Article

# Live birth rates in various subgroups of poor ovarian responders fulfilling the Bologna criteria

## Gurkan Bozdag [a], Mehtap Polat [b], Irem Yarali [b], Hakan Yarali [a,b,*]

[a] *Department of Obstetrics and Gynecology, School of Medicine, Hacettepe University, Ankara, Turkey*
[b] *Anatolia Women's Health and IVF Center, Ankara, Turkey*

Hakan Yarali, MD, is Professor at Hacettepe University, Ankara, Turkey, in the Department of Obstetrics and Gynecology, and is also founder of the Anatolia IVF Center. He completed his clinical fellowship at the University of British Columbia, in the Division of Reproductive Endocrinology & Infertility. His special interests are polycystic ovary syndrome, pre-implantation genetic screening and clinical aspects of assisted reproductive technologies.

> **KEY MESSAGE**
> Patients fulfilling Bologna criteria might not be homogenous with regard to number of oocytes harvested and live birth rate. Younger poor ovarian responders may have a better outcome if they reach the embryo transfer stage. Each additional oocyte retrieved lowers the risk of cycle cancellation and enhances live birth rate.

A B S T R A C T

The European Society of Human Reproduction and Embryology published Bologna criteria to generate a definition of poor ovarian responders (PORs). However, there are few data on whether PORs are homogenous for ovarian response or live birth rates (LBRs). In this retrospective study, 821 patients fulfilling Bologna criteria and undergoing intracytoplasmic sperm injection were stratified into four groups: Group A: female age ≥40 with a previous poor response (cycle cancelled or ≤3 oocytes) (105 patients, 123 cycles); Group B: female age ≥40 with an antral follicle count (AFC) < 7 (159 patients, 253 cycles); Group C: AFC <7 with a previous poor response (350 patients, 575 cycles); and Group D: female age ≥40 with an AFC <7 and previous poor response (207 patients, 306 cycles). Cluster data analysis was performed. Although median number of oocytes was higher in Group B ($P < 0.001$), higher implantation ($P = 0.024$) and LBR per embryo transfer ($P < 0.001$) or cycle ($P = 0.001$) were noted in Group C. We conclude that, once a patient fulfils Bologna criteria, prognosis is poor, with fewer than 10% recorded LBRs per cycle. However, the LBRs are not homogenous and 'young proven' PORs have the most favourable pregnancy outcome.

© 2017 Reproductive Healthcare Ltd. Published by Elsevier Ltd. All rights reserved.

## Introduction

Poor ovarian response is encountered in 9–25% of patients undergoing IVF (Venetis et al., 2010). There had been great diversity in the definition of poor ovarian responders (PORs) until the introduction of the European Society of Human Reproduction and Embryology (ESHRE) consensus, known as the Bologna criteria (Ferraretti et al., 2011). According to the ESHRE Bologna criteria (Ferraretti et al., 2011), at least two of the following three features must be fulfilled for POR classification: (i) advanced maternal age (≥40 years) or any other risk factor, (ii) a previous poor ovarian response (cycles cancelled or ≤3 oocytes with a conventional protocol), (iii) an abnormal ovarian reserve test (ORT) [with a maximum bilateral antral follicle count (AFC) of between 5 and 7 or anti-Müllerian hormone level of 0.5–1.3 ng/ml]. In the absence of advanced maternal age or abnormal ORT, two previous episodes of poor ovarian response after maximal stimulation are sufficient to define a patient as a POR (Ferraretti et al., 2011). Although the Bologna criteria have provided a useful criteria set for the definition of PORs, there has been criticism of the criteria regarding the lack of definition of risk factors (Younis et al., 2015), the threshold points chosen (Sallam et al., 2012) and, most importantly, the lack of homogeneity of pregnancy outcomes of various subgroups fulfilling the Bologna criteria (Papathanasiou, 2014).

There is a paucity of data on the live birth rates (LBRs) of various subgroups of PORs fulfilling the Bologna criteria and undergoing IVF (Papathanasiou, 2014). To our knowledge, there are only two studies reporting the IVF performance of various subgroups of PORs fulfilling the Bologna criteria, both reporting similar LBRs (Busnelli et al., 2015; La Marca et al., 2015). Obviously, these results should be validated in larger sample size studies. The main goal of this study was to evaluate whether various subgroups of PORs fulfilling the Bologna criteria have comparable prognoses regarding cycle cancellation and LBRs at intracytoplasmic sperm injection (ICSI) and embryo transfer cycles.

## Materials and methods

The database containing detailed clinical and laboratory information on all ICSI treatment cycles performed at the Anatolia IVF and Women's Health Center during the period between August 2005 and August 2014 (n = 14,709 cycles) was analysed. In the current retrospective cohort study, all the data per cycle were entered into the database prospectively. All the patient files, as well as cycle characteristics, were scrutinized manually and those cycles fulfilling the Bologna criteria were identified. The null hypothesis was that different subgroups of PORs fulfilling Bologna criteria have similar LBRs.

The exclusion criteria included: azoospermia necessitating surgical retrieval of spermatozoa (n = 1095), structural or numerical chromosomal errors necessitating pre-implantation genetic diagnosis or screening (n = 121) and frozen embryo transfer cycles (n = 373). Since the ovarian reserve testing, as well as starting dose of FSH and number of oocytes harvested in those ICSI cycles performed elsewhere could not be precisely validated based on the couple's medical history, only those first and subsequent cycles that were performed at our centre were included. In other words, those couples with a history of a prior IVF/ICSI attempt(s) elsewhere were excluded from the current analysis.

Because AFC performed in the early follicular phase is the primary tool for the assessment of the ORT at our clinic, no anti-Müllerian hormone (AMH) data were included in the current analysis. All women underwent AFC assessment at the second or third day of menses immediately before starting ovarian stimulation. Every round or oval structure within the margin of 2 to 10 mm was considered to be an antral follicle. The threshold for normalcy for AFC was taken as 7.

Women underwent IVF/ICSI cycles using microdose flare-up (Lucrin, Abbott, Istanbul, Turkey) or multi-dose flexible GnRH antagonist (Cetrotide, Merck, Istanbul, Turkey) protocol, based on the physician's preference and in the manner described elsewhere (Yarali et al., 2009). All women with expected or proven poor ovarian response underwent ovarian stimulation with a starting gonadotrophin dose of ≥300 IU/day (300–450 IU/day). Ovarian stimulation was performed with the use of recombinant FSH (Gonal-F, Merck) and/or HP-HMG (Menopur, Ferring, Istanbul, Turkey), based on the physician's preference. Ovarian response was monitored with frequent serum oestradiol measurements and transvaginal ultrasounds. The criterion for HCG (Ovitrelle, Merck or Pregnyl, MSD, Istanbul, Turkey) administration was the presence of at least one follicle exceeding 17 mm in diameter.

Oocyte retrieval was carried out under general anaesthesia using vaginal ultrasound-guided puncture of follicles 34 to 36 h after HCG administration. Standard procedures were followed for gamete–embryo handling and ICSI. Embryo transfer was performed in all cases using a soft catheter under ultrasound guidance. Daily vaginal progesterone gel (Crinone, Merck) was administered for luteal phase support, starting 1 day after oocyte retrieval and continued until fetal cardiac activity was confirmed.

Four distinct subgroups of patients fulfilling the Bologna criteria were generated: Group A: female age ≥40 with a previous poor ovarian response (cycle cancelled or ≤3 oocytes) (105 patients; 123 cycles); Group B: female age ≥40 with AFC <7 (159 patients; 253 cycles); Group C: AFC <7 with a previous poor ovarian response (350 patients; 575 cycles); and Group D: female age ≥40 with AFC <7 and a previous poor ovarian response (207 patients; 306 cycles). All comparisons between these four groups were made on a per-cycle basis.

Clinical pregnancy was defined as visualization of a gestational sac with fetal cardiac activity at ultrasonography and/or confirmation of chorionic villi at the pathology specimen. Live birth was defined as a birth of a live baby exceeding 24 weeks of gestation. Pregnancy rates were given as per cycle commenced and per embryo transfer attempt. We calculated the implantation rate for a given patient (individual implantation rate) by dividing the number of sacs with fetal cardiac activity by the number of embryos transferred, as reported (Ben-Shlomo et al., 1997). We then summed the individual implantation rates and divided by the number of embryo transfer attempts in each group.

Normally distributed parametric variables confirmed by the Kolmogorov–Smirnov and Shapiro–Wilk tests were compared by analysis of variance (ANOVA) with the Bonferroni method for post hoc analysis by using Statistics Package for Social Sciences (ver. 21.0; SPSS Inc., Chicago). Non-normally distributed metric variables were analysed by the Kruskal–Wallis and Mann–Whitney U-tests. The chi-squared test was used to analyse nominal variables in the form of frequency tables. Binary logistic regression analysis with the forward conditional method was used to delineate the independent variable(s) for live birth. $P < 0.05$ was considered statistically significant. Values were expressed as medians (minimum–maximum), unless otherwise stated.

In order to account for inclusion of multiple cycles of a patient, we performed cluster analysis. For this purpose, we performed multilevel linear regression analysis for the continuous embryological data (number of oocyte–cumulus complexes, metaphase-II oocytes and number of embryos transferred) and multilevel logistic regression analysis for the categorical data (fertilization rate, embryos <10% fragmentation with ≥7 blastomeres, number of cancelled cycles, number of embryos transferred, clinical pregnancy/cycle, clinical pregnancy/embryo transfer, live birth/cycle, live birth/embryo transfer and implantation rate). Due to their multi-nominal structure, Bologna criteria were considered as a dummy variable.

Ethical approval was obtained from Hacettepe University with the file number of GO 16/89 on 14 June 2016.

## Results

In our database, 821 patients fulfilling the Bologna criteria who underwent 1257 ICSI cycles were enrolled. The demographic features of the four subgroups of patients are outlined in **Table 1**. As might be expected, the median female age was significantly younger in Group C, in which advanced female age was not a criterion, compared with the other three groups ($P < 0.001$). There were slight but statistically significant differences regarding AFC among Groups B, C and D; the median AFC was highest in Group A when compared with Groups B, C and D ($P < 0.001$). Total FSH consumption was significantly higher in Group C when compared with Groups A and B ($P < 0.001$).

With ANOVA (without cluster data analysis), the total number of oocytes and metaphase-II oocytes was significantly higher in Group B when compared with Groups A, C and D (**Table 2**, $P < 0.001$). However, the fertilization rates and rates of top quality embryos were comparable among the four groups. The median number of embryos transferred was two in Group B, significantly higher than in Groups A, C and D (**Table 2**, $P < 0.001$). The rate of cycle cancellation before embryo transfer was significantly lower in Group B as compared with Groups C and D (**Table 2**, $P < 0.001$). There was a significantly higher implantation rate in Group C when compared with Group B ($P = 0.024$). Live birth rate per cycle ($P = 0.001$) or transfer ($P < 0.001$) in Group C was significantly higher than Group D.

The results obtained by ANOVA and cluster data analyses were concordant for all the studied variables, with the exception of fertilization rate only.

When univariate analysis was performed, the variables that were significantly different among cycles that did or did not end up with live birth were female age, AFC, duration of stimulation (days), total dose of FSH used (IU) and metaphase-II oocytes (data not shown). However, body mass index, duration of infertility oestradiol level on the day of HCG, number of embryos <10% fragmentation with ≥7 blastomeres and number of embryos transferred were not noted to be significant at univariate analysis. When the logistic regression analysis was set with live birth as the dependent variable, and the significant variables at the univariate analysis as independent variables, female age ($P < 0.001$; OR = 0.89, 95% CI: 0.84–0.92), total dose of FSH used ($P = 0.001$; OR = 0.98, 95% CI: 0.96–0.99) and metaphase-II oocytes harvested ($P < 0.001$; OR = 1.48, 95% CI: 1.26–1.74) remained significant.

By stratifying the number of oocytes collected, all patients were assessed together regarding the outcome (**Table 3**). As might be expected, the cycle cancellation rate declined and the LBR increased steadily with each additional oocyte yield (**Table 3**). When one harvested oocyte is taken as the reference, the risk ratios for cycle cancellation and live birth are also given in **Table 3**. Of note, the implantation rate was constant irrespective of the number of oocytes retrieved (**Table 3**).

## Discussion

The Bologna criteria, although not perfect, have been a major advance in the definition of PORs. With the Bologna criteria, considering that two out of three criteria suffice to make the diagnosis, 13 subgroups of patients may be generated (**Table 4**) (Ferraretti and Gianaroli, 2014; Papathanasiou, 2014). There is a paucity of data as to whether these subgroups of patients have comparable IVF/ICSI outcomes. In our study, we noted that the four subgroups of PORs we selected for investigation are not homogenous for ovarian stimulation performance and LBRs. Of the four studied subgroups, the subgroup presenting AFC < 7 and a previous poor ovarian response (Group C)

| Table 1 – The demographic features of four subgroups of poor ovarian responders fulfilling the Bologna criteria. | | | | | |
|---|---|---|---|---|---|
| Variable | Group A (Age ≥40 + previous poor ovarian response) | Group B (Age ≥40 + AFC <7) | Group C (Previous poor ovarian response + AFC <7) | Group D (Age ≥40 + AFC <7 + previous poor ovarian response) | P-value |
| No. of patients | 105 | 159 | 350 | 207 | |
| No. of cycles | 123 | 253 | 575 | 306 | |
| Female age (years) | 42 (40–47) | 42 (40–46) | 35 (21–39)[a] | 42 (40–46) | <0.001 |
| AFC | 8 (7–11)[b] | 4 (0–6)[c] | 4 (0–6)[a] | 3 (0–6)[d] | <0.001 |
| Body mass index (kg/m²) | 26.5 (18.3–39.7) | 25.8 (19.4–44.2) | 25.0 (16.2–48.3)[e] | 25.6 (17.1–44.2) | 0.001 |
| Duration of infertility (months) | 51 (2–504) | 48 (2–300) | 48 (1–288) | 36 (1–420)[e] | 0.002 |
| Duration of stimulation (days) | 9 (5–17) | 10 (4–18) | 10 (4–20)[e] | 10 (4–16) | <0.001 |
| Total dose of FSH used (IU) | 4050 (1800–7200) | 4050 (600–8400) | 4500 (1500–11,700)[e] | 4050 (1800–9300) | <0.001 |
| Oestradiol level on the day of HCG (pg/ml) | 536.0 (71–3294) | 978.0 (9–4844)[c] | 555.5 (33–3042)[f] | 443.0 (21–4871) | <0.001 |

Values are expressed as median (minimum–maximum).
For post hoc analysis, $P < 0.008$ was defined as statistical significance.
AFC = antral follicle count; HCG = human chorionic gonadotrophin.
[a,b,c,d] Statistically different from all other groups.
[e] Statistically different from Groups A and B.
[f] Statistically different from Groups B and D.

**Table 2 – The embryologic data and pregnancy rates of four subgroups of poor ovarian responders fulfilling the Bologna criteria. For each variable in a row, above-given values refer to median (minimum–maximum) and below-given values to coefficient (95% CI) as estimated by multilevel logistic regression analysis.**

| Variable | Group A (Age ≥40 + previous poor ovarian response) | Group B (Age ≥40 + AFC <7) | Group C (Previous poor ovarian response + AFC <7) | Group D (Age ≥40 + AFC <7 + previous poor ovarian response) | P-value |
|---|---|---|---|---|---|
| No. of oocyte–cumulus complexes | 2 (0–3)[a] | 5 (0–8)[b] | 2 (0–3) | 2 (0–3) | <0.001 |
| | −0.44 (−0.74; −0.14)[1] | 3.28 (3.12; 3.44)[2] | −1.03 (−1.21; −0.84)[3] | −0.96 (−1.16; −0.75)[4] | 0.004[1]/<0.001[2,3,4] |
| No. of metaphase-II oocytes | 2 (0–3) | 4 (0–8)[b] | 1 (0–3) | (0–3) | <0.001 |
| | −0.27 (−0.54; 0.003)[1] | 2.35 (2.19; 2.52)[2] | −0.83 (−0.99; −0.67)[3] | −0.66 (−0.84; −0.47)[4] | 0.047[1]/<0.001[2,3,4] |
| Fertilization rate (%) | 68.3 | 63.8 | 66.8 | 63.2 | 0.391 |
| | −0.005 (−0.18; 0.17) | 0.64 (0.52; 0.77)[1] | −0.30 (−0.40; −0.20)[2] | −0.17 (−0.29; −0.05)[3] | < 0.001[1,2]/0.007[3] |
| Embryos <10% fragmentation with ≥7 blastomeres (%) | 40.2 | 45.9 | 46.3 | 42.2 | 0.550 |
| | −0.07 (−0.18; 0.03) | 0.04 (−0.04; 0.11) | 0.02 (−0.05; 0.09) | −0.02 (−0.10; 0.07) | |
| No. of cancelled cycles (%) | 46 (37.4) | 82 (32.4)[c] | 270 (47.0) | 153 (50.0) | <0.001 |
| | −0.35 (−0.82; 0.13) | −0.71 (−1.07; −0.35)[1] | 0.28 (−0.003; 0.56)[2] | 0.40 (0.07; 0.73)[3] | <0.001[1]/0.052[2]/0.016[3] |
| No. of embryos transferred | 1 (1–3) | 2 (1–5)[b] | 1 (1–3) | 1 (1–3) | <0.001 |
| | −0.12 (−0.30; 0.07) | 0.70 (0.57; 0.82)[1] | −0.33 (−0.44; −0.21)[2] | −0.23 (−0.37; −0.09)[3] | <0.001[1,2]/0.001[3] |
| Clinical pregnancy/cycle (n, %) | 6 (4.9) | 22 (8.7) | 66 (11.5)[d] | 14 (4.6) | 0.002 |
| | −0.83 (−1.89; 0.22) | 0.07 (−0.54; 0.68) | 0.84 (0.26; 1.42)[1] | −1.02 (−1.80; −0.23)[2] | 0.004[1]/0.011[2] |
| Clinical pregnancy/embryo transfer (n, %) | 6 (7.8) | 22 (12.9) | 66 (21.6)[e] | 14 (9.2) | <0.001 |
| | −0.88 (−1.80; 0.05)[1] | −0.27 (−0.82; 0.27) | 0.89 (0.41; 1.37)[2] | −0.76 (−1.41; −0.10)[3] | 0.063[1]/<0.001[2]/0.024[3] |
| Live birth/cycle (n, %) | 4 (3.3) | 16 (6.3) | 50 (8.7)[d] | 7 (2.3) | 0.001 |
| | −0.72 (−1.74; 0.31) | 0.04 (−0.53; 0.61) | 0.84 (0.36; 1.32)[1] | −1.22 (−2.01; −0.43)[2] | 0.001[1]/0.002[2] |
| Live birth/embryo transfer (n, %) | 4 (5.2) | 16 (9.4) | 50 (16.4)[d] | 7 (4.6) | <0.001 |
| | −0.87 (−1.91; 0.16) | −0.22 (−0.80; 0.36) | 1.00 (0.50; 1.49)[1] | −1.11 (−1.90; −0.31)[2] | <0.001[1]/0.007[2] |
| Implantation rate (%) | 6.0 | 5.8 | 14.4[f] | 6.3 | 0.024 |
| | 0.02 (−0.22; 0.27) | −0.21 (−0.36; −0.06)[1] | 0.10 (−0.02; 0.23) | 0.06 (−0.12; 0.24) | 0.006[1] |

For post hoc analysis, P < 0.008 was defined as statistically significant.
AFC = antral follicle count.
[a] Statistically different from Groups B and D.
[b] Statistically different from all other groups.
[c] Statistically different from Groups C and D.
[d] Statistically different from Group D.
[e] Statistically different from Groups A and D.
[f] Statistically different from Group B.
[1,2,3,4] Statistically different from all other groups.

was found to be associated with the most favourable live birth and implantation rates. This is not unexpected, in fact, as this subgroup is that of the 'young proven' PORs, and female age is the most important determinant of live birth (Oudendijk et al., 2012). Interestingly, the expected PORs (Group B) had the highest median number of oocytes retrieved and hence number of embryos transferred. This finding is also not unexpected, because the ORTs are not 100% ac-

curate (Broer et al., 2013). However, in this subgroup, retrieval of a higher number of oocytes did not reflect higher pregnancy outcomes and there was significantly lower implantation rate when compared with young PORs, as referred by Group C. This finding once again highlights the prognostic value of female age for final pregnancy outcome (Broer et al., 2013). Oocyte aneuploidy is responsible for diminished LBRs in assisted reproductive technology in women

**Table 3 – The cycle cancellation rate, number of embryos transferred and pregnancy outcome in poor ovarian responders fulfilling the Bologna criteria with regard to number of oocytes collected.**

| | No. of oocytes | | | | | P-value |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | |
| Cycle cancellation (%, n) | 65.5 (258/394) | 37.4 (176/471) | 21.5 (118/548) | 16.2 (98/606) | 12.7 (91/714) | <0.001 |
| RR (95% CI) for cycle cancellation | 1 | 0.31 (0.24–0.42) | 0.15 (0.11–0.19) | 0.10 (0.08–0.14) | 0.08 (0.06–0.10) | |
| No. of embryos transferred[a] | 1.0 (0.9) | 1.4 (0.5) | 1.6 (0.6) | 1.8 (0.7) | 2.0 (0.8) | <0.001 |
| Live birth/cycle (%, n) | 3.8 (15/394) | 8.5 (40/471) | 12.4 (68/549) | 16.7 (101/606) | 22.1 (158/715) | <0.001 |
| RR (95% CI) for live birth/cycle | 1 | 2.35 (1.28–4.31) | 3.57 (2.01–6.35) | 5.05 (2.89–8.83) | 7.17 (4.15–12.37) | |
| Live birth/embryo transfer (%, n) | 11.0 (15/136) | 13.6 (40/295) | 15.8 (68/431) | 19.9 (101/508) | 25.3 (158/624) | <0.001 |
| RR (95% CI) for live birth/embryo transfer | 1 | 1.27 (0.67–2.38) | 1.51 (0.83–2.74) | 2.00 (1.12–3.57) | 2.74 (1.55–4.82) | |
| Implantation rate (%) | 13.7 | 10.8 | 12.2 | 12.7 | 11.1 | NS |

NS = not statistically significant; RR = risk ratio (calculated with reference to one oocyte).
[a] Values are given as mean (SD).

**Table 4 – Phenotypes as may be generated from the Bologna criteria.**

| Feature (Bologna criteria) | Phenotypes | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Advanced maternal age (1A) | + | – | + | – | – | + | – | – | + | – | – | + | – |
| Other risk factor (1B) | – | + | – | + | – | – | + | – | – | + | – | – | + |
| One previous poor ovarian response with a conventional protocol (2) | + | + | – | – | + | + | + | – | – | – | + | + | + |
| Abnormal ovarian reserve test (3) | – | – | + | + | + | + | + | – | – | – | – | – | – |
| Two previous episodes of poor ovarian response after maximal stimulation (4) | – | – | – | – | – | – | – | + | + | + | + | + | + |
| Defined groups in the current study | A | B | | C | D | | | | | | | | |

of advanced maternal age (Hassold and Hunt, 2001). According to a review of 15,169 consecutive trophectoderm biopsies evaluated with comprehensive chromosomal screening, the rate of aneuploidy increases from 34.5% to 75.1% in women aged 35 and 42, respectively (Franasiak et al., 2014).

It is essential that a classification scheme, being simple, clearly defined and reproducible, should define a homogenous group of patients with similar prognoses for treatment outcomes. Such a universal definition would permit the conducting of clinical research and development of evidence-based efficient protocols or modalities for treatment in these patients. However, to our knowledge, there are only two retrospective studies on the IVF/ICSI performance of the various subgroups of patients fulfilling the Bologna criteria (Busnelli et al., 2015; La Marca et al., 2015). In a retrospective study, Busnelli et al. recently reported the IVF/ICSI outcome of 362 PORs fulfilling the Bologna criteria and enrolled from two clinics in Italy (Busnelli et al., 2015). The main outcome measure was LBR per started cycle and five subgroups were generated which included (i) anamnestic risk factors and one previous poor ovarian response; (ii) anamnestic risk factors and an abnormal ORT; (iii) an abnormal ORT and one previous poor ovarian response; (iv) anamnestic risk factors, an abnormal ORT and one previous poor ovarian response; (v) two episodes of poor ovarian response after maximal stimulation. Overall, the LBR was 6% and did not differ significantly between the five patient subgroups. Age, serum FSH and AFC were not significantly associated with live birth.

La Marca et al. (2015), with a similar retrospective study design, reported the IVF/ICSI performance of five subgroups of PORs. For a total of 210 women with poor ovarian response fulfilling the Bologna criteria, five subgroups were formed. These included: (i) female age >40 years and a previous cycle with <4 oocytes; (ii) female age >40 years and abnormal ORT; (iii) abnormal ORT and a previous cycle with <4 oocytes; (iv) female age >40 years, a previous cycle with <4 oocytes and abnormal ORT; (v) two cycles with <4 oocytes. The rates of women reaching the oocyte pick-up and embryo transfer phases were comparable among the five subgroups. The clinical outcomes and LBRs were also comparable among the five subgroups. The LBR ranged from 5.5% to 7.4% among the five subgroups. The authors concluded that patients with a diagnosis of POR according to the Bologna criteria had uniformly poor prognosis and the various subgroups represented a homogenous population with similar clinical outcomes (La Marca et al., 2015).

There may be several reasons for the discordancy between our results and those reported by these previous two studies (Busnelli et al., 2015; La Marca et al., 2015). First, limited sample size renders a study more prone to type II error. The sample size of our study was 1257, 362 in the Busnelli et al. (2015) study, and 210 in the La Marca et al. (2015) study. Second, there might be differences in patient selection criteria between these studies, which might affect pregnancy outcome. In the original Bologna criteria, the definition of 'any other

risk factor' for poor ovarian response is not clearly defined and has been a source of criticism (Ferraretti et al., 2011). In the Busnelli et al. (2015) study, 'anamnestic risk factor' for poor response was considered an inclusion criterion whereas in the study by La Marca et al. (2015), patients with risk factors (previous ovarian surgery, presence of ovarian cysts, history of pelvic inflammatory disease) were excluded. In our study, we also excluded patients with such risk factors. Obviously, differences in inclusion criteria may contribute to discrepancies in results between different studies. Third, the differences in the usage of ovarian reserve biomarkers, the assays used to measure AMH and their thresholds for normalcy may account for divergent results. In the study conducted by Busnelli et al. (2015), the thresholds for AFC or AMH were set to 5 and 0.5 ng/ml, respectively. In the study by La Marca et al. (2015), these respective figures were 7 and 1 ng/ml. In our study, AMH was not used and the threshold for AFC was set to 7.

In general, PORs fulfilling the Bologna criteria return a discouraging proportion of LBRs per started cycle of less than 10% (Busnelli et al., 2015; Kedem et al., 2014; La Marca et al., 2015; Polyzos et al., 2012, 2014). In line with previous studies, the LBRs per started cycle ranged from 2.3% to 8.7% in our study. The number of oocytes retrieved is an important surrogate marker for live birth in this patient population. The LBRs per started cycle were 3.8%, 8.5%, 12.4%, 16.7% and 22.1% with one, two, three, four and five oocytes retrieved, respectively. However, once the patient reached embryo transfer, the implantation rates were comparable regardless of the number of oocytes retrieved. In our database, when logistic regression analysis was performed to predict the independent variables for live birth, female age, total FSH dose consumption and metaphase-II oocytes remained significant in the model, as expected.

Exclusion of those patients who had prior IVF/ICSI attempts elsewhere may be a limitation of the current study, potentially introducing selection bias. Although the database was generated in a prospective fashion, the retrospective nature of the study and retrospective assignment of those patients to fulfil the Bologna criteria before 2011 (Ferraretti et al., 2011) may be another limitation of the current study. Inclusion of multiple cycles of a patient may also be a limitation of the current series. However, cluster data analysis was performed to account for this and to cluster non-independent data. Interestingly, all the comparisons with and without cluster data analysis were concordant except fertilization rate, although we have no biologically plausible explanation for this finding. Further studies with larger sample sizes are warranted to establish whether there is a difference in prognosis of different subgroups of poor ovarian responder patients fulfilling Bologna criteria.

We conclude that, once a patient fulfils the diagnosis of poor ovarian response according to the Bologna criteria, the prognosis is poor, with LBRs per started cycle of less than 10%. However, the LBR is not homogenous among this patient population and 'young proven' PORs

have the most favourable pregnancy outcomes. The different re-ported outcomes of subgroups of PORs may be helpful when counselling those patients and may act as a potential guide to treat-ment plans.

REFERENCES

Ben-Shlomo, I., Wiener-Megnagi, Z., Golan, J., Eyali, V., Geslevich, J., Shalev, E., 1997. Individual implantation rate: proposal for a new index for evaluation of assisted reproduction results. Fertil. Steril. 68, 816–819.

Broer, S.L., van Disseldorp, J., Broeze, K.A., Dolleman, M., Opmeer, B.C., Bossuyt, P., Eijkemans, M.J., Mol, B.W., Broekmans, F.J., IMPORT Study Group, 2013. Added value of ovarian reserve testing on patient characteristics in the prediction of ovarian response and ongoing pregnancy: an individual patient data approach. Hum. Reprod. Update 19, 26–36.

Busnelli, A., Papaleo, E., Del Prato, D., La Vecchia, I., Iachini, E., Paffoni, A., Candiani, M., Somigliana, E., 2015. A retrospective evaluation of prognosis and cost-effectiveness of IVF in poor responders according to the Bologna criteria. Hum. Reprod. 30, 315–322.

Ferraretti, A.P., Gianaroli, L., 2014. The Bologna criteria for the definition of poor ovarian responders: is there a need for revision? Hum. Reprod. 29, 1842–1845.

Ferraretti, A.P., La Marca, A., Fauser, B.C., Tarlatzis, B., Nargund, G., Gianaroli, L., ESHRE Working Group on Poor Ovarian Response Definition, 2011. ESHRE consensus on the definition of 'poor response' to ovarian stimulation for in vitro fertilization: the Bologna criteria. Hum. Reprod. 26, 1616–1624.

Franasiak, J.M., Forman, E.J., Hong, K.H., Werner, M.D., Upham, K.M., Treff, N.R., Scott, R.T., Jr., 2014. The nature of aneuploidy with increasing age of the female partner: a review of 15,169 consecutive trophectoderm biopsies evaluated with comprehensive chromosomal screening. Fertil. Steril. 101, 656–663, e1.

Hassold, T., Hunt, P., 2001. To err (meiotically) is human: the genesis of human aneuploidy. Nat. Rev. Genet. 2, 280–291.

Kedem, A., Tsur, A., Haas, J., Yerushalmi, G.M., Hourvitz, A., Machtinger, R., Orvieto, R., 2014. Is the modified natural in vitro fertilization cycle justified in patients with 'genuine' poor response to controlled ovarian hyperstimulation? Fertil. Steril. 101, 1624–1628.

La Marca, A., Grisendi, V., Giulini, S., Sighinolfi, G., Tirelli, A., Argento, C., Re, C., Tagliasacchi, D., Marsella, T., Sunkara, S.K., 2015. Live birth rates in the different combinations of the Bologna criteria poor ovarian responders: a validation study. J. Assist. Reprod. Genet. 32, 931–937.

Oudendijk, J.F., Yarde, F., Eijkemans, M.J., Broekmans, F.J., Broer, S.L., 2012. The poor responder in IVF: is the prognosis always poor?: a systematic review. Hum. Reprod. Update 18, 1–11.

Papathanasiou, A., 2014. Implementing the ESHRE 'poor responder' criteria in research studies: methodological implications. Hum. Reprod. 29, 1835–1838.

Polyzos, N.P., Blockeel, C., Verpoest, W., De Vos, M., Stoop, D., Vloeberghs, V., Camus, M., Devroey, P., Tournaye, H., 2012. Live birth rates following natural cycle IVF in women with poor ovarian response according to the Bologna criteria. Hum. Reprod. 27, 3481–3486.

Polyzos, N.P., Nwoye, M., Corona, R., Blockeel, C., Stoop, D., Haentjens, P., Camus, M., Tournaye, H., 2014. Live birth rates in Bologna poor responders treated with ovarian stimulation for IVF/ICSI. Reprod. Biomed. Online 28, 469–474.

Sallam, H.N., Ezzeldin, F., Agameya, A.F., Abdel-Rahman, A.F., El-Garem, Y., 2012. The definition of 'poor response': Bologna criteria. Hum. Reprod. 27, 626–627, author reply 627–8.

Venetis, C.A., Kolibianakis, E.M., Tarlatzi, T.B., Tarlatzis, B.C., 2010. Evidence-based management of poor ovarian response. Ann. N. Y. Acad. Sci. 1205, 199–206.

Yarali, H., Esinler, I., Polat, M., Bozdag, G., Tiras, B., 2009. Antagonist/letrozole protocol in poor ovarian responders for intracytoplasmic sperm injection: a comparative study with the microdose flare-up protocol. Fertil. Steril. 92, 231–235.

Younis, J.S., Ben-Ami, M., Ben-Shlomo, I., 2015. The Bologna criteria for poor ovarian response: a contemporary critical appraisal. J. Ovarian Res. 8, 76.