analysis and **b** Bland–Altman plot for the comparison of progesterone concentrations measured using the Elecsys® Progesterone Gen III assay and LC–MS/MS. *2SD* two standard deviations, *LC–MS/MS* liquid chromatography–tandem-mass spectrometry

Data availability The data sets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Compliance with ethical standards

Conflict of interest N.P.P. has received honoraria for lecturing and/or research grants from MSD, Ferring International, Merck, Roche Diagnostics, IBSA, BESINS International, and Theramex. E.A. has nothing to disclose. P.D. has received honoraria for lecturing and/or research grants from MSD, Ferring International, and Merck. H.T. has received research grants from Merck Serono, MSD, Goodlife, Cook, Roche, Besins, Ferring, Research Fund of Flanders (FWO) and Mirtha (now Allergan), and consultancy fees from Finox, Abbott, ObsEva and Merck. J.S. has nothing to disclose. H.D., G.B., G.M., and W.V., are employees of Roche Diagnostics. E.B. has received fees from Ferring Pharmaceuticals, Merck Serono, MSD, and Roche Diagnostics, and fees and grant support from Finox.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Bosch E, Labarta E, Crespo J, Simón C, Remohí J, Jenkins J, Pellicer A (2010) Circulating progesterone levels and ongoing pregnancy rates in controlled ovarian stimulation cycles for in vitro fertilization: analysis of over 4000 cycles. Hum Reprod 25:2092–2100. https://doi.org/10.1093/humrep/deq125

2. Venetis CA, Kolibianakis EM, Bosdou JK, Tarlatzis BC (2013) Progesterone elevation and probability of pregnancy after IVF: a systematic review and meta-analysis of over 60 000 cycles. Hum Reprod Update 19:433–457. https://doi.org/10.1093/humupd/dmt014

3. Patton PE, Lim JY, Hickok LR, Kettel LM, Larson JM, Pau KY (2014) Precision of progesterone measurements with the use of automated immunoassay analyzers and the impact on clinical decisions for in vitro fertilization. Fertil Steril 101:1629–1636. https://doi.org/10.1016/j.fertnstert.2014.02.037

4. Huang CC, Lien YR, Chen HF, Chen MJ, Shieh CJ, Yao YL, Chang CH, Chen SU, Yang YS (2012) The duration of pre-ovulatory serum progesterone elevation before hCG administration affects the outcome of IVF/ICSI cycles. Hum Reprod 27:2036–2045. https://doi.org/10.1093/humrep/des141

5. Ochsenkühn R, Arzberger A, von Schönfeldt V, Gallwas J, Rogenhofer N, Crispin A, Thaler CJ, Noss U (2012) Subtle progesterone rise on the day of human chorionic gonadotropin administration is associated with lower live birth rates in women undergoing assisted reproductive technology: a retrospective study with 2555 fresh embryo transfers. Fertil Steril 98:347–354. https://doi.org/10.1016/j.fertnstert.2012.04.041

6. Xu B, Li Z, Zhang H, Jin L, Li Y, Ai J, Zhu G (2012) Serum progesterone level effects on the outcome of in vitro fertilization in patients with different ovarian response: an analysis of more than 10,000 cycles. Fertil Steril 97:1321–1327. https://doi.org/10.1016/j.fertnstert.2012.03.014**(e1–4)**

7. Coucke W, Devleeschouwer N, Libeer JC, Schiettecatte J, Martin M, Smitz J (2007) Accuracy and reproducibility of automated estradiol-17beta and progesterone assays using native serum samples: results obtained in the Belgian external assessment scheme. Hum Reprod 22:3204–3209. https://doi.org/10.1093/humrep/dem322

8. Rosner W, Hankinson SE, Sluss PM, Vesper HW, Wierman ME (2013) Challenges to the measurement of estradiol: an endocrine society position statement. J Clin Endocrinol Metab 98:1376–1387. https://doi.org/10.1210/jc.2012-3780

9. Drakopoulos P, Blockeel C, Stoop D, Camus M, de Vos M, Tournaye H, Polyzos NP (2016) Conventional ovarian stimulation and single embryo transfer for IVF/ICSI. How many oocytes do we need to maximize cumulative live birth rates after utilization of all fresh and frozen embryos? Hum Reprod 31:370–376. https://doi.org/10.1093/humrep/dev316

10. Thienpont LM, Verhaeghe PG, Van Brussel KA, De Leenheer AP (1988) Estradiol-17 beta quantified in serum by isotope dilution-gas chromatography-mass spectrometry: reversed-phase C18 high-performance liquid chromatography compared with immuno-affinity chromatography for sample pretreatment. Clin Chem 34:2066–2069

11. Taieb J, Mendez Lozano DH, Benattar C, Messaoudi C, Poüs C (2007) Enlightenment about the new Architect-i2000 estradiol (Abbott Laboratories) immunoassay during in vitro fertilization. Clin Biochem 40:1423–1426. https://doi.org/10.1016/j.clinbiochem.2007.09.009

12. Ashmita J, Vikas S, Swati G (2017) The impact of progesterone level on day of hCG injection in IVF cycles on clinical pregnancy rate. J Hum Reprod Sci 10:265–270. https://doi.org/10.4103/0974-1208.223278

13. Hill MJ, Healy MW, Richter KS, Parikh T, Devine K, DeCherney AH, Levy M, Widra E, Patounakis G (2018) Defining thresholds for abnormal premature progesterone levels during ovarian stimulation for assisted reproduction technologies. Fertil Steril 110:671–679.e2. https://doi.org/10.1016/j.fertnstert.2018.05.007

14. Devroey P, Polyzos NP, Blockeel C (2011) An OHSS-free clinic by segmentation of IVF treatment. Hum Reprod 26:2593–2597. https://doi.org/10.1093/humrep/der251

15. Healy MW, Yamasaki M, Patounakis G, Richter KS, Devine K, DeCherney AH, Hill MJ (2017) The slow growing embryo and premature progesterone elevation: compounding factors for embryo-endometrial asynchrony. Hum Reprod 32:362–367. https://doi.org/10.1093/humrep/dew296

16. Mittal S, Gupta P, Malhotra N, Singh N (2014) Serum estradiol as a predictor of success of in vitro fertilization. J Obstet Gynaecol India 64:124–129. https://doi.org/10.1007/s13224-013-0470-7

17. Silverberg KM, Burns WN, Olive DL, Riehl RM, Schenken RS (1991) Serum progesterone levels predict success of in vitro fertilization/embryo transfer in patients stimulated with leuprolide acetate and human menopausal gonadotropins. J Clin Endocrinol Metab 73:797–803. https://doi.org/10.1210/jcem-73-4-797

18. Fanchin R, de Ziegler D, Taieb J, Hazout A, Frydman R (1993) Premature elevation of plasma progesterone alters pregnancy rates of in vitro fertilization and embryo transfer. Fertil Steril 59:1090–1094. https://doi.org/10.1016/s0015-0282(16)55933-0

19. Bosch E, Valencia I, Escudero E, Crespo J, Simón C, Remohí J, Pellicer A (2003) Premature luteinization during gonadotropin-releasing hormone antagonist cycles and its relationship with in vitro fertilization outcome. Fertil Steril 80:1444–1449. https://doi.org/10.1016/j.fertnstert.2003.07.002

20. Papanikolaou EG, Pados G, Grimbizis G, Bili E, Kyriazi L, Polyzos NP, Humaidan P, Tournaye H, Tarlatzis B (2012) GnRH-agonist versus GnRH-antagonist IVF cycles: is the reproductive outcome affected by the incidence of progesterone elevation on the day of HCG triggering? A randomized prospective study. Hum Reprod 27:1822–1828. https://doi.org/10.1093/humrep/des066

21. Check JH, Amui J, Choe JK, Brasile D (2009) Relationship of serum progesterone (P) level the day after human chorionic gonadotropin (hCG) injection on outcome following in vitro fertilization-embryo transfer (IVF-ET). Clin Exp Obstet Gynecol 36:214–215

22. Huang R, Fang C, Xu S, Yi Y, Liang X (2012) Premature progesterone rise negatively correlated with live birth rate in IVF cycles with GnRH agonist: an analysis of 2,566 cycles. Fertil Steril 98:664–670.e2. https://doi.org/10.1016/j.fertnstert.2012.05.024

23. Schoolcraft W, Sinton E, Schlenker T, Huynh D, Hamilton F, Meldrum DR (1991) Lower pregnancy rate with premature luteinization during pituitary suppression with leuprolide acetate. Fertil Steril 55:563–566. https://doi.org/10.1016/S0015-0282(16)54186-7

24. Ubaldi F, Smitz J, Wisanto A, Joris H, Schiettecatte J, Derde MP, Borkham E, Van Steirteghem A, Devroey P (1995) Oocyte and embryo quality as well as pregnancy rate in intracytoplasmic sperm injection are not affected by high follicular phase serum progesterone. Hum Reprod 10:3091–3096. https://doi.org/10.1093/oxfordjournals.humrep.a135864

25. Urman B, Alatas C, Aksoy S, Mercan R, Isiklar A, Balaban B (1999) Elevated serum progesterone level on the day of human chorionic gonadotropin administration does not adversely affect implantation rates after intracytoplasmic sperm injection and embryo transfer. Fertil Steril 72:975–979. https://doi.org/10.1016/s0015-0282(99)00421-5

26. Venetis CA, Kolibianakis EM, Papanikolaou E, Bontis J, Devroey P, Tarlatzis BC (2007) Is progesterone elevation on the day of human chorionic gonadotropin administration associated with the probability of pregnancy in in vitro fertilization? A systematic review and meta-analysis. Hum Reprod Update 13:343–355. https://doi.org/10.1093/humupd/dmm007

27. Kolibianakis EM, Venetis CA, Bontis J, Tarlatzis BC (2012) Significantly lower pregnancy rates in the presence of progesterone elevation in patients treated with GnRH antagonists and gonadotropins: a systematic review and meta-analysis. Curr Pharm Biotechnol 13:464–470. https://doi.org/10.2174/138920112799361927

28. Wu AHB (2006) Tietz clinical guide to laboratory tests, 4th edn. WB Saunders Co, St Louis, p 894