

# The Vienna consensus: report of an expert meeting on the development of art laboratory performance indicators<sup>†‡</sup>

## ESHRE Special Interest Group of Embryology<sup>1,\*</sup> and Alpha Scientists in Reproductive Medicine<sup>2,\*</sup>

<sup>1</sup>European Society of Human Reproduction and Embryology, Meerstraat 60, B-1852 Grimbergen, Belgium <sup>2</sup>ALPHA Scientists in Reproductive Medicine, 19 Mayıs Mah. 19 Mayıs Cad. Nova Baran Center No:4 34360 Sisli, Istanbul, Turkey

\*Correspondence address. E-mail: coticchio.biogenesi@grupposandonato.it (G.C.)/zsolt.peter.nagy@gmail.com (Z.P.)

Submitted on June 15, 2017; resubmitted on June 15, 2017; editorial decision on July 13, 2017; accepted on August 3, 2017

**STUDY QUESTION:** What are appropriate performance indicators (PIs) for ART laboratories for use in monitoring 'fresh' IVF and ICSI cycles?

**SUMMARY ANSWER:** Minimum performance (competence) levels and aspirational (benchmark) values were recommended for a total of 19 indicators, including 12 key PIs (KPIs), five PIs and two reference indicators (RIs).

**WHAT IS ALREADY KNOWN:** PIs are necessary for systematic monitoring of the laboratory and an important element within the Quality Management System. However, there are no established PIs for ART laboratories and there is very little evidence on the topic.

**STUDY DESIGN, SIZE, DURATION:** This is the report of a 2-day consensus meeting of expert professionals. As a starting point for the discussion, two surveys were organized to collect information on indicators used in IVF laboratories. During the meeting, the results of the surveys, scientific evidence (where available), and personal clinical experience were integrated into presentations by experts on specific topics. After presentation, each proposed indicator was discussed until consensus was reached within the panel.

**PARTICIPANTS/MATERIALS, SETTING, METHODS:** Expert professionals representing different countries and settings convened in the consensus meeting.

**MAIN RESULTS AND THE ROLE OF CHANCE:** The paper is divided in two parts: the workshop report and the recommendations of the expert panel. The second part reflects the discussion on each of the indicators, with the agreed definition, competence level and benchmark value for each of the 19 indicators, including 12 KPIs, 5 PIs and 2 RIs.

**LIMITATIONS, REASONS FOR CAUTION:** The KPIs are mainly based on expert opinion. Future research may warrant an update of the recommended KPIs, their definition and the competence level and benchmark values.

**WIDER IMPLICATIONS OF THE FINDINGS:** Based on the information presented, each ART laboratory should select its own set of KPIs founded on laboratory organization, and processes.

**STUDY FUNDING/COMPETING INTEREST(S):** The consensus meeting and writing of the paper was supported by funds from ESHRE and Alpha. Alpha gratefully acknowledges the following organizations for their financial support, through the provision of unrestricted educational grants: Global Fertility Alliance, Merck, Origio and Vitrolife. There are no conflicts of interest to disclose.

**Key words:** embryo / IVF / KPI / oocyte / performance indicator / sperm

<sup>†</sup>ESHRE Pages content is not externally peer reviewed. This manuscript has been approved by the Executive Committee of ESHRE.

<sup>‡</sup>This article has been co-published with permission in *HROpen* and *Reproductive Biomedicine Online*.

## WHAT DOES THIS MEAN FOR PATIENTS?

Performance indicators (PI) are seen as a good way to check that healthcare is working well and within safety guidelines. There are not any existing PIs set across the board for IVF laboratories. This report looks at PIs for laboratories, and is based on a 2 day professional meeting and two surveys which aimed to find out more about how IVF laboratories are currently working. The aim was to set up key PI (KPIs) for IVF laboratories and an expert panel made some recommendations.

The report looked at a number of processes in an IVF laboratory including egg collection, how eggs are evaluated and sperm is assessed, fertilization rates after IVF and ICSI (where the sperm is injected directly into the egg), how well embryos develop, freezing rates and treatment outcomes. For each of these processes KPIs were defined.

The expert panel concluded that it was important for each laboratory to monitor and check their own performance regularly. The panel outlined good practice and suggested a list of 19 PIs for IVF laboratories. The panel also suggested that in the future staff at each laboratory should select their own indicators based on the way that they worked.

## Introduction

Performance indicators (PIs) are objective measures for evaluating critical healthcare domains (patient safety, effectiveness, equity, patient-centeredness, timeliness and efficiency) (Kohn et al., 2000). In the setting of a clinical laboratory, quality indicators are necessary for systematically monitoring and evaluating the laboratory's contribution to patient care (ISO15189-2012) and they represent an important element within the Quality Management System (QMS) (Mortimer and Mortimer, 2015; ESHRE Guideline Group on Good Practice in IVF Labs et al., 2016). Currently, there are no established PIs for ART laboratories, and there is very little published evidence on the topic.

Any PI should be reliable and robust, and routine data collection for the indicator should be straightforward. Furthermore, the biological or technical process to be monitored should be defined, and relevant qualifiers, confounders and endpoints should be identified. Key PIs (KPIs) are indicators deemed essential for evaluating the introduction of a technique or process; establishing minimum standards for proficiency; monitoring ongoing performance within a QMS (for internal quality control (IQC), external quality assurance (EQA)); benchmarking and quality improvement. In general, the results of a series of KPIs will provide an adequate overview of the most important steps in the IVF laboratory process (Salinas et al., 2010).

The aim of the consensus meeting and report was to establish KPIs for ART laboratories for the use in monitoring 'fresh' IVF and ICSI cycles and provide the basis for several of the quantitative performance criteria needed to create competency profiles for Clinical Embryologists. More specifically, the purpose was to achieve an international consensus regarding: first, a minimum list of IVF laboratory indicators and KPIs that in the future can be further extended and/or revised; second, specific definitions for these indicators (including necessary case inclusion/exclusion criteria; and calculation formulae); and third, recommended values for each KPI (minimum 'competency' limit; and 'aspirational goal' benchmark).

Based on the information presented here, each laboratory should develop its own set of KPIs founded on laboratory organization and processes, and develop a systematic, transparent and consistent approach to data collection and analysis and calculation of KPIs (Mayer et al., 2003; Salinas et al., 2010; Mortimer and Mortimer, 2015; ESHRE Guideline Group on Good Practice in IVF Labs et al., 2016).

## Methodology

This report is the result of a 2-day consensus meeting of expert professionals (participants are listed in Table I). As a starting point for the discussion at the meeting, two surveys were organized to collect information on indicators used in IVF laboratories. The first, the 'Alpha survey', was sent to national and international societies of ART laboratory directors and

**Table I Consensus workshop participants and contributors.**

Name	Affiliations
Susanna Apter	Fertilitetscentrum Stockholm, Sweden
Basak Balaban	American Hospital of Istanbul, Turkey
Alison Campbell*	CARE Fertility Group, UK
Jim Catt	Optimal IVF, Melbourne, Australia
Giovanni Coticchio	Biogenesi, Monza, Italy
Maria José de los Santos*	IVI Valencia, Valencia, Spain
Sophie Debrock*	Leuven University Fertility Centre, Leuven, Belgium
Thomas Ebner*	Kepler University, Linz, Austria
Stephen Harbottle	Cambridge IVF, UK
Ciara Hughes	Rotunda IVF, Dublin, Ireland
Ronny Janssens	Centre for Reproductive Medicine, Brussels, Belgium
Nathalie Le Clef	ESHRE Central Office, Grimbergen, Belgium
Kersti Lundin	Sahlgrenska University Hospital, Sweden
Cristina Magli*	SISMER, Bologna, Italy
David Mortimer*	Oozoa Biomedical, Vancouver, Canada
Sharon Mortimer	Oozoa Biomedical, Vancouver, Canada
Zsolt Peter Nagy	Reproductive Biology Associates, Atlanta, USA
Johan Smitz*	Centre for Reproductive Medicine, Brussels, Belgium
Arne Sunde	St Olav's University Hospital, Trondheim, Norway
Nathalie Vermeulen	ESHRE Central Office, Grimbergen, Belgium

\*Presenters at the consensus workshop.

Clinical Embryologists, and to the members of the ESHRE committee of national representatives. Eighteen responses (with opinions from Australia, Austria, Belgium, Bulgaria, Canada, Croatia, France, Germany, Ireland, Italy, Japan, Slovenia, Sweden, South Africa, Turkey, UK and USA) out of 34 sent were received, and its results informed the expert panel on minimum expected, or competence, values (i.e. values that any laboratory should be able to achieve), and aspirational, or benchmark, values (i.e. values that can be employed as a best practice goal), for a range of quality indicators. Where possible, responses were based on standardized information (national collected data, or large data sets), but in most countries such data are not available. Another survey, the 'ESHRE survey', provided information on current practice (how many KPIs are measured, frequency of measurement, characteristics of a reference population for KPIs) and the degree of importance of some indicators. This survey was sent to 2413 members of the ESHRE Special Interest Group (SIG) of Embryology, and 384 responses were received. In addition, where relevant, published data were collected from a literature search and summarized, although for most indicators, especially in ART, there is a general lack of evidence to support their importance, scientific soundness and usefulness (Shahangian and Snyder, 2009).

During the consensus meeting, the results of the surveys, scientific evidence and personal clinical experience were integrated into presentations by experts on specific topics. For each indicator, information was presented in a fixed format: definition, rationale, qualifiers, formula, data sources, KPI strengths and weaknesses, frequency of data collection and reference values for minimum expected and target value based on 50 and 75 percentile values, respectively. After the presentation for the topic, each proposed indicator was discussed until consensus was reached within the group.

After the meeting, a report was prepared describing the presentations (workshop report) and the consensus points. After approval of the report by the meeting participants, the national and international societies that contributed to the questionnaires were invited to review the report and submit comments. The final version of the manuscript was approved by the Executive Committees of ESHRE and Alpha before publication.

This paper is divided into two parts: the workshop report and the recommendations of the expert panel.

## Results (Workshop report)

### Effects of ovarian stimulation on embryology parameters

The methods of ovarian stimulation have been evolving since the earliest days of clinical IVF, in the search for the best stimulation protocol. With that goal in mind, there has been an enormous effort to develop the best pharmaceuticals and protocols, but in practice, economic factors as well as prevailing opinion can influence treatment decisions beyond consideration of the patient's endocrine background.

Despite the thousands of smaller studies on this subject in the literature, there are very few large multicentre RCTs beyond those organized by pharmaceutical companies, and these do not consider 'non-standard' groups of patients. Furthermore, the meta-analyses that have been conducted are often unclear about their inclusion criteria. As a result, the prevailing approach is to use what works within each clinic, and is often uncorrected for the patients' weight, BMI or endocrine background. Some clinics also do not monitor cycles, for financial reasons, which can have an influence on the cycle outcome.

In general, a 'good' stimulation is one that produces a homogeneous cohort of mature oocytes, with the least inconvenience and risk to the patient, and results in the birth of a healthy singleton. From the Clinical Embryologist's perspective, a good stimulation results in the retrieval

of well-expanded cumulus-oocyte complexes (COCs), as is expected from each follicle >14 mm in diameter, with a high proportion of metaphase II (MII) oocytes (Scott *et al.*, 1989; Ectors *et al.*, 1997; Nogueira *et al.*, 2006). On the other hand, a poor stimulation, caused for example by sub-optimal decisions regarding timing or stimulation dose, is one that may result in a high rate of abnormal COC morphology observed at the time of oocyte retrieval, possibly resulting in an increased rate of abnormal fertilization (e.g. one pronucleus [PN], 3PN, etc.) and/or decreased rate of normal fertilization and an increased aneuploidy rate (Soares *et al.*, 2003).

Aggressive ovarian stimulation has effects on the patient's well-being, by increasing the risk of ovarian hyperstimulation syndrome, (Delvigne, 2009; Rizk, 2009) as well as on the endometrium and the ovaries. There are studies showing an increased likelihood of embryo aneuploidy in connection with aggressive ovarian stimulation, even in patients younger than 35 years, including post zygotic segregation errors and maternal segregation errors (Baart *et al.*, 2007; Haaf *et al.*, 2009), as well as having a negative impact on the maintenance of genomic imprints during early embryogenesis (Fauque *et al.*, 2007; Denomme and Mann, 2012; Saenz-de-Juano *et al.*, 2016). It has been shown that in patients belonging to either high or low strata of antimüllerian hormone levels an inverse correlation exists between the daily dose of recombinant human FSH used in the stimulation and the proportion of blastocyst formation (Arce *et al.*, 2014).

To determine whether there is a stimulation method that could yield a higher number of competent oocytes, one first needs to consider the effects of LH and FSH as the principal drivers of ovarian stimulation, and their pharmacodynamics. In a natural cycle, FSH receptor expression peaks during the early follicular phase then declines, while LH receptor expression increases from mid-follicular phase, indicating that LH is likely to be involved in follicular development (Jeppesen *et al.*, 2012). This could explain, at least in part, why follicular recruitment is compromised in women with profound pituitary downregulation (Ferraretti *et al.*, 2004). The role of the theca cells in ovarian responsiveness to FSH is also illustrated by compromised follicular recruitment in women older than 35 years (Piltonen *et al.*, 2003; Humaidan *et al.*, 2004; Hugues *et al.*, 2010), particularly those with reduced ovarian sensitivity to FSH (Davison *et al.*, 2005) and reduced ovarian capacity to secrete androgens under basal LH conditions (Spencer *et al.*, 2007). In addition, it has been shown that LH induces epidermal growth factor-like factors in the mural granulosa. Amphiregulin, one of these factors, has been correlated with good oocyte quality (Zamah *et al.*, 2010; Sugimura *et al.*, 2015).

The time of oocyte retrieval relative to the time of the ovulation trigger is typically in the range of 34–38 h. In a meta-analysis including five RCTs with a total of 895 women, it was found that a time of oocyte retrieval relative to the time of the ovulation trigger of >36 h compared to <36 h resulted in a higher oocyte maturation rate, but no difference in fertilization rate, implantation rate or pregnancy rate (Wang *et al.*, 2011). Deviations from the locally established protocol should be documented and taken into consideration.

In conclusion, to optimize outcomes, there is a need for individualization of the treatment protocol. As stimulation can affect a range of systems, closely monitoring the patient's response could reduce many of the risks associated with ovarian stimulation, as stimulation can affect a range of systems, which will also have an impact on the success of treatment.

Because of the interlinked effects of ovarian stimulation on oocyte quality and developmental competence, it is logical that the most successful clinics are those in which the embryologists and clinicians speak with each other and communicate regularly and effectively about outcomes related to stimulation (and other procedures) (Van Voorhis et al., 2010).

## Oocytes

Not all oocytes collected from a patient following ovarian stimulation for ART will have the same developmental competence, which is illustrated by the observation that only 5% of oocytes collected eventually result in a live birth (Lemmen et al., 2016). Intrinsic oocyte competence is derived not only from the degree of nuclear maturity of the oocytes, but also from their cytoplasmic maturity (Patrizio and Sakkas, 2009; Garrido et al., 2011; Lemmen et al., 2016). Furthermore, oocyte developmental competence is affected by a range of intrinsic patient-related and external factors. These patient-related factors include age, BMI, lifestyle factors and type of infertility. The external factors include ovarian stimulation, laboratory procedures (such as oocyte retrieval, denudation, cryopreservation, preparation for ICSI), culture conditions (temperature, pH, pO<sub>2</sub>), environmental conditions (light, air quality, humidity) and culture medium.

Against that background, the question is whether any laboratory indicator can provide a measure of the intrinsic oocyte competence at the time of oocyte retrieval, as all of the subsequent events could be influenced by laboratory procedures, and/or by the genetic contribution of the spermatozoon. In other words, is quality measurable for oocytes, or perhaps more pertinently, is there any measure that could pinpoint where a dysfunction occurred during the long process of oocyte development?

There are a number of possible markers for oocyte competence, but these are largely research-based, and have not found widespread application in clinical service. These include assessment of biochemical markers in follicular fluid, gene expression studies of follicular cells and oxygen uptake assessments (Nagy et al., 2009; Nel-Themaat and Nagy, 2011). Other markers, such as assessment of oocyte morphology, spindle imaging, and polar body (PB) biopsy, can be incorporated into clinical service, but this is not a universal approach (Patrizio et al., 2007; Rienzi et al., 2011; Braga et al., 2013).

When the results of the Alpha and ESHRE surveys were combined, respondents identified oocyte recovery rate and oocyte maturity rate as the most important indicators for oocytes. Although, strictly speaking, they do not provide an indication of laboratory performance, they do provide an estimate of response to stimulation, and therefore a general estimation of the likely developmental competence of the oocytes.

In the Alpha survey, oocyte recovery rate was defined as the likelihood of aspirating a COC from each follicle over a certain size as measured on the day of triggering. The rationale for this is the expectation that those follicles that have achieved a certain size, with a good response to FSH and a sufficient number of LH receptors in follicular cells, will respond appropriately to the ovulation signaling cascade, resulting in the release of the COC into the follicular fluid, thereby facilitating its aspiration. A concern with this potential indicator was its reliance on the accuracy of follicular scanning, and the need for a consistent time interval between ovulation trigger and oocyte retrieval. However, while a range of follicle sizes was identified in the survey as the 'ideal' size for triggering, the expected recovery rates were

remarkably similar, generally ranging from 70 to 80% as the competence level, and 85–100% as the benchmark value.

Oocyte maturity rate is generally related to nuclear maturity, being defined as the proportion of oocytes at MII stage. Its potential value is as a marker of the efficiency of ovarian stimulation and triggering. Of the Alpha survey respondents, 80% indicated that their laboratory determined the MII rate, with median competence and benchmark values of 75 and 90%, respectively. It was noted that the timing of this assessment is an important factor, as it is not possible to assess oocyte maturity at the time of oocyte retrieval in the case of insemination by routine IVF. Since this assessment requires the removal of the cumulus and corona cells, it can be performed at the time of denudation of the oocytes prior to ICSI, but for a universal competence and benchmark value to be established, a consistent time interval between the time of trigger and the time of cumulus cell removal would be required (e.g. 40 ± 1 h post ovulation triggering).

A third potential indicator, oocyte grade, was defined as the proportion of COCs with expanded cumulus at the time of oocyte retrieval. As ovulation triggers cumulus expansion by mediating the synthesis of hyaluronic acid and the organization of a stable cellular matrix (Russell and Salustri, 2006), this indicator provides an assessment of the quality of communication between the oocyte and its cumulus cells. The potential problems associated with the use of this indicator are a lack of objective criteria for making this assessment, and a concern that there is not always a good correlation between nuclear maturity and cumulus cell expansion (Balaban and Urman, 2006).

Other indicators that were proposed but not considered valuable or reliable by the survey respondents were: rate of degenerated (or empty) zonae pellucidae; rate of germinal vesicle oocyte recovery; definition of the minimum number of follicles to justify flushing; and oocyte degeneration rate at the time of oocyte retrieval. Oocyte degeneration rate after removal of cumulus cells is discussed in the section on ICSI.

Overall, the responses to the surveys highlighted the lack of consistent data for the evaluation of oocyte quality and competence, and identified an opportunity for national and international registries to promote the collection of this information.

## Sperm KPIs

Proposed andrology laboratory PIs were sperm recovery rate, and sperm motility post-wash. In addition, sperm parameters were discussed in relation to the decision for IUI, IVF or ICSI.

Survey responses revealed such wide ranges in perceived semen analysis minima for suitability for IUI, IVF or ICSI, as well as expected sperm recovery post-wash, as to make it impossible to determine robust recommendations for competency and benchmark values for any of these criteria.

With regard to sperm preparation, it is possible that respondents were confused when reporting the 'recovery rate', and the substantial variability in terms of the expected/required number of spermatozoa in the final preparation likely included confusion between percentage recovery ('yield') values and the actual number of spermatozoa (millions); hence these data were considered unreliable. However, the expected proportion of motile spermatozoa in the final washed preparation showed coherence across the respondents, with both median and mode values of 90% for competency and 95% for the benchmark.

A major issue when considering semen analysis data is that many ART laboratories do not employ methods that meet the minimum standards required by either the World Health Organization (WHO) or the ESHRE SIG Andrology (Björndahl *et al.*, 2010; World Health Organization, 2010). Therefore, reported values for sperm concentration and motility must be understood to have high uncertainty of measurement, and hence need to be considered as inaccurate and unreliable (Sanchez-Pozo *et al.*, 2013; Björndahl *et al.*, 2016). As a consequence, any association between semen analysis characteristics, yield and fertility potential will remain unclear if based on studies using inappropriate semen analysis techniques. From a best practice standpoint, any clinical laboratory providing semen analysis or post-preparation values that are to be used for diagnostic or treatment management purposes should participate in an EQA programme which provides a comparison between the participating laboratories' results and established reference ('correct answer') values so as to permit quality improvement in laboratory work (Björndahl *et al.*, 2010).

Moreover, there is a general concern that semen analysis reference values have little or no value for ART procedures (Björndahl, 2011). In particular, the WHO reference values for sperm concentration, motility and vitality were derived from populations of men who had achieved *in vivo* conceptions (Cooper *et al.*, 2010), and therefore these cut-off values have no *a priori* relevance in regard to ART patients, and hence the need or suitability for any form of ART treatment should not be decided based on these reference values.

The Tygerberg Strict Criteria for normal sperm morphology were derived in regard to ART success (Kruger *et al.*, 1988; Coetzee *et al.*, 1998), so these cut-offs might be pertinent in differentiating between the need for IUI, IVF or ICSI—although concern regarding measurement uncertainty cannot be ignored (Menkveld *et al.*, 2011). While a cut-off of 4% normal forms might help to define sub-populations of patients with differing prognoses, at the level of individual patients a result of 4% based on 200 spermatozoa evaluated is not very informative since the result has a statistical expectation ranging from 2 to 8% (Björndahl *et al.*, 2010), and to be able to differentiate between 3 and 5% with statistical robustness would require the assessment to have been made by evaluating over 1500 spermatozoa.

Because of the limitations of semen analysis cut-off values, a decision on the suitability of IUI or routine IVF for a couple should be made based on post-preparation sperm numbers/concentration and their motility, ideally assessed during a pre-treatment 'trial wash' (while still taking into account the uncertainty of measurement). In case of ICSI, there should not be any cut-off based on semen analysis characteristics, the only logical criterion would be having sufficient (in comparison with the expected number of oocytes) spermatozoa that are, ideally, viable, and preferably motile or positive in the hypo-osmotic swelling (HOS) test (Nagy *et al.*, 1995). As a general principle, laboratories should develop and apply their own criteria for deciding on IUI, IVF or ICSI, based on the couple's clinical situation and reproductive history rather than semen analysis.

Regarding andrology laboratory PIs, results from the Alpha survey indicated that only post-preparation sperm motility would be a valuable indicator, as it monitors the effectiveness of the sperm washing procedure. Therefore, post-wash sperm motility should be monitored for fresh ejaculate specimens that show normozoospermia as per the WHO5 guidelines (World Health Organization, 2010), but still taking into account the poor reliability of sperm motility data, non-robust

classification of semen samples based on the high uncertainty of semen analysis data, and variability in sperm preparation methods.

Sperm recovery rate, defined as the percentage recovery of progressively motile sperm after washing as compared to pre-washing (Björndahl *et al.*, 2010), can be used as a laboratory KPI, providing useful information for inter-operator comparison and proficiency testing. However, given the high uncertainty in counting and the different protocols for sperm preparation (notably with density gradient washing resulting in higher recovery rates as compared to direct swim-up from semen), no competence values can be provided. Laboratories should develop their own standards according to their own clinical and laboratory practice.

## Fertilization after insemination by ICSI

Although several potential KPIs have been identified in ICSI, the presentation focused on the four most pertinent, i.e. normal fertilization rate, oocyte degeneration rate, poor fertilization rate and failed fertilization rate.

The definition used most often for the ICSI normal fertilization rate is the proportion of injected oocytes with 2PN the day after injection, except for the Spanish Registry and the Istanbul Consensus which include the observation of two PBs in the definition (Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology, 2011). The suggested competence and benchmark values for this indicator were consistent among respondents ranging from 60 to 80%, and 70 to 100%, respectively. The UK's Association of Clinical Embryologists proposed benchmark for the 2PN rate is >65% including only patients below 40 years of age with at least three oocytes collected (Hughes and Association of Clinical Embryologists, 2012). From the literature, it was found that ICSI results in an average fertilization rate of 70% (Heindryckx *et al.*, 2005; Kashir *et al.*, 2010), which was similar to 68.7% based on over 20 000 unselected MII oocytes at the CARE Fertility laboratory, UK (personal communication Alison Campbell). ICSI fertilization rate is a commonly reported and effective indicator that is informative of gamete quality and operator competence. ICSI 2PN rate does depend on the various criteria used for performing ICSI, which can be considered a weakness of the indicator.

The ICSI damage rate or oocyte degeneration rate was ranked as important. In the Alpha survey, the minimum expected value and target value ranged from 3 to 30% and 0 to 10%, respectively. Oocyte damage can be observed at three time points during the ICSI process: from the start at stripping, during ICSI, or at the fertilization assessment on Day 1. Oocyte damage probably occurs most frequently during injection, but without immediate signs of damage, this is not detected until the fertilization check. In addition, as both the damage detected at injection and at fertilization check reflect damage from the ICSI process, these should not be recorded and calculated separately. Damage at denudation/stripping can be monitored separately as it mainly reflects operator's competency, but it has a very low frequency. ICSI damage rate is therefore defined as the number of oocytes damaged during ICSI, and/or observed at fertilization check over the number of injected oocytes. It is useful to monitor this indicator for operator competence, oocyte quality, and laboratory performance. The damage rate can also be indicative of technical problems (e.g. cumulus cell removal stress, vibration). Alternatively, the term ICSI oocyte survival rate can be used.

Poor fertilization rate is defined as the proportion of cycles in which <25% of the injected oocytes are fertilized. The responses from the



survey are much divided, ranging from 5 to 20% for the competence and from 0 to 15% for the benchmark value. Poor fertilization rate can give an indication of operator competence and reflect gamete quality.

Failed fertilization rate scored relatively low on importance in the surveys. Failed fertilization rate is defined as the proportion of cycles in which none of the injected oocytes are fertilized. The indicator can be informative of gamete quality/function and/or operator skill. A deficiency in the mechanism of oocyte activation is regarded as the principal cause of ICSI fertilization failure or abnormally low fertilization. Complete ('or virtually complete') fertilization failure with ICSI occurs in 1–5% of cycles (Liu et al., 1995; Flaherty et al., 1998; Mahutte and Arici, 2003; Yanagida, 2004; Kashir et al., 2010). From the Alpha survey there was a wide range in the competence values for this indicator ranging from 2 to 15% (median 5%), with a benchmark of <1%.

For these and other KPIs, a reference population could be relevant. With regard to ICSI fertilization rate, it could be relevant to exclude cases where reduced fertilization rates are anticipated, including *in vitro* matured metaphase I oocytes (although inconclusive data), artificially activated oocytes, the use of testicular sperm, and cases of globozoospermia and asthenozoospermia (Rubino et al., 2016). However, owing to the low prevalence, including these cases may not significantly affect the indicators in most clinics.

In conclusion, from the surveys and collected evidence, ICSI damage rate and ICSI normal fertilization rate are considered relevant and important KPIs, while the value of ICSI low/failed fertilization rate as a KPI is less clear. Oocyte maturity rate and IPN/3PN rate were not specifically discussed for ICSI. The ICSI rate, defined as the proportion of cycles that use ICSI, was not commonly recorded by the survey respondents, judging this as a less relevant PI.

## Fertilization after (routine) IVF insemination

This section deals with normal fertilization rate, polyspermy rate, poor fertilization rate and zygote morphology after routine IVF insemination.

Pronuclear formation occurs 1.5–2.0 h earlier in oocytes inseminated by ICSI compared with those inseminated by conventional IVF (Nagy et al., 1998; Montag et al., 2001). This should be taken into consideration when setting the time for fertilization check, relative to the time elapsed since insemination (recommended as  $17 \pm 1$  h) (Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology, 2011). A normally fertilized oocyte should have 2PN of similar size that are closely apposed and centrally located (Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology, 2011). Although the provenance of micronuclei remains unclear, their presence could be considered to be abnormal, as could the presence of pronuclei of different sizes. From the literature, the normal fertilization rate (i.e. the proportion of inseminated oocytes with 2PN at the time of the fertilization check on Day 1) is expected to be around 67%, with a range of reported values of 53–81%. This range was reflected in the Alpha survey results, where the median value for competency was 61% of inseminated oocytes, with a median benchmark value of 77%.

The presence of three or more pronuclei is indicative of an abnormal fertilization, arising from either nondisjunction (failure to extrude the second PB) or polyspermy. Polyspermy may be the result of either oocyte immaturity (causing failure of the cortical reaction), oocyte overmaturity and/or an extremely high concentration of motile

spermatozoa in the insemination volume (Wang et al., 2003). From the literature, the incidence of  $\geq 3$ PN is 4–7% in IVF (Joergensen et al., 2015). This agrees well with the median values from the Alpha survey of <9% for competence, and <4.5% as a benchmark.

The calculation of the total fertilization rate following IVF includes all fertilized oocytes with  $\geq 2$ PN. Although, as already stated, oocytes with >2PN are abnormally fertilized, this parameter provides an indication of the ability of the culture system to support sperm capacitation and sperm-oocyte interaction in IVF cycles. Of the published studies that included >100 oocytes, the median total fertilization rate was 76% (range 69–87%) in selected patient populations. This is similar to the Alpha survey results, which suggested a competency level of at least 63% and a benchmark of at least 84%.

Oocytes with a single pronucleus after insemination by IVF, which occurs in 1–5% of cases, can be indicative of fertilization and syngamy, asynchronous appearance of pronuclei (an extremely rare event, as evidenced by the use of time-lapse microscopy), or parthenogenetic activation (Levron et al., 1995). The incidence of diploidy in IPN oocytes following conventional IVF has been reported to be in the range of 45–50% (Sultan et al., 1995; Staessen and Van Steirteghem, 1997; Kai et al., 2015). In contrast, IPN oocytes arising after ICSI have a reported diploidy rate of only 7–14%, with genetic abnormalities in the subsequent embryos (Mateo et al., 2013).

The incidence of poor fertilization (<25% of inseminated COCs with 2PN) or total failure of fertilization (no oocytes with signs of fertilization) could be indicative of a problem with sperm function, too few motile spermatozoa during insemination, or failure of oocyte activation (Ebner et al., 2015). There is very little evidence in the literature regarding the expected incidence of either poor or failed fertilization. However, the Alpha survey results suggested competency and benchmark levels for poor and failed fertilization of 14 and 6%, and 8 and 4%, respectively.

Although the Istanbul Consensus made recommendations about grading zygote morphology (Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology, 2011), the question remains as to whether indicators referring to zygote morphology are useful, especially as differences in pronuclear pattern could be related to the insemination method and timing of the observation (Montag et al., 2001; Ebner et al., 2003).

## Cleavage-stage embryos

Proposed indicators for cleavage-stage embryos are early cleavage rate, cleavage rate, embryo development rates, embryo fragmentation rate, and rate of good quality embryos (embryo score or grade).

Early cleavage rate is defined as the proportion of cleaved zygotes at the early cleavage check on Day 1 ( $26 \pm 1$  h post-ICSI or  $28 \pm 1$  h post-IVF) (Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology, 2011), but other time-intervals after insemination have been used for assessing early cleavage. This indicator reflects the ability of the culture system to support early cleavage of fertilized oocytes and the viability and quality of the embryos (Shoukir et al., 1997). There are conflicting results on the importance of early cleavage. Studies have shown that early cleavage, together with other factors, can be used as an embryo selection method (Lundqvist et al., 2001; Ciray et al., 2005). Early cleavage rate has also been shown to correlate with blastocyst implantation and pregnancy rates (Shoukir et al., 1997; Balaban and Urman, 2003) and it is a better independent marker of implantation

potential than zygote morphology (Brezinova *et al.*, 2009). In contrast, early cleavage was not found to be an independent predictor of implantation in IVF patients with good prognosis (Thurin *et al.*, 2005). In addition, premature occurrence of early cleavage can be negatively, instead of positively, associated with embryo implantation potential (Meseguer *et al.*, 2011). Furthermore, it was not a reliable predictor for embryo implantation rate when good quality embryos are transferred (Sundstrom and Saldeen, 2008; de los Santos *et al.*, 2014), or when using a GnRH antagonist protocol (Yang *et al.*, 2015). In the Alpha survey, competence and benchmark values ranged from 10 to 50% and 15 to 60%, respectively.

Cleavage rate reflects the ability of the culture system to support cellular division of fertilized oocytes. It is an indicator of embryo viability and has the ability to detect culture media contaminants. The presence of non-cleaved embryos or arrested zygotes on Day 3 is associated with a decrease in quality of the remaining cohort, but without a negative impact on clinical outcome (Machtiger *et al.*, 2015). Cleavage rate is considered important and is widely monitored, and defined as the proportion of zygotes which cleave to become embryos on Day 2 at  $44 \pm 1$  h post-insemination (Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology, 2011). Cleavage rate should be calculated not only in the total population, but also in reference groups (IVF versus ICSI, female age, ejaculated versus surgically retrieved sperm), and controlled for confounders (the timing of observation and oocyte maturity). Also, the presence of a refractile body in the oocyte is associated with reduced cleavage rates and impaired embryo development, while the cytoplasmic granularity did not seem to have an effect on embryo development (Fancsovsits *et al.*, 2012). Cleavage rate should be calculated frequently in a laboratory (at least once per month). In the Alpha survey, the competence values ranged from 80 to 95%, with a benchmark of 90 to 100%.

Embryo development rate is defined as the proportion of 4-cell embryos on Day 2 among the 2PN zygotes (measured at  $44 \pm 1$  h post-insemination), the proportion of 8-cell embryos on Day 3 (measured at  $68 \pm 1$  h post-insemination), and the proportion of morula-stage embryos on Day 4 ( $92 \pm 2$  h post-insemination) (Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology, 2011). This indicator reflects the ability of the culture system to support cleavage according to the expected developmental stages and the quality and viability of embryos, especially for Day 2 or 3 transfer, while less important for blastocyst transfer. Possible confounders are the timing of laboratory observations and the type of culture media used. Although dependent on iatrogenic factors such as the culture conditions, embryo development rate is an important indicator; in well-defined categories of patients, it reflects the overall laboratory performance. It was stated that sufficient numbers of embryos or longer data collection are necessary, as this indicator is influenced by short-term variations. Calculation of a sliding mean can be helpful for detecting long-term variations. The value and practicability of the morula check on Day 4 was questioned as centers performing Day 5 embryo transfer often do not assess the embryo development at Day 4. Also, instead of assessing the number of 4-cell embryos on Day 2, or 8-cell embryos on Day 3, a combination of 4-cell and 8-cell embryos on Days 2 and 3 can be used. This was analyzed in a study by van Royen, who thereby characterized a top quality embryo as having 4- to 5-cells on Day 2 and  $\geq 7$  cells on Day 3 (Van Royen *et al.*, 1999). In larger centers, assessing the embryos at the specified time points can be difficult, and counting 4- and 5-cell embryos on Day 2, and 7-, 8- and 9-cell embryos on Day 3 may be more relevant.

The rate of good quality embryos is defined as the proportion of Days 2 and 3 embryos with high score or grade. Many different scoring systems exist, based on different variables, including cell number, fragmentation, cell size and multinucleation (Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology, 2011) and blastomere nuclear status (Fauque *et al.*, 2013). A recent study evaluating which set of embryo variables is most predictive for live birth rate (LBR) reported that blastomere number, proportion of mononucleated blastomeres, degree of fragmentation and variation in blastomere size were significantly associated with LBR in univariate analysis, while symmetry of the embryo was not (Rhenman *et al.*, 2015). Furthermore, the grading systems are not robust, but can be used as internal quality assessment parameters. The importance of this parameter is also affected by the limited significance of the fragmentation rate.

Embryo fragmentation rate, defined as the proportion of Days 2 and 3 embryos with  $<10\%$  fragmentation, reflects the quality and viability of embryos. From the Alpha survey, the competence value for this parameter ranged from 20 to 90% (median 50%) and the benchmark from 30 to 90% (median 70%). These large ranges underline the difficulties with this parameter: embryo fragmentation rate is reported to be a subjective parameter and difficult to evaluate as one has to differentiate between a cell and a fragment and then estimate the relative proportion of fragments (Paternot *et al.*, 2011).

Embryo utilization rate is defined as the number of embryos utilized (transferred or cryopreserved) per number of 2PN zygotes in the same cycle. This parameter is often presented in studies, but competence and benchmark values cannot be calculated because of its dependence on strategies for embryo transfer and cryopreservation, as well as patient request.

In conclusion, embryo cleavage rate and embryo development rate are extremely important indicators, while early cleavage rate, rate of good quality embryos and embryo fragmentation rate are less important as quality indicators.

## Blastocyst development

In the case of blastocyst-stage embryo transfer, several parameters were suggested as indicators: blastocyst development rate, good blastocyst development rate, the proportion of good quality blastocyst and Day 5 embryo transfer rate.

The blastocyst development rate, defined as the proportion of 2PN zygotes (not just of cleaved zygotes) which are at the blastocyst-stage at Day 5 ( $116 \pm 2$  h post-insemination) (Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology, 2011), was rated important because it reflects the efficiency of the whole culture system. Blastocyst development rates can be calculated on Day 5, 6 or Days 5/6 combined. Assessment and calculation on Day 5, consistent with previous consensus, is preferred based on limited numbers of embryos available on Day 6. The competence and benchmark values for blastocyst development rate on Day 5 ranged from 25 to 60% and 44 to 80%, respectively. The blastocyst development rate is an objective parameter, but dependent on the assessment of blastocyst morphology, which is straightforward in case of good quality blastocysts, but can be challenging for embryos showing an attempt of cavitation. Confounders can be the timing of laboratory observation, the culture medium and the culture conditions (in particular the  $pO_2$  concentration).

A good blastocyst development rate is defined as the proportion of 2PN zygotes which are good quality blastocysts on Day 5 ( $116 \pm 2$  h post-insemination) (Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology, 2011). Similar to blastocyst development rate, measuring this as an indicator only on Day 5 could make the indicator more robust. Blastocyst quality should be based on three factors, namely blastocoele expansion, appearance of trophoctoderm (TE) and appearance of inner cell mass (ICM). Although all three parameters have been shown to be significantly correlated to pregnancy and LBR (Van den Abbeel et al., 2013), only TE was found to be a statistically significant independent predictor of live birth outcome after adjustment for known confounders (Ahlstrom et al., 2011). Even though the ICM is important, a strong TE layer is essential at this stage of embryo development, allowing successful hatching and implantation (Ahlstrom et al., 2011). In the Alpha survey, the competence and benchmark values for good blastocyst development rate on Day 5 ranged from 15 to 45% and 25 to 80%, respectively.

The proportion of good quality blastocysts can be calculated from the blastocyst development rates and good blastocyst development rates. There is no evidence pertaining to the significance of this parameter in the literature, and no data from the Alpha survey.

Day 5 embryo transfer rate was defined as the proportion of cycles with  $\geq 1$  2PN zygotes on Day 1 that had  $\geq 1$  blastocyst for transfer on Day 5. From the Alpha survey there was a large variation in the values for competence and benchmarks (ranging from 25 to 90% and 40 to 100%, respectively), which is assumed to be related to differences in the time of blastocyst assessment, and different grade of expansion. In addition, this parameter depends on different policies for transfer in different settings. PGD/PGS (globally indicated as preimplantation genetic testing—PGT—that includes PGD for single gene disorders or for chromosome structural abnormalities, and PGS for aneuploidy) cycles should be excluded from this calculation. Some possible confounders are the timing of laboratory observation (ideally at  $116 \pm 2$  h post-insemination), but in some laboratories the time of observation depends on the timing of embryo transfer (physician availability), the culture medium and other culture conditions (e.g.  $pO_2$  concentration), and the grade of blastocyst expansion.

#### *Fresh cleavage or fresh blastocyst transfer?*

Recently, a Cochrane review summarized evidence from RCTs comparing the reproductive outcomes after fresh cleavage-stage versus fresh blastocyst-stage embryo transfer (Glujovsky et al., 2016). Based on low-quality evidence, they concluded that the LBR following fresh transfer was higher in the blastocyst transfer group as compared to the cleavage-stage transfer group [odds ratio (OR) 1.48, 95% CI 1.20–1.82; 13 RCTs, 1630 women,  $I^2 = 45\%$ ]. This is translated in a LBR of 29% after fresh cleavage-stage transfer, and between 32 and 42% after fresh blastocyst-stage transfer. The five RCTs that reported cumulative pregnancy rates after fresh and frozen transfers, showed no significant difference after one round of oocyte retrieval (based on very low-quality evidence). The main limitation was serious risk of bias, associated with failure to describe acceptable methods of randomization, and unclear or high risk of attrition bias (Glujovsky et al., 2016). Another review reported that blastocyst (Day 5/6) transfer in a fresh IVF/ICSI treatment cycle significantly increased LBR (OR 1.77; 95% CI 1.32–2.37), clinical pregnancy rate, implantation rate, and ongoing pregnancy rate, and reduced first trimester miscarriage rate,

in comparison with cleavage-stage (Day 2/3) embryos transfer based on seven RCTs and 1446 cases (Wang and Sun, 2014).

### Implantation rate and LBR

Implantation rate is judged an important indicator that reflects the overall performance of the laboratory and an overall low implantation rate is a serious alert. Implantation rate is defined as the number of gestational sacs observed divided by the number of embryos (cleavage-stage or blastocysts) transferred (Zegers-Hochschild et al., 2009), or as the proportion of fetal hearts relative to the number of embryos transferred (Alpha Scientists In Reproductive Medicine, 2012). Implantation rate reflects the efficiency of the whole culture system, but it can be influenced by uterine receptivity, and by different policies for embryo transfer in different centers.

LBR may be considered as the ultimate KPI for checking IVF clinic performance and defined as the likelihood of a baby to be born per embryo transferred. LBR is largely affected by a series of clinical maternal factors pertaining to post-implantation development, rather than reflecting laboratory performance. This parameter can be calculated annually, but it is often difficult to collect the data.

### PGT: PGD/PGS

The indicators proposed for PGD/PGS (PGT) were successful biopsy rate/tubing rate, rate of no biopsy, proportion of samples submitted to analysis where no results were available (no DNA was detected), and embryos not found on warming. Other indicators suggested in the survey were 'survival after warming', embryo transfer per PGT cycle, and miscarriage rate, but there were very few responses for these variables in the Alpha survey.

A successful biopsy rate is defined as the proportion of biopsied and tubed/fixed samples where DNA is detected. It was suggested to use this parameter combining the tubing rate (the proportion of embryos where cells were tubed) and the proportion of samples submitted to analysis where no DNA was detected, as the tubing of the cells can often not be inspected visually and will only be detected by the presence of DNA after amplification. A benchmark for the tubing rate of 95% was reported in the Alpha survey. Confounders for the successful biopsy rate are embryo quality, and the criteria for biopsy (for Day three embryos: presence of a visible nucleus; for Day 5/6: grade of hatching, TE quality). In the data of the ESHRE PGD Consortium, of 254 820 samples that were biopsied, 91.3% were diagnosed (De Rycke et al., 2015). It is important to underline that, as already mentioned, these data mainly refer to Day 3 biopsies and that in the case of PGT for aneuploidy (generally known as PGS) the analysis of biopsies was mainly by fluorescence *in situ* hybridization (FISH), a technique requiring cells to be fixed on a glass slide. FISH is now being replaced by other methods providing 24-chromosome testing, which require tubing of cell biopsies.

The rate of no biopsy is defined as the proportion of intended PGD/PGS (PGT) cases where there were no embryos available to biopsy. This parameter was rated important, but it reflects patient-related factors and the ability of the culture system to support cleavage/blastocyst formation, rather than the performance ability of the laboratory to perform a PGD/PGS (PGT) treatment/analysis. Furthermore, the parameter is different whether biopsy is performed at the cleavage or blastocyst-stage, and depends on the timing of the



laboratory observation, culture medium and culture conditions (e.g. pO<sub>2</sub> concentration), criteria for biopsy, the time point of assisted hatching (Day 3 or 5), and patient selection. The parameter should also be calculated for PGD and PGS separately, based on the difference in patient populations. From the Alpha survey, the median competence value and benchmark were 20 and 10%, respectively. In the last data collection of the ESHRE PGD Consortium, out of 45163 reported cycles, 2.8% were canceled before biopsy (De Rycke *et al.*, 2015). However, as this data collection covers cycles performed up to 2010, this figure refers to Day 3 biopsies, implying that possibly a higher value could be expected for Day 5 biopsies.

For a number of samples, no results are available after amplification. Results can be inconclusive even with a strong DNA amplification band in the agarose gel, which can be due to early fragmentation of DNA. In single gene analysis, results can be inconclusive if information is found only on a reduced number of markers, or the gene of interest failed to amplify. Although this indicator is not strictly related to the performance of the laboratory, it is relevant to inform clinicians on the total number of embryos where PGD/PGS (PGT) results are available. The number of samples where no results are available after amplification was reported in the Alpha survey with a benchmark of <5%, which corresponds with the converse of the tubing rate.

Nowadays, the majority of PGT cycles are based on blastocyst biopsy and cryopreservation. The proportion of embryos not found on warming and embryos degenerated after warming mainly reflects operator skills. In the Alpha survey, the median competence values were 3 and 10% for not found and degenerated embryos, respectively. These indicators can also be reported as one combined indicator. It was mentioned that clinicians should be aware that although laboratories strive for 100%, not all embryos submitted to PGD/PGS (PGT) will be recovered after warming.

Reference values for implantation rate and LBR after PGT can be derived from the data from the ESHRE PGD Consortium. The implantation rate was 26% for PGD and 22% in PGS, with delivery rates (per oocyte retrieval) of 21 and 14%, respectively (De Rycke *et al.*, 2015).

### Indicators for cryopreservation: addition to the previous consensus

Blastocyst re-expansion rate is defined as the proportion of warmed blastocysts that show re-expansion within a defined time period (e.g. 2 h) (Alpha scientists in reproductive medicine, 2012). Recent evidence shows an impact on the performance results depending on the quality/expansion of the blastocysts which are cryopreserved (Cobo *et al.*, 2012). Also, in blastocyst fresh transfer, from multivariate analysis it was shown that the odds of live birth increased by 36% for each grade of expansion ( $P = 0.0061$ ) and decreased by 29% for blastocysts with grade B TE compared with Grade A TE ( $P = 0.0099$ ). Furthermore, after thawing, the odds of live birth increased by 39% ( $P = 0.0042$ ) for each 10% increase in degree of re-expansion. Therefore, blastocoel expansion and TE grade were selected as the most significant pre-freeze morphological predictors of live birth and degree of re-expansion was selected as the best post-thaw parameter for prediction of live birth (Ahlstrom *et al.*, 2013). Confounding factors are time of observation, female age and fertilization method. These observations do not include embryos that had been biopsied on Day 3 as they have a different hatching dynamic (Lopes *et al.*, 2015).

## Recommendations of the expert panel

### General comments

- (1) Regarding frequency of data collection for indicators, it was the consensus opinion that this should be done, ideally, on a monthly basis. However, it was recognized that this is not always practical, based on caseload, and therefore either a longer timeframe or a specific predetermined number of cases might be used instead. The minimum number will depend on the stability of the indicator and will need to be developed by the laboratory, although an initial dataset of 30 cases could be used as a guideline. Nonetheless laboratories should remain vigilant and respond promptly to unexpected fluctuations.
- (2) The discussions identified three different types of indicators: reference indicators (RIs), PIs and KPIs.
  - RIs were related to the oocytes coming into the laboratory, and so were proxy indicators of the response to ovarian stimulation.
  - PIs were those for which data should be documented and stored, even if they are not routinely reported in a control chart.
  - KPIs were those related to the 'core business' of the ART laboratory.
- (3) The values for indicators are presented as competency and benchmark values, as was done for the cryopreservation consensus (Alpha Scientists In Reproductive Medicine, 2012). The gap between the competency and the benchmark values is the 'desirable range'.
- (4) It was the opinion of the expert panel that ovarian stimulation can have an impact on the overall treatment cycle but is less likely to have an impact on any single laboratory PI.
- (5) To apply the recommended values:
  - The time of oocyte retrieval relative to the time of the ovulation trigger is typically in the range of 34–38 h (mostly 36 h). Deviations from the locally established protocol should be documented and taken into consideration.
  - The timing of all observations should be made as recommended in the Istanbul consensus (Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology, 2011).
- (6) Individual clinics should decide whether it is more practical to subdivide their results into specific patient groups for KPI and PI determinations, based on their clinical practice. The indicator values presented here were derived relative to cycles that met the criteria for a 'reference population'. With the exception of indicators with specific qualifiers identified, these criteria were:
  - Female patients <40 years old;
  - own fresh oocytes;
  - ejaculated spermatozoa (fresh or frozen);
  - no PGD/PGS (PGT);
  - all insemination methods (i.e. routine IVF and ICSI).
- (7) It was the opinion of the panel that national and international registries should be encouraged to gather data that can be used for the derivation of KPI standard values.
- (8) Any selection procedure in addition to embryo morphology, for example PGD/PGS (PGT) or time-lapse imaging, is not expected to increase the cumulative pregnancy/LBR, but in the case of PGS it may result in reduced time to pregnancy/live birth.

### Indicators

The results of the discussions have been summarized for the majority of the indicators. The values for indicators have also been presented in

**Table II** RIs for identifying performance of the ART laboratory.

RI	Calculation	Benchmark value
Proportion of oocytes recovered (stimulated cycles)	$\frac{\text{no. oocytes retrieved}}{\text{no. follicles on day of trigger}} \times 100$	80–95% of follicles measured
Proportion of MII oocytes at ICSI	$\frac{\text{no. MII oocytes at ICSI}}{\text{no. COCs retrieved}} \times 100$	75–90%

MI, metaphase II; RI, reference indicators; COC, cumulus-oocyte complex.

**Table III** PIs for the ART laboratory.

PI	Calculation	Competency value (%)	Benchmark value (%)
Sperm motility post-preparation (for IVF and IUI)	$\frac{\text{progressively motile sperm}}{\text{all sperm counted}} \times 100$	90	≥95
IVF polyspermy rate	$\frac{\text{no. fertilized oocytes with > 2PN}}{\text{no. COCs inseminated}} \times 100$		<6
I PN rate (IVF)	$\frac{\text{no. I PN oocytes}}{\text{no. COCs inseminated}} \times 100$		<5
I PN rate (ICSI)	$\frac{\text{no. I PN oocytes}}{\text{no. MII oocytes injected}} \times 100$		<3
Good blastocyst development rate	$\frac{\text{no. good quality blastocysts on Day 5}}{\text{no. 2PN/2PB oocytes on Day 1}} \times 100$	≥30	≥40

PN, pronucleus; PI, performance indicator; PB, polar body.

Tables II–IV, but these should be read in association with the summary for each indicator.

#### Proportion of oocytes recovered

The proportion of oocytes recovered RI (Table II) is defined as the number of oocytes retrieved as a function of the number of ovarian follicles seen at ultrasound assessment. It is useful as a measure of whether the quantity of oocytes is maximized. The values are not influenced by laboratory practice, and so cannot be held to be laboratory PIs, but values outside the expected range could prompt an investigation of any changes in stimulation practice. Having this information is an important factor in troubleshooting. The expected range is 80–95% of follicles measured in stimulated cycles.

#### Proportion of MII oocytes at ICSI

The proportion of MII oocytes at ICSI RI (Table II) is defined as the proportion of oocytes that have nuclear maturity at the time of injection, and so acts as a proxy indication of the effectiveness of ovarian stimulation. It is not a laboratory PI, as values are not influenced by laboratory practice, but rather reflects factors that influence the competence of oocytes coming into the laboratory. The expected range is 75–90% at 40 ± 1 h post-trigger for all COCs retrieved. Values outside this range could prompt a review of any changes in ovarian stimulation, triggering or follicle aspiration practice, as changes in the proportion of MII oocytes could be a factor in changes in fertilization rates and/or embryo development. Instability in this value could indicate changes in the stimulation, resulting in a higher proportion of either immature or post-mature oocytes.

#### Notes:

- Since this value is expected to be stable, laboratories may choose ‘reporting by exception’—only reporting on it when it falls outside the expected range.
- It should be noted that nuclear maturity does not necessarily indicate cytoplasmic maturity of the oocyte (Sundstrom and Nilsson, 1988; Eppig, 1996; Coticchio et al., 2012).
- Clinics should consider whether they should sub-divide their assessment of this indicator based on patient demographics.
- Good communication between laboratory and clinic (cycle planning and cycle review) were cited as being vital to excellent outcomes in IVF programmes (Van Voorhis et al., 2010).

#### Semen analysis characteristics

**Sperm concentration, motility and vitality.** Unless semen analyses are performed employing analytical methods as per the ESHRE SIG Andrology (Björndahl et al., 2010) or WHO5 (World Health Organization, 2010) the results for sperm concentration, motility and vitality will be subject to unacceptably high uncertainty of measurement (Sanchez-Pozo et al., 2013; Björndahl et al., 2016).

**Sperm morphology.** Sperm morphology assessment is subjective and so is dependent on consistent training. Since the Tygerberg Strict Criteria cut-off of 4% normal forms was derived in relation to IVF success, it could be pertinent in differentiating between the need for IUI, IVF or ICSI (Mortimer and Menkveld, 2001; Menkveld, 2010). However, the current visual evaluation of 200 or 400 spermatozoa used in the vast majority of laboratories to assess ‘percentage normal forms’ has such a large uncertainty of

**Table IV** KPIs for the ART laboratory.

KPI	Calculation	Competency value (%)	Benchmark value (%)
ICSI damage rate	$\frac{\text{no. damaged or degenerated}}{\text{all oocytes injected}} \times 100$	≤10	≤5
ICSI normal fertilization rate	$\frac{\text{no. oocytes with 2PN and 2PB}}{\text{no. MII oocytes injected}} \times 100$	≥65	≥80
IVF normal fertilization rate	$\frac{\text{no. oocytes with 2PN and 2PB}}{\text{no. COCs inseminated}} \times 100$	≥60	≥75
Failed fertilization rate (IVF)	$\frac{\text{no. cycles with no evidence of fertilization}}{\text{no. of stimulated IVF cycles}} \times 100$		<5
Cleavage rate	$\frac{\text{no. cleaved embryos Day 2}}{\text{no. 2PN/2PB oocytes on Day 1}} \times 100$	≥95	≥99
Day 2 Embryo development rate	$\frac{\text{no. 4-cell embryos on Day 2}}{\text{no. normally fertilized oocytes}^a} \times 100$	≥50	≥80
Day 3 Embryo development rate	$\frac{\text{no. eight cell embryos on Day 3}}{\text{no. normally fertilized oocytes}^a} \times 100$	≥45	≥70
Blastocyst development rate	$\frac{\text{no. blastocysts Day 5}}{\text{no. normally fertilized oocytes}^a} \times 100$	≥40	≥60
Successful biopsy rate	$\frac{\text{no. biopsies with DNA detected}}{\text{no. biopsies performed}} \times 100$	≥90	≥95
Blastocyst cryosurvival rate	$\frac{\text{no. blastocysts appearing intact}}{\text{no. blastocysts warmed}} \times 100$	≥90	≥99
Implantation rate (cleavage-stage) <sup>b</sup>	$\frac{\text{no. sacs seen on ultrasound}^c}{\text{no. embryos transferred}} \times 100$	≥25	≥35
Implantation rate (blastocyst-stage) <sup>b</sup>	$\frac{\text{no. sacs seen on ultrasound}^c}{\text{no. blastocysts transferred}} \times 100$	≥35	≥60

<sup>a</sup>Defined as oocytes with 2PN and 2PB on Day 1.

<sup>b</sup>Based on total number of embryos transferred to *all* patients in the reference group, not just those for whom an implantation occurred.

<sup>c</sup>Definition reached after discussion, as some felt that no. fetal hearts / no. embryos transferred was a more meaningful indicator.

KPI, key performance indicator.

measurement that it cannot be considered a reliable predictor for IVF success/failure for individual men (Kvist and Björndahl, 2002; Björndahl et al.). Unless determined using a more robust methodology, sperm normal forms should not be used to direct ART treatment options.

#### Sperm motility post-preparation

The sperm motility post-preparation PI is defined as the proportion of progressively motile spermatozoa in the sperm preparation for insemination, and includes only fresh normozoospermic ejaculate specimens. Sperm motility after washing should be very high, and low values would indicate problems with the sperm preparation procedure. In this case, progressive motility is defined as spermatozoa that are moving with net space gain of the head, and so includes hyperactivated spermatozoa. Although there was excellent agreement among the survey responses, potential weaknesses of this PI include: possible poor reliability of sperm percentage motility data; non-robust classification of cases based on uncertainty of semen analysis data; variability in sperm preparation method used; and abnormal response of the sperm to the preparation method used. The reference values were competence 90% and benchmark ≥95%.

#### Notes:

- There is no sperm recovery rate KPI recommended because this is so heavily dependent on the processing method.

#### – Recommendations for IUI or IVF treatment:

It was the recommendation of the expert panel that decisions regarding a man's suitability or need for an appropriate ART treatment modality (IUI, IVF or ICSI) should be based on sperm numbers and motility assessments determined in a pre-treatment 'trial preparation'. The competency threshold value was agreed as at least 90% progressive motility post-wash with a benchmark of at least 95% progressive motility.

#### – Recommendations for ICSI treatment:

There was no cut-off value recommended for ICSI treatment, beyond the spermatozoa ideally being alive. In this case, the best evidence of vitality is motility, while the second-best evidence, in the absence of motility, is HOS-test positive (Björndahl et al.), although other methodologies exist.

#### ICSI damage rate

The ICSI Damage Rate KPI (Table IV) is defined as the proportion of oocytes that is damaged during the ICSI injection, or have degenerated by the time of fertilization assessment on Day 1. It is informative of gamete quality and/or operator skill, and excludes damage from oocyte stripping, which should be very rare. The results can be skewed by the patient mix or the stimulation protocols used, so all cycles should be included, to reduce the relative impact of these variables. The Alpha survey gave similar median and mode values for each level. These values were agreed by the expert panel in relation to those

recommended by the Association of Clinical Embryologists (Hughes and Association of Clinical Embryologists, 2012). The reference values for ICSI damage rate are: competence  $\leq 10\%$ ; benchmark  $\leq 5\%$ .

#### ICSI normal fertilization rate

The ICSI Normal Fertilization Rate KPI (Table IV) is defined as the number of fertilized oocytes on Day 1 (presence of 2PN and 2PB assessed at  $17 \pm 1$  h post-injection) (Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology, 2011), as a function of all MII oocytes injected. This is a common, broad, effective indicator of good laboratory practice, as it is informative of gamete quality and/or operator skill. This KPI includes ejaculated spermatozoa only (fresh or frozen) as results may be lower with surgically retrieved spermatozoa, and excludes *in vitro* matured oocytes, as well as thawed/warmed oocytes (this was dealt with in the cryopreservation consensus) (Alpha Scientists In Reproductive Medicine, 2012). Reference values were agreed by the expert panel: competence  $\geq 65\%$ ; benchmark  $\geq 80\%$ .

##### Notes:

- *Total ICSI failed fertilization rate:* It was the consensus of the expert panel that complete failure to achieve normal fertilization in an ICSI cycle did not need to be a PI, but should be reported by exception, meaning that ‘every case’ should be investigated. This includes only stimulated cycles, as natural cycles are expected to have only one oocyte.
- *Poor ICSI fertilization rate:* Although opinion regarding the expected incidence of cycles with ICSI fertilization rates  $<25\%$  was sought in the Alpha survey, the consensus of the expert panel was to exclude this indicator, as it did not add to the information already collected.
- *Giant oocytes should not be injected due to published evidence of chromosomal abnormality* (Balakier et al., 2002; Rosenbusch et al., 2002; Lehner et al., 2015).
- *Regarding smooth endoplasmic reticulum (SER) positive oocytes,* more recent publications of outcomes suggest that the Istanbul Consensus recommendation not to inject/inseminate these oocytes may need to be revisited (Mateizel et al., 2013). It was the opinion of the expert panel that in the meantime, the decision to inject SER positive oocytes should be reviewed by the clinical team on a case-by-case basis. Follow-up of results, including pregnancy outcome and babies born, in case of insemination and transfer of the resulting embryos should be performed.

#### Normal IVF fertilization rate

The Normal IVF Fertilization Rate KPI (Table IV) is defined as the number of fertilized oocytes on Day 1 (presence of 2PN and 2PB assessed at  $17 \pm 1$  h post-insemination (Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology, 2011)), as a function of all COCs inseminated. IVF normal fertilization rate is an important indicator of laboratory performance, as it relies upon effective gamete handling and culture, and so is a measure of the whole IVF system. The reference values are: competency  $\geq 60\%$ ; benchmark  $\geq 75\%$ . It should be noted that the benchmark value was determined based on a rounding of the product of the benchmark rates for MII oocytes (90%) and fertilization rate of MII oocytes (80%).

#### IVF polyspermy rate

IVF Polyspermy Rate PI (Table III) is defined as the proportion of inseminated oocytes with more than two pronuclei on Day 1 ( $17 \pm 1$  h

post-insemination) and is needed to provide the information to interpret any observed variations in the normal fertilization rate. It was the consensus that polyspermy rate should be  $<6\%$ . Observed values above this rate should be reported and investigated.

#### IPN rate following IVF or ICSI

The IPN Rate Following IVF or ICSI PI (Table III) is defined as the proportion of inseminated oocytes with one pronucleus on Day 1 ( $17 \pm 1$  h post-insemination). It can provide a marker of a problem in gamete handling or culture conditions and so should be low under normal conditions. It was the consensus that the IPN rate should be  $<5\%$  for IVF cycles, and  $<3\%$  for ICSI cycles. The difference between IVF and ICSI is related to the pre-selection of oocytes prior to ICSI injection. Observed values above this rate should be reported and investigated.

#### Failed fertilization rate (IVF cycles)

The failed fertilization rate (IVF cycles) KPI (Table IV) is defined as the proportion of IVF cycles (excludes ICSI cycles) with no evidence of fertilization (i.e. 0 oocytes with  $\geq 2$ PN) on Day 1 ( $17 \pm 1$  h post-insemination). It can provide a marker of a problem in gamete quality (sperm function, oocyte activation, gamete receptors), sperm processing, or in the number of spermatozoa used for insemination. It should be low, under normal conditions. Based on the Alpha survey, and on the values recommended by the Association of Clinical Embryologists (Hughes and Association of Clinical Embryologists, 2012), it was the consensus that the IVF failed fertilization rate should be  $<5\%$  for stimulated cycles. Observed values above this rate should be reported and investigated.

#### Zygote grade (IVF cycles)

Zygote grade is an evaluation of the quality of the fertilized oocyte, conducted  $17 \pm 1$  h post-insemination. It was the consensus that there were not enough data to recommend indicator values for this measure. In the discussion related to micronuclei, it was agreed that while micronuclei are abnormal, there is no evidence to confirm the identity of these dynamic manifestations.

#### Early cleavage rate

Early cleavage rate is the proportion of fertilized oocytes that have undergone the first round of cleavage by  $26 \pm 1$  h post-insemination by ICSI or  $28 \pm 1$  h post-insemination by IVF. There is evidence that early cleavage, together with other factors, can be used as embryo selection method as it has been correlated with implantation rate (Shoukir et al., 1997; Lundqvist et al., 2001; Balaban and Urman, 2003; Ciray et al., 2005; Brezinova et al., 2009). However, as it is not routinely calculated, it was the consensus that while this indicator can be useful for troubleshooting purposes, there were no recommendations for expected values.

#### Cleavage rate

The cleavage rate KPI (Table IV) is defined as the proportion of zygotes that cleave to become embryos on Day 2 ( $44 \pm 1$  h post-insemination) (Alpha Scientists in Reproductive Medicine and ESHRE Special Interest Group of Embryology, 2011). It provides an indication of the ability of the culture system to support cleavage of fertilized oocytes (i.e. with occurrence of cellular division), and of embryo viability, so a low cleavage rate could be a warning that the culture system



has been impacted by an extrinsic factor. Furthermore, there is evidence that the presence of at least one non-cleaved embryo is predictive of reduced embryo quality for the remaining cohort (Machtinger et al., 2015). The reference values are: competency > 95%; benchmark > 99%.

#### Embryo development rate

The KPI embryo development rate (Table IV) is defined as the proportion of cleaved embryos at the 4-cell stage on Day 2 ( $44 \pm 1$  h post-insemination) or at the 8-cell stage on Day 3 ( $68 \pm 1$  h post-insemination) per normally fertilized oocyte. This is an indicator of the ability of the culture system to support cleavage according to the expected stages, as well as providing an indication of the viability and quality of embryos. This KPI just considers the developmental stage of the embryo, regardless of grade, because developmental stage has been reported to give the highest degree of predictive power, and has the advantage of being an objective measure. It was acknowledged that not all clinics consider the specific cell stages defined here, and that the culture system used can affect the kinetics of embryo development.

The reference values, based on the median and maximum results from the participants' laboratories are: Day 2, competency  $\geq 50\%$  and benchmark  $\geq 80\%$ ; and Day 3, competency  $\geq 45\%$  and benchmark  $\geq 70\%$ .

#### Notes:

- It was the decision of the expert panel not to consider Day 4 embryo development rate.
- The Embryo Development Rate KPI is most relevant to clinics that transfer embryos on Day 2 or 3.

#### Embryo and blastocyst utilization rates

These potential indicators were defined as the number of embryos (or blastocysts) suitable for transfer or cryopreservation as a function of the number of normally fertilized (2PN) oocytes observed on Day 1. Although it was agreed that these indicators could be of value for internal laboratory comparison, the consensus opinion was that because there are so many differences in laboratory and clinical practice, it was not practical to suggest any values for these indicators.

#### Blastocyst development rate

The KPI blastocyst development rate (Table IV) is defined as the proportion of blastocysts observed at  $116 \pm 2$  h post-insemination as a function of the number of normally fertilized oocytes. It estimates the ability of the culture system to support blastocyst formation from fertilized oocytes (i.e. with formation of ICM, TE and a blastocoele cavity), and provides an indication of embryo viability. It should be noted that this definition only considers blastocyst formation, with no consideration of blastocyst-stage or blastocyst quality.

The reference values are: competency  $\geq 40\%$ ; benchmark  $\geq 60\%$  (Day 5). A possible additional PI might be the development of an additional 10–15% blastocysts by  $140 \pm 2$  h post-insemination (i.e. by Day 6).

#### Good blastocyst development rate

The good blastocyst development rate PI (Table III) is defined as the number of good quality blastocysts as a function of the number of normally fertilized oocytes. Blastocyst quality is as defined in the Istanbul consensus (Alpha Scientists in Reproductive Medicine and ESHRE

Special Interest Group of Embryology, 2011). This indicator estimates the ability of the culture system to support the formation of high-grade blastocysts from fertilized oocytes (i.e. with formation of ICM, TE and a blastocoele cavity), and an indication of embryo viability. The reference values are: competency  $\geq 30\%$ ; benchmark  $\geq 40\%$  (Day 5). A possible additional PI might be the development of an additional 5–15% blastocysts by  $140 \pm 2$  h post-insemination (i.e. by Day 6) (depending upon the culture system).

#### Proportion of good blastocysts

This PI is defined as the proportion of blastocysts with a grade of 'good' or higher. There was no discussion of reference values for this indicator, as they can be inferred from the preceding two Indicators.

#### Day 5 embryo transfer rate

This PI is defined as the proportion of cycles with at least one utilizable blastocyst on Day 5 relative to the presence of at least one 2PN oocyte on Day 1, to allow for the inclusion of cycles in which the decision has been made to cryopreserve all embryos. This indicator reflects the efficiency of the whole culture system, but is only relevant for those clinics that have a blanket strategy of Day 5 transfers. It was the consensus opinion that as there are too many clinic-specific variables, including different embryo transfer policies in different centers, clinics should develop their own expectations for this indicator, depending on when the decision to go to Day 5 transfer is made (e.g. Day 0 vs Day 3).

#### Implantation rate

For this consensus, the implantation rate KPI (Table IV) is defined as the number of gestational sacs divided by the total number of embryos transferred, irrespective of whether a pregnancy was established (Zegers-Hochschild et al., 2009). There was some dissent about the use of sacs, rather than fetal hearts—which was felt to be a more meaningful indicator of implantation rate—however sacs are used by most national/international registries. Following discussion, it was agreed to use sacs for the purpose of this consensus, but that the use of fetal heart as the numerator should be revisited in the future.

Implantation rate provides an indication of the overall performance of the laboratory so an overall low implantation rate is a serious sign of a systemic problem. Values would be expected to be lower for Days 2 and 3 transfers than for blastocyst transfers. In addition, results will be influenced by clinical factors (e.g. uterine receptivity) and the existence of different policies in different centers for deciding the day of embryo transfer.

#### Reference values for implantation rates

*Planned transfers of cleavage-stage embryos (Day 2 or 3):* competency  $\geq 25\%$ , benchmark  $\geq 35\%$ .

*Blastocyst transfers:* competency  $\geq 35\%$ , benchmark  $\geq 60\%$  (the panel was divided between 55% and 60%, but agreed that 60% was an aspirational goal).

*Note:* These values could be affected if there are a large number of patients in the cohort who have had a large number of previous, unsuccessful cycles or significant clinical adverse factors. In addition, individual clinics may wish to further stratify their results based on patient age groups.

### Live birth rate

It was the consensus that there are too many other variables to support the use of LBR as a laboratory indicator for either cleavage-stage or blastocyst-stage embryo transfers.

### Successful biopsy rate

The KPI successful biopsy rate (Table IV) is defined as the proportion of biopsied and tubed/fixed samples where DNA is detected. It is a measure of the ability of embryologists to transfer the biopsied samples to test tubes, as proven by positive DNA amplification.

Based on data from the surveys and the PGD Consortium, which reported a 91% diagnosis rate in 254,820 biopsies (De Rycke et al., 2015), the reference values were: competency  $\geq 90\%$ ; benchmark  $\geq 95\%$  (Table IV).

### No biopsy

This indicator was defined as the proportion of treatment cycles that had no embryos available for biopsy. It was the consensus that not having good quality blastocysts for biopsy is not an indication of the quality of the PGD/PGS service. It is an expression of blastocyst development rates, and so no further values were developed for this indicator.

### No embryos found on warming/degenerated on warming

Nowadays, the majority of PGT cycles are based on blastocyst biopsy and cryopreservation. It was agreed that no embryos found on warming/degenerated on warming should be a KPI, as it provides a reflection of operator skills, and/or the device used. However, since not finding the embryo is a rare event, the panel was unable to estimate a competency value, as the value would be very low.

Owing to greater experience with blastocyst vitrification, the rate of degeneration on warming should now be lower than that estimated in the previous cryopreservation consensus (Alpha Scientists In Reproductive Medicine, 2012). Similarly, it was the consensus that re-expansion does not differ between (warmed) biopsied and non-biopsied blastocysts.

Notwithstanding some device differences, the reference rates for blastocyst cryosurvival could now reasonably be expected to be: competency  $\geq 90\%$ , benchmark  $\geq 99\%$  (Table IV).

### Implantation rate of biopsied embryos

It was the consensus that the implantation rate for blastocysts biopsied for PGS should exceed that expected for the age-matched patient population in the same clinic. From the literature, a meta-analysis reported an improvement of 30% sustained implantation rate after the transfer of PGS-selected blastocysts relative to controls (Scott et al., 2013; Dahdouh et al., 2015).

### Time-lapse imaging

Despite the increasing number of IVF cycles incorporating time-lapse imaging for embryo assessment and selection, the panel considered it premature to propose time-lapse related PIs for the IVF laboratory, owing to the limited and varied data associating precise timings of human embryo development with viability or good laboratory practice. It was accepted, however, that because of the detailed morphological and kinetic information collected per embryo, time-lapse imaging may

prove to be a future early warning tool for compromised culture conditions, providing a (intra-)laboratory PI, specifically if a change in mean timings for embryos to reach developmental milestones may be detected more readily and rapidly than with standard assessment methods.

A time-lapse assay of mouse embryo development linked specific morphokinetic changes to toxicity of mineral oil. This demonstrates the sensitivity of mouse embryo cleavage timings to the quality of the culture environment, and the potential value of time-lapse in detecting such changes (Wolff et al., 2013). To date, no studies of this type have been performed on human embryos.

The current recommendation, therefore, was that clinics may wish to establish time-lapse KPIs and benchmarks based on their own experience. It was suggested that the frequency of anomalous cleavage events, such as trichotomous mitosis (direct cleavage to three cells), which is known to be relatively common (6–8%) and associated with reduced implantation potential, could be monitored for future use as a KPI (Rubio et al., 2012; Athayde Wirka et al., 2014). In addition, it was considered that clinics using validated time-lapse algorithms for embryo selection may develop benchmarks associated with the proportion of embryos which ranked highest using morphokinetic modeling. Large-scale population studies are, however, required in order to identify which, if any, morphokinetic markers can be developed into universal and useful laboratory KPIs. It is also important to point out that morphokinetic embryo performance is believed to be highly associated with clinical and laboratory practices, and may, therefore, be difficult to compare between laboratories.

## Acknowledgements

We thank the respondents to the questionnaires for their valuable contribution to this consensus. This manuscript was prepared by N.V. and Sharon Mortimer.

## Authors' roles

The document resulted from a joint effort of the ESHRE Special Interest Group of Embryology and Alpha Scientists in Reproductive Medicine. All members of the expert panel contributed equally in writing the paper and critical reading.

## Funding

The consensus meeting and writing of the paper was supported by funds from ESHRE and Alpha. Alpha gratefully acknowledges the following organizations for their financial support, through the provision of unrestricted educational grants: Global Fertility Alliance, Merck, Origio and Vitrolife.

## Conflict of interest

There were no conflicts of interest disclosed.

## References

- Ahlstrom A, Westin C, Reismer E, Wikland M, Hardarson T. Trophoctoderm morphology: an important parameter for predicting live birth after single blastocyst transfer. *Hum Reprod* 2011;**26**:3289–3296.
- Ahlstrom A, Westin C, Wikland M, Hardarson T. Prediction of live birth in frozen-thawed single blastocyst transfer cycles by pre-freeze and post-thaw morphology. *Hum Reprod* 2013;**28**:1199–1209.
- Alpha Scientists In Reproductive Medicine. The Alpha consensus meeting on cryopreservation key performance indicators and benchmarks: proceedings of an expert meeting. *Reprod Biomed Online* 2012;**25**:146–167.
- Alpha Scientists in Reproductive Medicine, ESHRE Special Interest Group of Embryology. The Istanbul consensus workshop on embryo assessment: proceedings of an expert meeting. *Hum Reprod* 2011;**26**:1270–1283. and *Reprod Biomed Online* 2011;**22**:1632–46 (simultaneous publication).
- Arce JC, Andersen AN, Fernandez-Sanchez M, Visnova H, Bosch E, Garcia-Velasco JA, Barri P, de Sutter P, Klein BM, Fauser BC. Ovarian response to recombinant human follicle-stimulating hormone: a randomized, antimullerian hormone-stratified, dose-response trial in women undergoing in vitro fertilization/intracytoplasmic sperm injection. *Fertil Steril* 2014;**102**:1633–1640.e1635.
- Athayde Wirka K, Chen AA, Conaghan J, Ivani G, Gvakharia M, Behr B, Suraj V, Tan L, Shen S. Atypical embryo phenotypes identified by time-lapse microscopy: high prevalence and association with embryo development. *Fertil Steril* 2014;**101**:1637–1648.e1631–1635.
- Baart EB, Martini E, Eijkemans MJ, Van Opstal D, Beckers NG, Verhoeff A, Macklon NS, Fauser BC. Milder ovarian stimulation for in-vitro fertilization reduces aneuploidy in the human preimplantation embryo: a randomized controlled trial. *Hum Reprod* 2007;**22**:980–988.
- Balaban B, Urman B. Embryo culture as a diagnostic tool. *Reprod Biomed Online* 2003;**7**:671–682.
- Balaban B, Urman B. Effect of oocyte morphology on embryo development and implantation. *Reprod Biomed Online* 2006;**12**:608–615.
- Balakier H, Bouman D, Sojecki A, Librach C, Squire JA. Morphological and cytogenetic analysis of human giant oocytes and giant embryos. *Hum Reprod* 2002;**17**:2394–2401.
- Björndahl L. What is normal semen quality? On the use and abuse of reference limits for the interpretation of semen analysis results. *Hum Fertil (Camb)* 2011;**14**:179–186.
- Björndahl L, Barratt CL, Mortimer D, Jouannet P. 'How to count sperm properly': checklist for acceptability of studies based on human semen analysis. *Hum Reprod* 2016;**31**:227–232.
- Björndahl L, Mortimer D, Barratt CL, Castilla JA, Menkveld R, Kvist U, Alvarez JG, Haugen TB. *A Practical Guide to Basic Laboratory Andrology*. Cambridge, UK: Cambridge University Press, 2010.
- Braga DP, Setti AS, Figueira Rde C, Machado RB, Iaconelli AJr, Borges E Jr. Influence of oocyte dysmorphisms on blastocyst formation and quality. *Fertil Steril* 2013;**100**:748–754.
- Brezinova J, Oborna I, Svobodova M, Fingerova H. Evaluation of day one embryo quality and IVF outcome—a comparison of two scoring systems. *Reprod Biol Endocrinol* 2009;**7**:9.
- Ciray HN, Karagenc L, Ulug U, Bener F, Bahceci M. Use of both early cleavage and day 2 mononucleation to predict embryos with high implantation potential in intracytoplasmic sperm injection cycles. *Fertil Steril* 2005;**84**:1411–1416.
- Cobo A, de los Santos MJ, Castello D, Gamiz P, Campos P, Remohi J. Outcomes of vitrified early cleavage-stage and blastocyst-stage embryos in a cryopreservation program: evaluation of 3,150 warming cycles. *Fertil Steril* 2012;**98**:1138–1146.e1131.
- Coetzee K, Kruger TF, Lombard CJ. Predictive value of normal sperm morphology: a structured literature review. *Hum Reprod Update* 1998;**4**:73–82.
- Cooper TG, Noonan E, von Eckardstein S, Auger J, Baker HW, Behre HM, Haugen TB, Kruger T, Wang C, Mbizvo MT et al. World Health Organization reference values for human semen characteristics. *Hum Reprod Update* 2010;**16**:231–245.
- Coticchio G, Dal-Canto M, Guglielmo MC, Mignini-Renzini M, Fadini R. Human oocyte maturation in vitro. *Int J Dev Biol* 2012;**56**:909–918.
- Dahdouh EM, Balayla J, Garcia-Velasco JA. Comprehensive chromosome screening improves embryo selection: a meta-analysis. *Fertil Steril* 2015;**104**:1503–1512.
- Davison SL, Bell R, Donath S, Montalto JG, Davis SR. Androgen levels in adult females: changes with age, menopause, and oophorectomy. *J Clin Endocrinol Metab* 2005;**90**:3847–3853.
- de los Santos MJ, Arroyo G, Busquet A, Calderon G, Cuadros J, Hurtado de Mendoza MV, Moragas M, Herrero R, Ortiz A, Pons C et al. A multicenter prospective study to assess the effect of early cleavage on embryo quality, implantation, and live-birth rate. *Fertil Steril* 2014;**101**:981–987.
- De Rycke M, Bvela F, Goossens V, Moutou C, SenGupta SB, Traeger-Synodinos J, Coonen E. ESHRE PGD Consortium data collection XIII: cycles from January to December 2010 with pregnancy follow-up to October 2011. *Hum Reprod* 2015;**30**:1763–1789.
- Delvigne A. Symposium: Update on prediction and management of OHSS. Epidemiology of OHSS. *Reprod Biomed Online* 2009;**19**:8–13.
- Denomme MM, Mann MR. Genomic imprints as a model for the analysis of epigenetic stability during assisted reproductive technologies. *Reproduction* 2012;**144**:393–409.
- Ebner T, Montag M, Oocyte Activation Study G, Montag M, Van der Ven K, Van der Ven H, Ebner T, Shebl O, Oppelt P, Hirschhain J et al. Live birth after artificial oocyte activation using a ready-to-use ionophore: a prospective multicentre study. *Reprod Biomed Online* 2015;**30**:359–365.
- Ebner T, Moser M, Sommergruber M, Tews G. Selection based on morphological assessment of oocytes and embryos at different stages of preimplantation development: a review. *Hum Reprod Update* 2003;**9**:251–262.
- Ectors FJ, Vanderzwalmen P, Van Hoeck J, Nijs M, Verhaegen G, Delvigne A, Schoysman R, Leroy F. Relationship of human follicular diameter with oocyte fertilization and development after in-vitro fertilization or intracytoplasmic sperm injection. *Hum Reprod* 1997;**12**:2002–2005.
- Eppig JJ. Coordination of nuclear and cytoplasmic oocyte maturation in eutherian mammals. *Reprod Fertil Dev* 1996;**8**:485–489.
- ESHRE Guideline Group on Good Practice in IVF Labs, De los Santos MJ, Apter S, Coticchio G, Debrock S, Lundin K, Plancha CE, Prados F, Rienzi L, Verheyen G et al. Revised guidelines for good practice in IVF laboratories (2015). *Hum Reprod* 2016;**31**:685–686.
- Fancovits P, Tothne ZG, Murber A, Rigo J Jr, Urbancsek J. Importance of cytoplasmic granularity of human oocytes in in vitro fertilization treatments. *Acta Biol Hung* 2012;**63**:189–201.
- Fauque P, Audureau E, Leandri R, Delaroche L, Assouline S, Epelboin S, Jouannet P, Patrat C. Is the nuclear status of an embryo an independent factor to predict its ability to develop to term? *Fertil Steril* 2013;**99**:1299–1304.e1293.
- Fauque P, Jouannet P, Lesaffre C, Ripoché MA, Dandolo L, Vaiman D, Jammes H. Assisted reproductive technology affects developmental kinetics, H19 Imprinting Control Region methylation and H19 gene expression in individual mouse embryos. *BMC Dev Biol* 2007;**7**:116.
- Ferraretti AP, Gianaroli L, Magli MC, D'Angelo A, Farfalli V, Montanaro N. Exogenous luteinizing hormone in controlled ovarian hyperstimulation for assisted reproduction techniques. *Fertil Steril* 2004;**82**:1521–1526.
- Flaherty SP, Payne D, Matthews CD. Fertilization failures and abnormal fertilization after intracytoplasmic sperm injection. *Hum Reprod* 1998;**13**:155–164.
- Garrido N, Bellver J, Remohi J, Simon C, Pellicer A. Cumulative live-birth rates per total number of embryos needed to reach newborn in

- consecutive in vitro fertilization (IVF) cycles: a new approach to measuring the likelihood of IVF success. *Fertil Steril* 2011;**96**:40–46.
- Glujovsky D, Farquhar C, Quinteiro Retamar AM, Alvarez Sedo CR, Blake D. Cleavage stage versus blastocyst stage embryo transfer in assisted reproductive technology. *Cochrane Database Syst Rev* 2016:CD002118.
- Haaf T, Hahn A, Lambrecht A, Grossmann B, Schwaab E, Khanaga O, Hahn T, Tresch A, Schorsch M. A high oocyte yield for intracytoplasmic sperm injection treatment is associated with an increased chromosome error rate. *Fertil Steril* 2009;**91**:733–738.
- Heindryckx B, Van der Elst J, De Sutter P, Dhont M. Treatment option for sperm- or oocyte-related fertilization failure: assisted oocyte activation following diagnostic heterologous ICSI. *Hum Reprod* 2005;**20**:2237–2241.
- Hughes C. Association of Clinical Embryologists. Association of clinical embryologists—guidelines on good practice in clinical embryology laboratories 2012. *Hum Fertil (Camb)* 2012;**15**:174–189.
- Hugues JN, Theron-Gerard L, Coussieu C, Pasquier M, Dewailly D, Cedrin-Durnerin I. Assessment of theca cell function prior to controlled ovarian stimulation: the predictive value of serum basal/stimulated steroid levels. *Hum Reprod* 2010;**25**:228–234.
- Humaidan P, Bungum M, Bungum L, Yding Andersen C. Effects of recombinant LH supplementation in women undergoing assisted reproduction with GnRH agonist down-regulation and stimulation with recombinant FSH: an opening study. *Reprod Biomed Online* 2004;**8**:635–643.
- Jeppesen JV, Kristensen SG, Nielsen ME, Humaidan P, Dal Canto M, Fadini R, Schmidt KT, Ernst E, Yding Andersen C. LH-receptor gene expression in human granulosa and cumulus cells from antral and preovulatory follicles. *J Clin Endocrinol Metab* 2012;**97**:E1524–E1531.
- Joergensen MW, Labouriau R, Hindkjaer J, Stougaard M, Kolevraa S, Bolund L, Agerholm IE, Sunde L. The parental origin correlates with the karyotype of human embryos developing from tripronuclear zygotes. *Clin Exp Reprod Med* 2015;**42**:14–21.
- Kai Y, Iwata K, Iba Y, Mio Y. Diagnosis of abnormal human fertilization status based on pronuclear origin and/or centrosome number. *J Assist Reprod Genet* 2015;**32**:1589–1595.
- Kashir J, Heindryckx B, Jones C, De Sutter P, Parrington J, Coward K. Oocyte activation, phospholipase C zeta and human infertility. *Hum Reprod Update* 2010;**16**:690–703.
- Kohn LT, Corrigan JM, Donaldson MS. *To Err is Human: Building a Safer Health System*. Washington, DC: National Academies Press, 2000.
- Kruger TF, Acosta AA, Simmons KF, Swanson RJ, Matta JF, Oehninger S. Predictive value of abnormal sperm morphology in in vitro fertilization. *Fertil Steril* 1988;**49**:112–117.
- Kvist U, Bjorn Dahl L. *Manual on Basic Semen Analysis*. Oxford, UK: Oxford University Press, 2002.
- Lehner A, Kaszas Z, Murber A, Rigo J Jr, Urbancsek J, Fancsovsits P. Giant oocytes in human in vitro fertilization treatments. *Arch Gynecol Obstet* 2015;**292**:697–703.
- Lemmen JG, Rodriguez NM, Andreasen LD, Loft A, Ziebe S. The total pregnancy potential per oocyte aspiration after assisted reproduction—in how many cycles are biologically competent oocytes available? *J Assist Reprod Genet* 2016;**33**:849–854.
- Levron J, Munne S, Willadsen S, Rosenwaks Z, Cohen J. Male and female genomes associated in a single pronucleus in human zygotes. *Biol Reprod* 1995;**52**:653–657.
- Liu J, Nagy Z, Joris H, Tournaye H, Smits J, Camus M, Devroey P, Van Steirteghem A. Analysis of 76 total fertilization failure cycles out of 2732 intracytoplasmic sperm injection cycles. *Hum Reprod* 1995;**10**:2630–2636.
- Lopes AS, Frederickx V, Van Kerkhoven G, Campo R, Puttemans P, Gordts S. Survival, re-expansion and cell survival of human blastocysts following vitrification and warming using two vitrification systems. *J Assist Reprod Genet* 2015;**32**:83–90.
- Lundqvist M, Johansson U, Lundqvist O, Milton K, Westin C, Simberg N. Does pronuclear morphology and/or early cleavage rate predict embryo implantation potential? *Reprod Biomed Online* 2001;**2**:12–16.
- Machtinger R, Bormann CL, Ginsburg ES, Racowsky C. Is the presence of a non-cleaved embryo on day 3 associated with poorer quality of the remaining embryos in the cohort? *J Assist Reprod Genet* 2015;**32**:677–683.
- Mahutte NG, Arici A. Failed fertilization: is it predictable? *Curr Opin Obstet Gynecol* 2003;**15**:211–218.
- Mateizel I, Van Landuyt L, Tournaye H, Verheyen G. Deliveries of normal healthy babies from embryos originating from oocytes showing the presence of smooth endoplasmic reticulum aggregates. *Hum Reprod* 2013;**28**:2111–2117.
- Mateo S, Parriego M, Boada M, Vidal F, Coroleu B, Veiga A. In vitro development and chromosome constitution of embryos derived from mono-pronucleated zygotes after intracytoplasmic sperm injection. *Fertil Steril* 2013;**99**:897–902.e891.
- Mayer JF, Jones EL, Dowling-Lacey D, Nehchiri F, Muasher SJ, Gibbons WE, Oehninger SC. Total quality improvement in the IVF laboratory: choosing indicators of quality. *Reprod Biomed Online* 2003;**7**:695–699.
- Menkveld R. Clinical significance of the low normal sperm morphology value as proposed in the fifth edition of the WHO Laboratory Manual for the Examination and Processing of Human Semen. *Asian J Androl* 2010;**12**:47–58.
- Menkveld R, Holleboom CA, Rhemrev JP. Measurement and significance of sperm morphology. *Asian J Androl* 2011;**13**:59–68.
- Meseguer M, Herrero J, Tejera A, Hilligsoe KM, Ramsing NB, Remohi J. The use of morphokinetics as a predictor of embryo implantation. *Hum Reprod* 2011;**26**:2658–2671.
- Montag M, van der Ven H, German Pronuclear Morphology Study Group. Evaluation of pronuclear morphology as the only selection criterion for further embryo culture and transfer: results of a prospective multicentre study. *Hum Reprod* 2001;**16**:2384–2389.
- Mortimer D, Menkveld R. Sperm morphology assessment—historical perspectives and current opinions. *J Androl* 2001;**22**:192–205.
- Mortimer ST, Mortimer D. *Quality and Risk Management in the IVF Laboratory*. Cambridge, UK: Cambridge University Press, 2015.
- Nagy ZP, Janssenswillen C, Janssens R, De Vos A, Staessen C, Van de Velde H, Van Steirteghem AC. Timing of oocyte activation, pronucleus formation and cleavage in humans after intracytoplasmic sperm injection (ICSI) with testicular spermatozoa and after ICSI or in-vitro fertilization on sibling oocytes with ejaculated spermatozoa. *Hum Reprod* 1998;**13**:1606–1612.
- Nagy ZP, Jones-Colon S, Roos P, Botros L, Greco E, Dasig J, Behr B. Metabolomic assessment of oocyte viability. *Reprod Biomed Online* 2009;**18**:219–225.
- Nagy ZP, Liu J, Joris H, Verheyen G, Tournaye H, Camus M, Derde MC, Devroey P, Van Steirteghem AC. The result of intracytoplasmic sperm injection is not related to any of the three basic sperm parameters. *Hum Reprod* 1995;**10**:1123–1129.
- Nel-Themaat L, Nagy ZP. A review of the promises and pitfalls of oocyte and embryo metabolomics. *Placenta* 2011;**32**:S257–S263.
- Nogueira D, Friedler S, Schachter M, Raziel A, Ron-El R, Smits J. Oocyte maturity and preimplantation development in relation to follicle diameter in gonadotropin-releasing hormone agonist or antagonist treatments. *Fertil Steril* 2006;**85**:578–583.
- Paternot G, Wetzels AM, Thonon F, Vansteenbrugge A, Willemen D, Devroey J, Debrock S, D'Hooghe TM, Spiessens C. Intra- and interobserver analysis in the morphological assessment of early stage embryos during an IVF procedure: a multicentre study. *Reprod Biol Endocrinol* 2011;**9**:127.
- Patrizio P, Fragouli E, Bianchi V, Borini A, Wells D. Molecular methods for selection of the ideal oocyte. *Reprod Biomed Online* 2007;**15**:346–353.



- Patrizio P, Sakkas D. From oocyte to baby: a clinical evaluation of the biological efficiency of in vitro fertilization. *Fertil Steril* 2009;**91**:1061–1066.
- Piltonen T, Koivunen R, Ruokonen A, Tapanainen JS. Ovarian age-related responsiveness to human chorionic gonadotropin. *J Clin Endocrinol Metab* 2003;**88**:3327–3332.
- Rhenman A, Berglund L, Brodin T, Olovsson M, Milton K, Hadziosmanovic N, Holte J. Which set of embryo variables is most predictive for live birth? A prospective study in 6252 single embryo transfers to construct an embryo score for the ranking and selection of embryos. *Hum Reprod* 2015;**30**:28–36.
- Rienzi L, Vajta G, Ubaldi F. Predictive value of oocyte morphology in human IVF: a systematic review of the literature. *Hum Reprod Update* 2011;**17**:34–45.
- Rizk B. Symposium: Update on prediction and management of OHSS. Genetics of ovarian hyperstimulation syndrome. *Reprod Biomed Online* 2009;**19**:14–27.
- Rosenbusch B, Schneider M, Glaser B, Brucker C. Cytogenetic analysis of giant oocytes and zygotes to assess their relevance for the development of digynic triploidy. *Hum Reprod* 2002;**17**:2388–2393.
- Rubino P, Viganò P, Luddi A, Piomboni P. The ICSI procedure from past to future: a systematic review of the more controversial aspects. *Hum Reprod Update* 2016;**22**:194–227.
- Rubio I, Kuhlmann R, Agerholm I, Kirk J, Herrero J, Escriba MJ, Bellver J, Meseguer M. Limited implantation success of direct-cleaved human zygotes: a time-lapse study. *Fertil Steril* 2012;**98**:1458–1463.
- Russell DL, Salustri A. Extracellular matrix of the cumulus-oocyte complex. *Semin Reprod Med* 2006;**24**:217–227.
- Saenz-de-Juano MD, Billooye K, Smits J, Anckaert E. The loss of imprinted DNA methylation in mouse blastocysts is inflicted to a similar extent by in vitro follicle culture and ovulation induction. *Mol Hum Reprod* 2016;**22**:427–441.
- Salinas M, Lopez-Garrigos M, Gutierrez M, Lugo J, Sirvent JV, Uris J. Achieving continuous improvement in laboratory organization through performance measurements: a seven-year experience. *Clin Chem Lab Med* 2010;**48**:57–61.
- Sanchez-Pozo MC, Mendiola J, Serrano M, Mozas J, Bjorndahl L, Menkveld R, Lewis SE, Mortimer D, Jorgensen N, Barratt CL *et al*. Proposal of guidelines for the appraisal of SEMen QUALity studies (SEMQUA). *Hum Reprod* 2013;**28**:10–21.
- Scott RT, Hofmann GE, Muasher SJ, Acosta AA, Kreiner DK, Rosenwaks Z. Correlation of follicular diameter with oocyte recovery and maturity at the time of transvaginal follicular aspiration. *J In Vitro Fert Embryo Transf* 1989;**6**:73–75.
- Scott RT Jr, Upham KM, Forman EJ, Hong KH, Scott KL, Taylor D, Tao X, Treff NR. Blastocyst biopsy with comprehensive chromosome screening and fresh embryo transfer significantly increases in vitro fertilization implantation and delivery rates: a randomized controlled trial. *Fertil Steril* 2013;**100**:697–703.
- Shahangian S, Snyder SR. Laboratory medicine quality indicators: a review of the literature. *Am J Clin Pathol* 2009;**131**:418–431.
- Shoukir Y, Campana A, Farley T, Sakkas D. Early cleavage of in-vitro fertilized human embryos to the 2-cell stage: a novel indicator of embryo quality and viability. *Hum Reprod* 1997;**12**:1531–1536.
- Spencer JB, Klein M, Kumar A, Azziz R. The age-associated decline of androgens in reproductive age and menopausal Black and White women. *J Clin Endocrinol Metab* 2007;**92**:4730–4733.
- Soares SR, Rubio C, Rodrigo L, Simón C, Remohí J, Pellicer A. High frequency of chromosomal abnormalities in embryos obtained from oocyte donation cycles. *Fertil Steril* 2003;**80**:656–657.
- Staessen C, Van Steirteghem AC. The chromosomal constitution of embryos developing from abnormally fertilized oocytes after intracytoplasmic sperm injection and conventional in-vitro fertilization. *Hum Reprod* 1997;**12**:321–327.
- Sugimura S, Ritter LJ, Rose RD, Thompson JG, Smits J, Mottershead DG, Gilchrist RB. Promotion of EGF receptor signaling improves the quality of low developmental competence oocytes. *Dev Biol* 2015;**403**:139–149.
- Sultan KM, Munne S, Palermo GD, Alikani M, Cohen J. Chromosomal status of uni-pronuclear human zygotes following in-vitro fertilization and intracytoplasmic sperm injection. *Hum Reprod* 1995;**10**:132–136.
- Sundstrom P, Nilsson BO. Meiotic and cytoplasmic maturation of oocytes collected in stimulated cycles is asynchronous. *Hum Reprod* 1988;**3**:613–619.
- Sundstrom P, Saldeen P. Early embryo cleavage and day 2 mononucleation after intracytoplasmic sperm injection for predicting embryo implantation potential in single embryo transfer cycles. *Fertil Steril* 2008;**89**:475–477.
- Thurin A, Hardarson T, Hausken J, Jablonowska B, Lundin K, Pinborg A, Bergh C. Predictors of ongoing implantation in IVF in a good prognosis group of patients. *Hum Reprod* 2005;**20**:1876–1880.
- Van den Abbeel E, Balaban B, Ziebe S, Lundin K, Cuesta MJ, Klein BM, Helmgard L, Arce JC. Association between blastocyst morphology and outcome of single-blastocyst transfer. *Reprod Biomed Online* 2013;**27**:353–361.
- Van Royen E, Mangelschots K, De Neubourg D, Valkenburg M, Van de Meerssche M, Ryckaert G, Eestermans W, Gerris J. Characterization of a top quality embryo, a step towards single-embryo transfer. *Hum Reprod* 1999;**14**:2345–2349.
- Van Voorhis BJ, Thomas M, Surrey ES, Sparks A. What do consistently high-performing in vitro fertilization programs in the U.S. do? *Fertil Steril* 2010;**94**:1346–1349.
- Wang SS, Sun HX. Blastocyst transfer ameliorates live birth rate compared with cleavage-stage embryos transfer in fresh in vitro fertilization or intracytoplasmic sperm injection cycles: reviews and meta-analysis. *Yonsei Med J* 2014;**55**:815–825.
- Wang W, Zhang XH, Wang WH, Liu YL, Zhao LH, Xue SL, Yang KH. The time interval between hCG priming and oocyte retrieval in ART program: a meta-analysis. *J Assist Reprod Genet* 2011;**28**:901–910.
- Wang WH, Day BN, Wu GM. How does polyspermy happen in mammalian oocytes? *Microsc Res Tech* 2003;**61**:335–341.
- Wolff HS, Fredrickson JR, Walker DL, Morbeck DE. Advances in quality control: mouse embryo morphokinetics are sensitive markers of in vitro stress. *Hum Reprod* 2013;**28**:1776–1782.
- World Health Organization. *WHO Laboratory Manual for the Examination and Processing of Human Semen*, 5th edn. [http://whqlibdoc.who.int/publications/2010/9789241547789\\_eng.pdf](http://whqlibdoc.who.int/publications/2010/9789241547789_eng.pdf) (1 July 2017, date last accessed), 2010.
- Yanagida K. Complete fertilization failure in ICSI. *Hum Cell* 2004;**17**:187–193.
- Yang WJ, Hwang YC, Lin CS, Hwu YM, Lee RK, Hsiao SY. Embryonic early-cleavage rate is decreased with aging in GnRH agonist but not antagonist protocols. *J Assist Reprod Genet* 2015;**32**:789–795.
- Zamah AM, Hsieh M, Chen J, Vigne JL, Rosen MP, Cedars MI, Conti M. Human oocyte maturation is dependent on LH-stimulated accumulation of the epidermal growth factor-like growth factor, amphiregulin. *Hum Reprod* 2010;**25**:2569–2578.
- Zegers-Hochschild F, Adamson GD, de Mouzon J, Ishihara O, Mansour R, Nygren K, Sullivan E, van der Poel S, International Committee for Monitoring Assisted Reproductive Technology, World Health Organization. The International Committee for Monitoring Assisted Reproductive Technology (ICMART) and the World Health Organization (WHO) Revised Glossary on ART Terminology, 2009. *Hum Reprod* 2009;**24**:2683–2687.