RBMO

ARTICLE

Comparison of 36 assisted reproduction laboratories monitoring environmental conditions and instrument parameters using the same quality-control application

BIOGRAPHY

Giles Anthony Palmer is a Senior Clinical Embryologist, Business and Quality Manager at London Women's Clinic, Wales. He studied Genetics at Leeds University. His collaboration with Athens University resulted in the first births in Greece following embryo-biopsy for cystic fibrosis and beta thalassaemia. He has been accredited with Senior Embryologist Status by ESHRE.

Giles A Palmer[1], Caroline Kratka[2], Sarah Szvetecz[2], Geraldine Fiser[2], Stephen Fiser[2], Cody Sanders[2], Giles Tomkin[2], Mary Ann Szvetecz[2], Jacques Cohen[2,*]

KEY MESSAGE

IVF laboratories (n = 36) in 12 countries used a cloud-based application for instrument monitoring. A retrospective analysis was conducted to assess differences and similarities between laboratories. Major differences in instrument monitoring practices were found in periodic assessment, detail and number of instrument parameters. International standardization is needed.

ABSTRACT

Research question: Assisted reproduction laboratories record instrument performance periodically. No standardized guidelines have been produced for this activity despite mandatory auditing systems in several countries. This study of 36 laboratories in 12 different countries was conducted to assess differences and similarities between quality assurance programmes using an adaptable cloud-based quality-control app for instrument monitoring.

Design: A total of 36 deidentified IVF laboratories that subscribed to the same quality-assurance application were studied. Data were evaluated based on instrument types allocated to 10 domains: incubators, gas tanks, warming surfaces, refrigerators and freezers, cryo-storage, environment, water purification, peripheral equipment, checklists and miscellaneous.

Results: The incubator domain constituted the greatest proportion of parameters (35%), followed by surface warming instruments at 15%. Most incubator O_2 readings were monitored between 4.5 and 5.5%, and between 5.5 and 6.5% for CO_2. The altitude of the laboratory was poorly correlated with the CO_2 setting. Incubator display and measured values of gases and temperature by built-in sensors vary considerably compared with third-party sensors. A quality-control diligence score or mean average data points was calculated for each laboratory. This score is independent of number of instruments or laboratory size. Higher scores were associated with laboratories in countries with government regulations and mandatory auditing systems.

Conclusions: Major differences exist in instrument monitoring practices among laboratories. Although incubator monitoring is the largest domain, many other sensitive instruments are diligently monitored by most laboratories. International standardization and guidelines are needed.

[1] London Women's Clinic, London Women's Clinic Wales, 15 Windsor Place, Cardiff, Wales CF103 BY, UK
[2] Althea Science and IVFqc, ART Institute of Washington, New York City, NY, USA

*Corresponding author. E-mail address: jc@embryos.net (J. Cohen). https://doi.org/10.1016/j.rbmo.2019.03.204 1472-6483/
Declaration: A potential conflict of interest is declared by Geraldine Fiser, Stephen Fiser, Cody Sanders, Giles Tomkin and Mary Ann Szvetecz who are employed by Althea Science, Inc. and Jacques Cohen who is a founder and shareholder of Althea Science, Inc.

INTRODUCTION

The IVF laboratory has developed into a complex environment since its initial inception as a clinical discipline 40 years ago. The ever-increasing success rates are owed, in part, to understanding and controlling adverse environmental conditions (*Cohen, 2018*). During the first decades of assisted reproduction, laboratory details needed to be shared to improve guidelines (*Purdy, 1982; Dawson, 1997*). Embryologist apprenticeships were hard to find. Both the American Fertility Society (AFS) now called the American Society for Reproductive Medicine (ASRM) and the European Society for Human Reproduction and Embryology (ESHRE) began developing quality guidelines and published best practice recommendations many years ago (*Gianaroli et al., 2000*). Quality assurance signifies the standards and processes in laboratory practice. It has become a cornerstone of the profession. Continuous improvement through observation and corrective action after testing and monitoring laboratory equipment (quality control) have led to improved culture conditions and increased IVF success, moving clinical embryology from being simply observational and subjective to becoming an objective clinical science (*Matson, 1998*). Guidelines from governmental institutes and professional organizations have been revised over time (*ASRM, 2008; 2014; Magli et al., 2008; De los Santos et al., 2015*).

A new generation of embryologists, specially trained and accredited, are now in the work place. Awareness of the importance of quality assurance for laboratory systems, including disposables, culture media and instruments specifically designed for assisted reproduction, has increased. The knowledge and means are in place for universal adoption of detailed standards of laboratory management, but defined optimal limits for parameters, such as incubation and laboratory temperatures, are not available internationally; the College of American Pathologists (*CAP, 2018*) provides standards for US-based laboratories subscribed to their system, although details such as observation frequency may not always be addressed (*Hreinsson, 2013*; *Mortimer and Mortimer, 2015*). The global uniformity of quality-assurance standards regarding assessments and record keeping is unknown.

At least 200 confounders affect IVF success (*Pool et al., 2012*). Most of these are concerned with monitoring, staffing, equipment and procedures in the embryology laboratory. It is known that certain conditions, such as optimal temperature and its changes over time, may affect overall success (*Munne et al., 1997; Mortimer and Mortimer, 2015*). Daily or periodic checks of all instruments are recommended (*Esteves and Agarwal, 2013; Anifandis, 2013*), but systematic monitoring of the equipment and environment must be analysed first before patient care can be optimized.

A long list of instrument types are candidates for periodic monitoring. Among these are upright box incubators, benchtop and time-lapse incubators, culture media warming blocks, tube holders, air velocity and pressure gauges, centrifuges, IVF chambers, laminar flow hoods, refrigerators, freezers, Dewars, particle counters, pH sensors, safety systems, monitoring equipment and volatile organic compound sensors. Each new model of such instruments that becomes available will have its own challenges on how best to monitor its performance.

To achieve successful embryo development and clinical outcome, the embryos must be maintained in a stable environment (*Swain, 2014*). Although not well established by robust experimentation, the effects of temperature variation may have adverse consequences: while the pre-implantation embryo can develop within a range of temperatures (*Hong et al., 2012*), elevated temperature can cause spindle damage (*Sun et al., 2014*), aneuploidy (*Munne and Alikani, 2011*) or fragmentation (*Alikani et al., 2005*). When handling gametes and embryos outside incubators, the 'desired' surface temperature may vary among laboratories owing to differences in culture media procedures, details of handling, timing of procedures, air flow, use of micro chambers, room conditions and other factors. A clear consensus is not in place concerning the correct temperature for embryo development, the tolerance limits during manipulation and observation, and the significance of uncontrolled fluctuations.

An increasing number of studies have been published on the effect of poor air quality in the laboratory and its correlation with IVF success (*Esteves and Agarwal, 2013*; *Morbeck, 2015*; *Mortimer et al., 2018*).

Improved laboratory construction methods using clean room technology are slowly becoming commonplace, with many solutions for maintaining a clean air environment (*Varghese and Palmer, 2016*); however, it is unknown how prevalent the practices are for checking environmental conditions in fertility laboratories.

The merits of daily, twice daily logs or continuous monitoring are not always clear, and the benefits must be reflected upon. Corrective action after monitoring is essential whether the system is paper or cloud-based. A paper log system can offer all the benefits that a cloud-based system provides. The advantages of instrument logs are many, but, rather than using handwritten logs, instant electronic transcription may have several advantages. These include data security, instant reporting and cloud access when the embryologist is remote. Key parameters involved in handling, culturing and cryo-storage of gametes and embryos must be monitored, regardless of the data acquisition technology. Fluctuations or drift must also be studied so that adverse conditions can be recognized and remedied. This should be an integral part of the quality-assurance pathway.

Global quality standardizations are needed, but this requires willingness from the IVF community to accept that current standards are inadequate or absent (*Schoolcraft and Meseguer, 2017*). The aim of the present study was to investigate 36 IVF laboratories retrospectively using a cloud-based quality-control recording application. The present study provides an outline of current laboratory practices and may initiate debate about the importance of periodic monitoring of critical instruments and conditions among IVF laboratories internationally.

MATERIALS AND METHODS

Terminology

Instrument domain

'Instrument domain' is the broadest classification for an instrument. It refers to a set of instruments with similar purpose and functionality, e.g. incubators.

Instrument group
'Instrument group' refers to an instrument group that a laboratory has created during application set-up. This term groups instruments with more specific guidelines than 'instrument domain' would., e.g. Minc incubators.

Instrument
An instrument is a piece of equipment that a laboratory has chosen to assign to similar instruments, e.g. incubator 1.

Parameter
Parameter is a factor that a laboratory uses to measure a certain instrument, e.g. temperature, or a component of a checklist, such as 'incubator doors checked', 'dewars locked'. This measurable factor can be chosen from the quality-control application preference menu, such as numerical, Boolean or customizable, i.e. initials of staff members.

Data entry
Data entry represents actual data input for a given parameter, i.e. how many times does the user enter data? This includes repeat factors studied in an instrument group, whereas a parameter would not.

Web application subscription
IVF laboratories have signed up to a new instrument monitoring web application (Reflections App, Althea Science, New York, USA) available since January 2016. The web application allows monitoring and reporting of any equipment using digital data entry rather than handwritten entry or Microsoft Excel. Users can define the instrument type and instrument group and provide the parameters of choice for instrument assessment according to data type: decimal; integer; true/false (yes/no); multiple selection; date; and simple text. Once all instruments and parameters are defined, the order of daily or periodic instrument assessment is determined by a laboratory parameter input system (laboratory log), which is controlled by the laboratory users beforehand, but can be edited *ad hoc* (usually extra readings of one or more instruments). Number-specific parameters require a range such as a minimum and maximum acceptable value. As periodic assessment may differ in intensity and parameter usage, multiple lab logs can be created and named by the laboratory, e.g. daily lab log, intense detailed lab log, gas and liquid nitrogen detail lab log, weekend lab log. The initial daily sequence of data entry is to choose an identifiable lab log previously designed by the laboratory director or embryologists and enter the data in a user-specific order, recording entries until complete. Completion is indicated on the daily application dashboard. Recordings outside the parameter range, e.g. 35.9°C when 36.0°C is the minimum, are acceptable but will trigger alerts. The laboratories have several immediate reports available for quality-assurance purposes, i.e. comparisons of sensors, assessing fluctuations and summary assessments.

Data extraction
The data for this study were extracted from the Reflections application using an algorithm collating summary report data. Extra miscellaneous data were also collected. The data were downloaded as a single, cumulative Excel spreadsheet. The period of use covered 22 months from August 2015 to May 2017.

The laboratories had previously been de-identified using a randomized number system. Listed for each laboratory were the names of each instrument group, the names chosen by the laboratory for the instruments in that group and the data associated with each instrument. For each laboratory, the dates of data entry, the parameters that were measured and the actual inputs recorded were listed. Additional data collected were the minimum and maximum value accepted by that lab for a specific parameter (assuming the entry was numerical), the units in which the parameter was measured, and the laboratory personnel who had entered the data. The latter was not evaluated.

Data sorting and reorganization
After the extracted data were retrieved, they were sorted and reorganized to a more intuitive format. Ten large instrument domains of single or groups of instrument types were selected based on the following classifications: incubators (all types, including big box, small box, desktop/bench top, and time-lapse); warming surfaces (all types including flow hoods and microscope stages), blocks, and boxes (instruments that handle gametes outside of the incubator and must be kept at a particular temperature); gas tanks (the gas cylinders serving the incubators); refrigerators and freezers; cryo-storage (Dewars); laboratory environment, e.g. air quality and pressure, filters, generators, O_2 alarms; water purification systems; peripheral equipment, e.g. instruments with moving parts, centrifuges, pumps; checklist (including where operator action is logged for protocol compliance, typically in a yes/no fashion); and miscellaneous. All the instruments were classified into one of these instrument domains (TABLE 1). All relevant information about each instrument was transferred, including its name and the instrument group it was part of as chosen by the laboratory. Each data parameter for that instrument was classified by data type and its units and 'entry values' were recorded if applicable. These 'entry values' applied primarily to the multiple selection entry type, in which laboratories provided multiple but specific options to choose from. On rare occasions, however, numerical parameters included a text entry (often used as some sort of comment), and these were also recorded under the 'entry options' column. The refined data also listed the date of first collection, date of last collection, and frequency of collection, e.g. once daily, once every other day, twice a month, for each parameter for a particular instrument. Once all the data had been sorted, the total number of instrument groups, instruments, parameters and data entries were grouped for each laboratory. Note that the total number of parameters did not include repeat parameters for instruments in the same group. Total number of data entries included the entries for repeat parameters. Some laboratories were eliminated from the analysis for the following reasons: data collection was for a non-IVF lab; application set-up was incomplete; or laboratories that had less than 10 weeks of recorded data. In addition, specific parameters were excluded from an active laboratory if less than five data entries were recorded for that parameter. If all the parameters for an instrument were disregarded, the instrument itself was excluded.

Overview of refined data
An overview was created in which all the accepted laboratories (n = 36) and the core information about their quality-control processes were recorded. This core information included the total number of instruments overall; the total number of parameters overall; the total number of data points entered overall; its first and last date of data collection; the

TABLE 1 PERCENTAGE OF LABORATORIES THAT HAVE CREATED INSTRUMENT PARAMETERS IN A PARTICULAR DOMAIN

Domain	Laboratories using domain (%)
Incubators All types, including big box, small box, desktop/bench top and time-lapse	100.0
Warming surfaces Including blocks and boxes (instruments that handle gametes outside of the incubator and must be kept at a particular temperature)	91.7
Gas tanks The gas cylinders serving the incubators	41.7
Refrigerators and freezers	100.0
Cryo-storage Dewars and equipment that store gametes and embryos in liquid nitrogen or vapour	36.1
Laboratory environment Including equipment for air quality and pressure, filters, generators and O_2 alarms	88.9
Water purification systems	11.1
Peripheral equipment Instruments with moving parts, centrifuges and pumps	11.1
Checklist Including where operator action is logged for protocol compliance, typically in a yes/no answer or 'electronic signature'	44.4
Miscellaneous	13.9

total number of parameters dedicated to each instrument type separately; and the number of data entries dedicated to that instrument type. This information provided insight into the degree of quality control and how much was committed to each instrument type by each laboratory. Once this overview was completed, each instrument type and all its numerical parameters (with decimal or integer data values) were studied independently. This facilitated the use of quantifiable comparisons between the minimum and maximum recorded values for a particular instrument among participating laboratories. Checklist parameters were assessed separately from the numerical parameters in a qualitative fashion.

Statistics

Pearson's r correlation coefficient was calculated to determine the relationship between the mean CO_2 settings of incubators from all clinics with respect to their altitude.

RESULTS

Overall distribution of data entry

The data entries from 36 eligible laboratories were combined to determine the total number of data points (FIGURE 1A) and total number of parameters (FIGURE 1B) per instrument domain. Although 'data points' include all information for each instrument entered into the application, 'parameters' only counts unique parameters for an instrument group. The incubator domain constituted both the greatest portion of data entries (50%) and parameters (35%). This was followed by the warming surfaces domain in both analyses (11% and 15%, respectively). On the low contribution end for each analysis were the water systems and peripheral equipment domains. A contradictory analysis for a particular instrument domain can be insightful into the nature of that domain. For example, the cryo-storage domain accounts for 11% of total data entries but only 4% of parameters. This indicates that, although laboratories input data for cryo-storage frequently, the number of parameters they are tracking is relatively small. Contrast this with laboratory environment, which is responsible for a small percentage of total data points (4%), but a correspondingly larger percentage of parameters (10%). This relationship suggests that the laboratory environment may not be monitored as frequently as other instrument domains, but that a significant number of parameters are tracked by certain laboratories (32/36) for laboratory environment. The percentage of laboratories that have created parameters in a particular domain is presented in TABLE 1.

The data collected in FIGURE 1 and TABLE 1 are relevant because they provide insight into which instrument domains IVF laboratories tend to focus most quality control. Incubators seem to be an instrument domain that high importance is placed on universally. The data on parameter distribution are shown in FIGURE 2 again, but this time distinguished by laboratory. This separation by laboratory allowed for direct comparisons to be made of parameter number and therefore emphasis on instrument domain.

Comparisons of instrument domain input

Incubators

The oxygen (O_2) (TABLE 2) and carbon dioxide (CO_2) concentrations (TABLE 3) of incubators were compared. For each laboratory that measured the O_2, CO_2 concentration, or both, the averages of their entered values were taken. Most laboratories (81.8%) averaged an O_2 reading of between 4.5% and 5.5%. As for CO_2, most laboratories (72.2%) averaged between 5.5% and 6.5%. The participating laboratories in this study are from various locations across the globe and are therefore found in regions of different altitude, climate and humidity. Altitude will have a thinning effect of the air molecules. In order to maintain the pH, the relative percentage of CO_2 must be increased. Carbon dioxide molecules are more spread apart at increased altitude. The percentage of the incubator atmosphere occupied by CO_2 must be increased to produce the same number of CO_2 molecules as for an incubator located at sea level. To determine whether altitude is correlated with differences in mean percentage of CO_2 among laboratories in this study, the elevations for each participating laboratory were plotted against their mean percent CO_2 entry (FIGURE 3). Although a mild linear increase is present (correlation coefficient = 0.373), many of the data entries fall far from the trendline. This lack of strong correlation is perhaps due to a lack of data or the small number of eligible laboratories. A larger study may confirm whether the relative percentage of CO_2 is being adjusted for altitude.

Refrigerators

The averages of the display and measured temperatures recorded for refrigerators were calculated for each

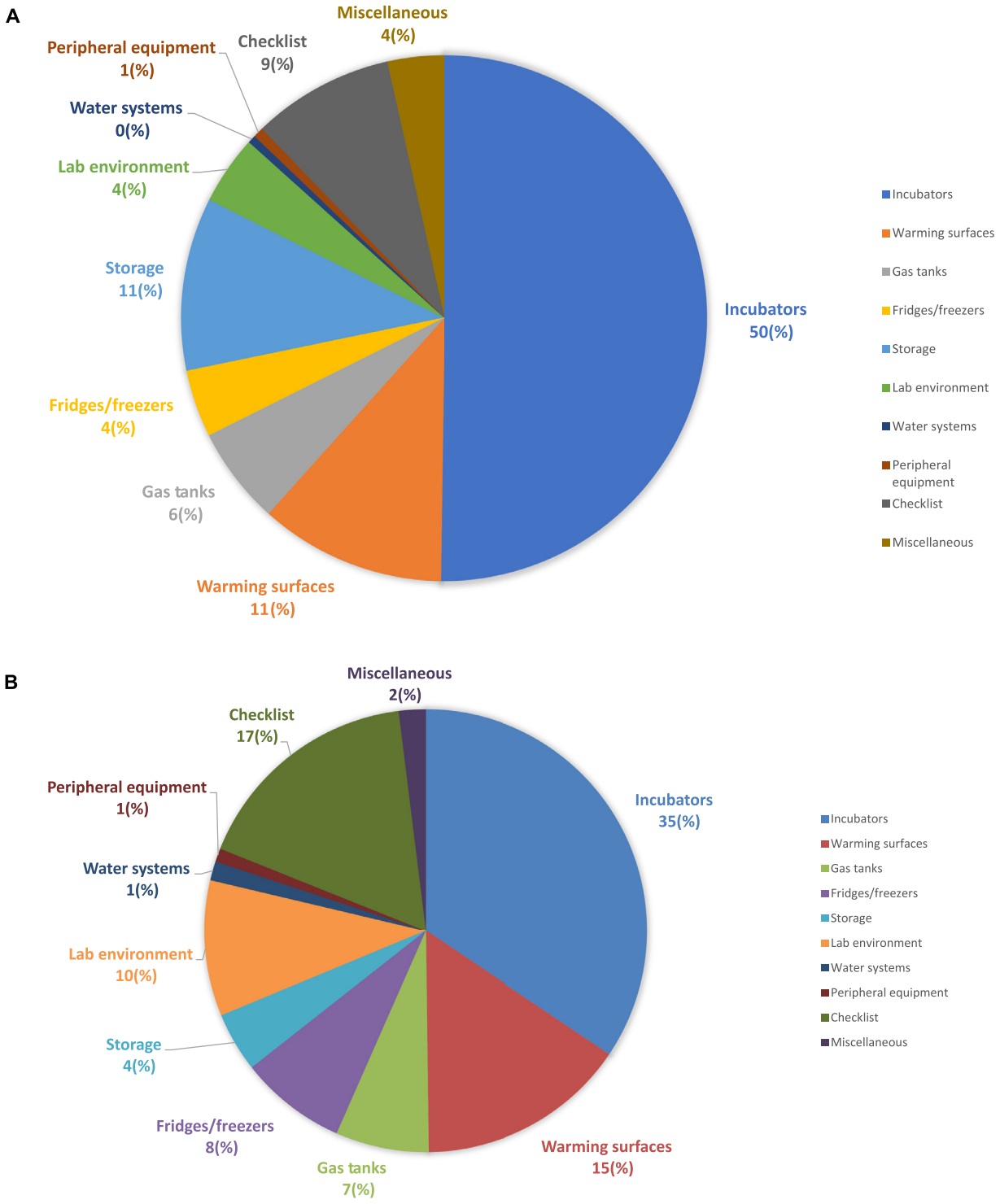

FIGURE 1 (A) Total number of data entries per domain. Data were not differentiated by laboratory but represent the input of all laboratories; (B) the total number of parameters designated towards each instrument domain. Data were not differentiated by laboratory but instead represent the input of all laboratories.

laboratory. Then, the mean of all the display temperature averages, and the mean of all the measured temperature averages, were recorded (TABLE 4). The minimum and maximum laboratory averages for measured temperature were also recorded, falling within 2.3 degrees Celsius of each other. Interestingly, the display and measured temperature means for all laboratories were over half a degree Celsius apart (0.61 C), although, they should ideally be the same. This confirms what has been assumed in a previous study (*Walker et al., 2013*). On average, the difference between the display and measured means in a particular laboratory was 0.67°C. An example of this discrepancy for one

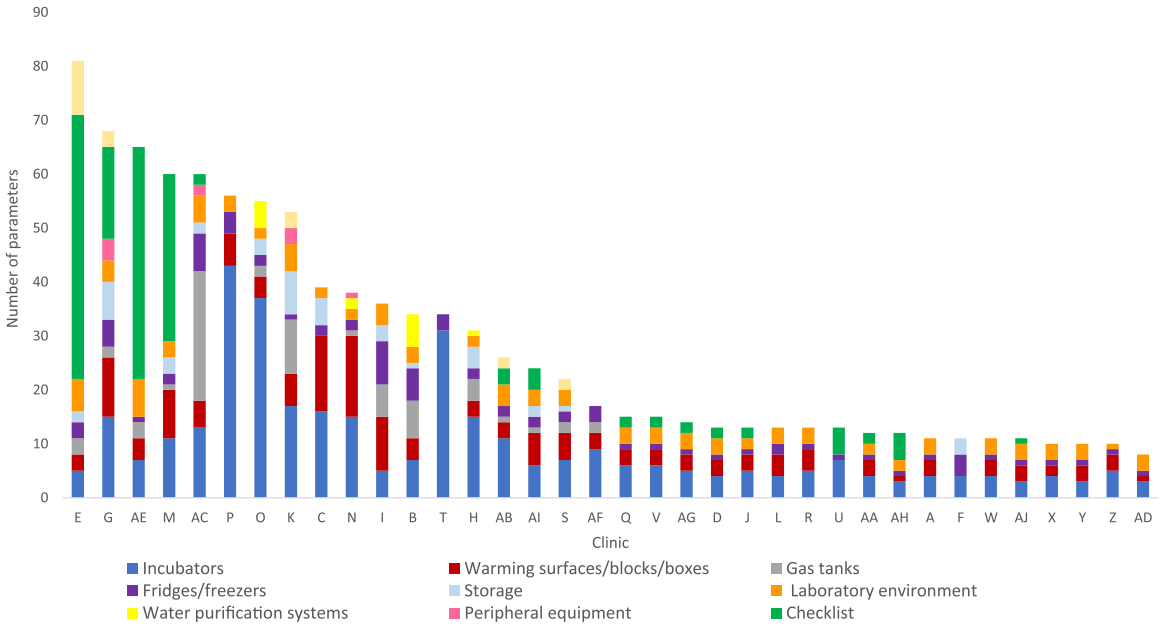

FIGURE 2 Parameter distribution by both instrument domain and participating laboratory. All laboratories were anonymized by one or more letters before the study.

laboratory can be seen in FIGURE 4A. Although the display temperature almost never varies, the measured temperature varies above or below the display value on a near daily basis.

Warming surfaces

The same calculations made on the refrigerator data were conducted on warming surfaces, specifically on surface microscopes, surface hoods and dry blocks (TABLE 4). The display and measured temperature means fell within 36.25–37.89°C, which was expected considering the optimal temperature for cell health. Again, the difference between the display and measured temperature means both overall and within laboratories was noticeable. An example of these differences over time is presented in FIGURE 4B.

TABLE 2 PER CENT OXYGEN IN INCUBATORS AS RECORDED BY LABORATORIES, ACCORDING TO CONCENTRATION RANGE

Oxygen concentration (%)	Incubators (%)
4.5 ≤ to < 5.0	38
5.0 ≤ to < 5.5	43.80
5.5 ≤ to < 6.0	0.00
6.0 ≤ to < 6.5	12.50
6.5 ≤ to < 7.0	0.00
7.0 ≤ to < 7.5	6.30

TABLE 3 PER CENT CARBON DIOXIDE IN INCUBATORS, AS RECORDED BY LABORATORIES, ACCORDING TO CONCENTRATION RANGE

CO_2 concentration (%)	Incubators (%)
4.5 ≤ to < 5.0	0
5.0 ≤ to < 5.5	2.80
5.5 ≤ to < 6.0	27.80
6.0 ≤ to < 6.5	44.40
6.5 ≤ to < 7.0	13.90
7.0 ≤ to < 7.5	5.60
7.5 ≤ to < 8.0	2.80

A score for measuring laboratory quality-control diligence

A surrogate marker for each laboratory's emphasis on monitoring was introduced. Each laboratory expressed per mean average data (MAD) points that were associated with the instruments (number of data entries/number of instruments) is shown in FIGURE 5. The values of the MAD score ranged from 1.28 to 4.51, with higher numbers reflecting increased quality control diligence.

DISCUSSION

This retrospective study examined the system of current laboratory monitoring and highlights the heterogeneity of practices among laboratories worldwide. The laboratories in this study were chosen from IVF clinics that were using the same electronic laboratory application. IVF laboratory equipment monitoring practices were assessed using the application's adaptation settings for each participating laboratory's equipment, types of parameters, tolerance limits and recording frequency. Although the number of participating clinics was small, we believe that the self-selection of clinics subscribing to an application like this, offered publicly, is a representative sample of current laboratory practice. An application assessing quality-control protocols must be versatile enough to accommodate a wide variety of instruments and assessment parameters. In the absence

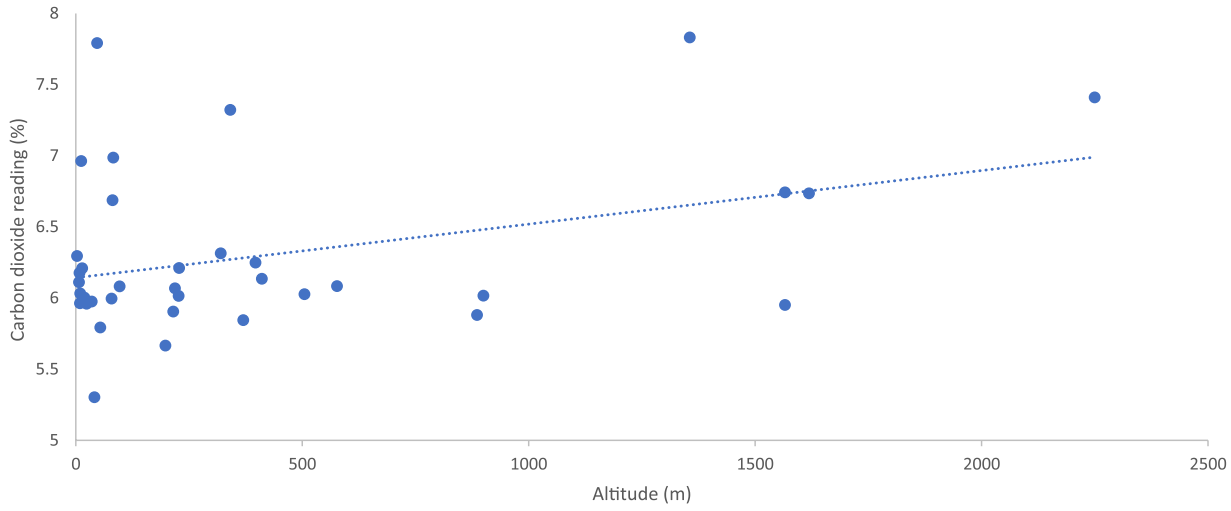

FIGURE 3 The effects of altitude on mean carbon dioxide reading in incubators. The correlation coefficient for a linear relationship was 0.373.

of regulatory guidelines, laboratories must be able to monitor equipment using a multitude of standards. The data must be recorded swiftly and evaluated using flexible reporting mechanisms.

Examples of other instruments or conditions monitored by the study group were laboratory temperature, laboratory humidity, volatile organic compounds, air filter pressure, laboratory O_2 levels, Dewar liquid nitrogen levels, cryo-storage vessel temperatures and various gas cylinder and manifold pressures.

Although checklists (on/off, +/–, yes/no) were mainly confined to protocol accomplishments such as checking of certain events, i.e. cleaning, restocking supplies, incubator water level observations, documenting laboratory close down routines and generally practices requiring Boolean answers or a choice from a 'picklist', such as the name of laboratory members.

An electronic means of data recording and reporting does not by itself improve quality management of a laboratory; it is the responsibility of laboratory staff to measure drift and variations in data readings and to correct deviations and non-conformities. Sharing data from an application like the one reported here may assist fine-tuning ranges of equipment performance and finding commonalities to help optimize laboratory quality control.

The problems that exist in current monitoring philosophies are both qualitative and quantitative. How frequently should critical instruments and conditions be monitored and what are the accepted parameters to meet good practice standards? This report highlights the differences in what instruments are monitored and also to what degree the laboratories are vigilant in their quality-assurance assessments. The time that the preimplantation embryo spends in the IVF laboratory is subject to conditions that have been adapted and improved for more than 40 years (*Cohen, 2018*). There is general agreement that both the equipment and laboratory environment must be working optimally. All IVF laboratory procedures require handling of gametes and embryos on surfaces or instruments that can sustain viability throughout. The data show that no two laboratories practised identically.

It is not surprising considering the importance for healthy cell growth *in vitro* that incubator parameters comprise the largest group of total entries added by laboratory staff, followed by warming surfaces. This illustrates their importance and critical use in the IVF laboratory. Culture conditions in the incubator and temperature-controlled warming surfaces are essential parameters to check daily, particularly because this does not require specialized equipment.

Perhaps a static (rather than dynamic) core temperature of 37°C is not optimal, but regardless, it has become mainstream. Nonetheless temperature must be monitored closely. It is known that differences exist in microscope stage warmers and incubators. It is possible that some heat retention devices or practices in successful clinics are performing more optimally than others

TABLE 4 COMPARATIVE DISPLAY AND SENSOR TEMPERATURES OF WARMING SURFACES, DRY BLOCKS AND REFRIGERATORS

	Surface microscopes	Surface hoods	Dry blocks	Refrigerators
'Display temperature' mean (°C)	37.74	37.89	36.60	4.34
'Measured temperature' mean (°C)	36.25	36.70	37.02	4.95
Maximum 'measured temperature' mean (°C)	38.67	38.12	37.57	5.83
Minimum 'measured temperature' mean (°C)	34.12	34.47	36.51	3.53
Difference between 'display' and 'measured' means (°C)	1.49	1.19	0.42	0.61
Average of differences between 'display' and 'measured' means (°C)	1.32	1.36	0.45	0.67

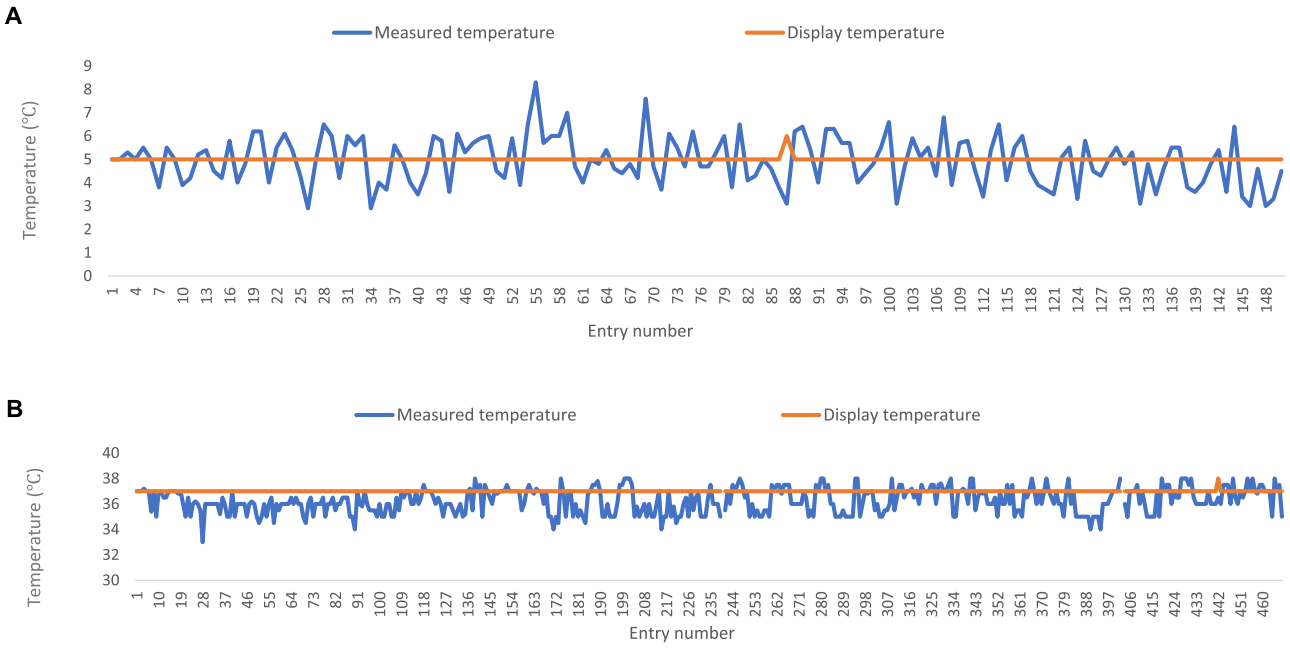

FIGURE 4 (A) Example of discrepancy between refrigerator display temperature and measured temperature; (B) example of discrepancy between surface hood display temperature and measured temperature.

(*Cooke et al., 2002*). It is also known that display temperatures should not be relied upon and incubators should be measured by an external standard (*Swain, 2014*). It is concerning that some laboratories may rely on the manufacturers' display temperature both for incubators and refrigerators.

Moreover, the mean display temperatures of microscope stages and hood surfaces suggest that the correct temperature to ensure proper operating temperatures in the dishes is often not set (TABLE 4). With reference to microscope stages, 20 laboratories in the study did not compare display temperatures with independent temperature measurements, whereas only four out of 16 adjusted the displayed temperature reading to correspond with the independently measured temperature. Reasons why little attention to such adjustment was paid are multiple, such as the variety of commercial and laboratory-made devices used on the microscope stage, room temperature or even the type of dish.

Embryo culture media must be stored according to the manufacturers' recommendations and daily recording of storage temperature may highlight its importance. All laboratories in the present study monitored refrigerators and freezers used for storage of culture media.

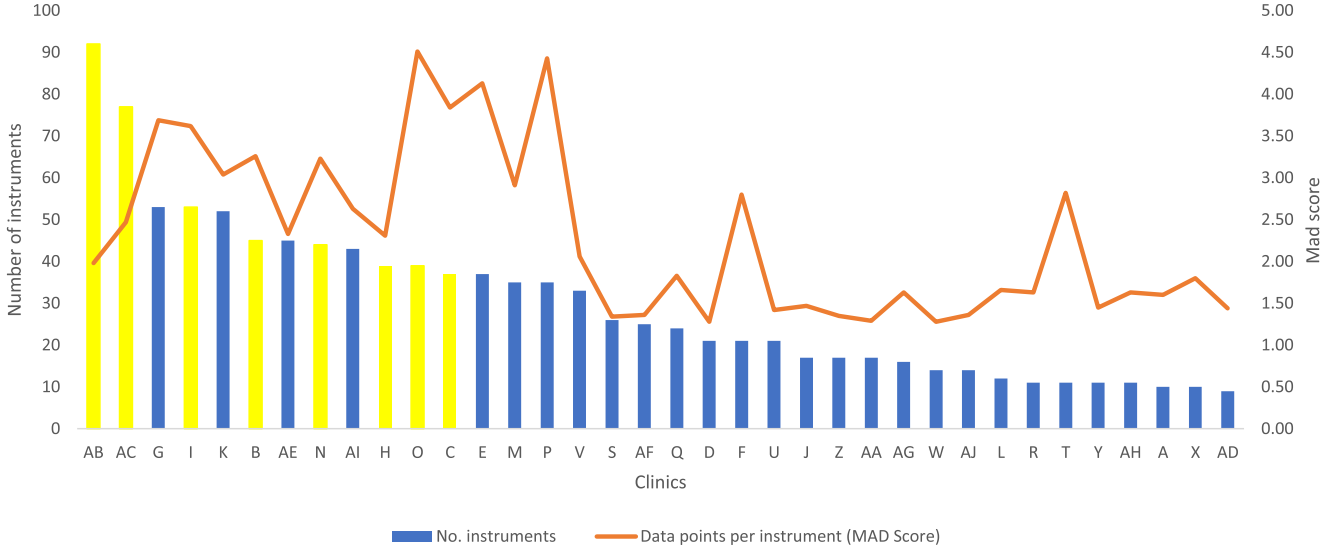

FIGURE 5 Mean average data scores: mean average data points in order of descending number of instruments present in the anonymized laboratories. Yellow bars represent clinics that submit data to The Society for Assisted Reproductive Technology.

Daily manual checks of liquid nitrogen storage vessels were relatively low (only 36% of clinics) and consisted of either manual temperature measurements or low-level measurements of liquid nitrogen. It is likely that the other clinics used an automatic alarm-based type of system or perhaps made visual checks of liquid nitrogen levels, but not maintained in the application log plan. It is possible that some clinics in this study used paper records to monitor Dewars. They may not have been aware of the flexibility of the application. Cryo-storage relates to 11% of total data entry but only 4% of parameters, indicating a low number of parameters being tracked to verify temperature compliance. Liquid nitrogen levels can obviously be monitored automatically and perhaps those data were not entered by all clinics under Dewar maintenance but rather filed under checklists. In light of recent events in which two clinics suffered catastrophic liquid nitrogen storage tank failure (*Alikani, 2018*), daily visual checks using a dip stick to measure LN2 height in addition to an automatic system can possibly evert disaster. Additional counter measures may be developed, which may detect deteriorating Dewar integrity (such as the presence of ice or condensation around the Dewar neck or side) and progressive liquid nitrogen evaporation. Such aspects may be monitored visually or using sensors.

The wide variety of peripheral instruments that were monitored indicates both the flexibility of the laboratory log and the ingenuity of the laboratory managers to customize the programme to include novel instruments and conditions. Although some attention was paid to the laboratory environment, this was not a universal custom. Low volatile organic compounds, humidity and temperature control make for a more conducive work environment for embryologists, and are important in maintaining optimum conditions for embryo culture and instrumentation. Despite this, there is little consensus on adequate laboratory room temperature and humidity. Higher than standard room temperature may be favourable to keep a stable 37°C environment during gamete and embryo manipulation, although there may be a slightly increased risk of contamination. Most laboratory instruments have an optimum room temperature range that must not be exceeded for them to function optimally. It is interesting that 15 of the 36 clinics were located in tropical or subtropical areas where controlling room temperature presents a challenge. The versatility of the system used here makes the IVF laboratory logs customizable, and this was evident when analysing the number and diversity of parameters.

Mean average data scores

Clinical embryologists validate, monitor and re-validate, but no direct measurement compares laboratory diligence. The MAD score introduced here could act as a simple bench mark to rate each laboratory according to its monitoring prowess. This score was used to compare laboratories' equipment monitoring dexterity or quality-control diligence.

The MAD score is the number of data entries per day divided by the number of instruments. The higher the score the more quality-control readings are being conducted regardless of the size and volume of work performed by the clinic. If daily checks are to be standardized, it would be fair to say that the MAD score would be similar in all laboratories big and small. It would not matter if the laboratory has less equipment, because the same number of checks per instrument would be recorded.

It is clear from the data presented that not all laboratories have similar quality control diligence; it is possible that high scoring laboratories are more successful and, in turn, attract more patients and thus are equipped with more instruments to accommodate this demand. This should be confirmed in additional studies.

Similarly, with reference to FIGURE 1 and TABLE 1, it would be interesting to evaluate laboratories that place an uncommon emphasis on particular instrument domains and look for potential correlations with key performance indicators.

Laboratories were included from all over the world where we know legislation on practices and IVF success reporting differs; interestingly, the UK and US laboratories fell among the top 50% high MAD score when the data were unblinded to country of origin. The relative sample size, however, must be taken into account, and this finding can only be confirmed in a larger study of this kind.

Temperature control

Although manufacturers' recommendations for storage of media have a wide temperature range of 2–8°C, this range does not apply to the safe range of temperature that gametes and embryos should endure. Only in recent years with the increasing use of both time-lapse incubators and single step media have embryos been subject to an uninterrupted environment (*Meseguer et al., 2012*). Frequent opening of incubator doors and internal variations of temperature relative to location within shelves in 'box' incubators has been reported (*Fujiwara et al., 2007; Anifandis, 2013; Walker et al., 2013*), whereas video surveillance inside compact incubators provokes little disturbance during culture. Few studies; however, have been conducted on the ideal temperature for IVF embryos (*Hong et al., 2012, 2013 and 2014*), despite numerous animal studies showing tubal temperatures at least 1°C lower than the core temperature (*Baak et al., 2016*).

Embryologists carry out tasks in the laboratory with the perception of working at optimal temperature. Meticulous temperature monitoring must be undertaken to control temperatures while handling culture dishes outside the incubator. Any change in consumables or protocols must be met with the immediate validation of the operating temperature. The type and thickness of plasticware may alter the culture conditions owing to its heat retention capacity, whereas the temperature of the culture droplet or well may differ due to the volume of medium (*Cooke et al., 2002*).

Highlighted in this study are the differences between display temperatures and actual measured values. Values from internal sensors in both refrigerated and laminar-flow hood surfaces remain almost constant over time (FIGURE 4a and FIGURE 4b), illustrating the lack of sensitivity in the display. It can be clearly seen from independent measurements considerable periodic fluctuations occur in temperature (*Blomfield et al., 2016*). Acknowledgement of perceived versus real temperatures is imperative for maintaining a robust internal (quality assurance) programme when conducting daily readings of temperature.

In addition, the external reading from an independent probe is dependent

on its location and how the location is measured (tube, dish, surface attachment and characteristics, height of dish, rim footprint). Probes located in different areas may read differently. This may explain the differences between measured and display temperatures in refrigerators and freezers. It is advisable to re-calibrate the independent measuring tool with a calibrated device. Frequent and periodic calibration by certified companies should be implemented. Laboratories may be measuring differently in a dish or using a thermocouple attached to a surface. Without knowing how each approach affects outcomes, it is important to consider repeatability and consistency of the measurement approach. Once the optimal location and calibration of the independent measuring device is assured, a display reading may only serve to alert the staff that there is a possible malfunction.

Oxygen and carbon dioxide measurements

By comparing the gaseous environment for embryo incubation, this study showed two approaches by IVF teams. First, great variation was found in oxygen levels during embryo culture in laboratories using reduced oxygen levels. The culturing of mammalian embryos in incubators with reduced O_2 concentration is beneficial to their development (*Wale and Gardner, 2016*). Human embryo culture in a more physiological O_2 tension is associated with improved blastocyst quality (*Hoff et al., 2008*) and live birth rates (*Meintjes et al., 2009*). Despite this, it is not practised universally; a previous online survey involving 54 different countries revealed that less than 25% of human embryo culture is carried out under physiological oxygen levels (*Christianson et al., 2004*).

Most clinics in this study reported an operational O_2 concentration in their incubators of between 4.5 to less than 5.5%. Although 5–6% is most often cited in published research, some laboratories were using concentrations recorded in the higher range of 7.0% to less than 7.5%. This range may reflect the choice of laboratories to reduce the high cost of using low oxygen tension. It is also possible that embryos may exhibit hypoxia at low oxygen concentrations (*De los Santos et al., 2013*).

All laboratories either logged or measured incubator carbon dioxide levels. Checking concentrations of CO_2 may be the most practical indicator of a stable intracellular pH of the developing embryo. This parameter was expected to show variations among laboratories using different media formulations, and may occasionally need to be adjusted according to the batch of culture medium. An important concern is accuracy of CO_2 measurements because measuring devices must be calibrated periodically.

Interestingly, it is known that the CO_2 partial pressure (which determines the amount of CO_2 that dissolves in the culture medium) changes with altitude. As the elevation rises and atmospheric pressure drops, the relative percentage of CO_2 must be increased to produce the equivalent pH at elevated altitude compared with that at sea level.

Clinics participated in this study from all over the world with elevations from sea level to 2230 m, and we expected a closer correlation between the rise of effective CO_2 concentration and altitude. Although a mild linear increase was present, many values fell far from the trend line. This may derive partially from the small numbers of eligible clinics in this study, but it is still a cause for concern that several laboratories have not adjusted incubation CO_2 relative to their altitude. It should be made clear that increasing percent CO_2 levels at higher altitudes to reflect the correct operating pH for a specific formulation should not be detrimental and must be accompanied with external pH measurements.

Air quality

During the first decades of IVF, the adverse effect of poor air quality within the laboratory was generally anecdotal. It was later shown that highly volatile organic compound levels can exist in laboratory air and incubator environments (*Cohen et al., 1997*). This heightened awareness has led to common practices such as 'off-gassing' of plasticware and 'running or burning in' new equipment (releasing residual volatile organic compounds from the manufacturing process) outside of the critical area (*Wiemer, 2017; Mortimer et al., 2018*). Introduction of cleanroom standards, such as those used in the biotechnology industry, has been proposed to better protect gametes and embryos subjected to the laboratory environment (*Palmer, 2013*), and regulators such as the European Union's Tissue and Cells Directive (*Anonymous, EUTCD; 2004*/EC) stipulate quality and safety requirements with the critical point of the doctrine being clean air.

Despite mounting evidence for the importance of good laboratory air supply, the number of clinics following these guidelines was expected to be higher than the reports that only four laboratories transcribed monitoring laboratory air and only two measured the presence of volatile organic compounds using the application. It remains to be seen, however, if the portable volatile organic compounds measuring devices and particle counters are sensitive enough to warrant widescale use, without periodic third-party testing of the laboratory environment using gas chromatography and mass spectroscopy.

Periodic versus continuous monitoring

Maintaining critical conditions through 24 h of remote (continuous) monitoring is mandated by some national authorities, such as the UK's Human Fertilisation and Embryology Authority (*HFEA, 2017*). Out of routine working hours, alerts on instrument performance are important to ensure that critical parameters are maintained between limits at all times, but daily static monitoring also has great importance. Some laboratories may rely completely on remote monitoring systems that alarm and notify when parameters disagree with the set range. Other laboratories that are manually monitoring daily changes believe that this provides optimal instrument performance. Scoring drift and parameter fluctuations can be an invaluable reference for assessing instrument performance. Immediate access is now possible using a quality control data application to assess and report performance.

Although this may arm the embryologist with vital information for routine inspections, audits and trouble-shooting, little is known about the possibility of data overload in tracking IVF-laboratory instruments. Although instruments such as incubators and microscope heating pads are known confounders of IVF outcomes, whether weekly, daily, bi-daily or continuous assessment correlates differentially with outcomes such as fertilization and development is unknown. Measuring this consistently, once daily, may seem the most convenient and efficient, but this needs further investigation.

One of the aims of having a common system as reported here is to study whether optimal approaches to equipment monitoring policies have been established. It is now possible to correlate quality-control diligence parameters, such as refrigerator conditions during storage of media batches, with outcomes such as fertilization, embryo development and implantation. This may require a large data set from clinics located in countries with mandated national outcome reporting systems. Preferably, there should be an integration of quality-control data with electronic medical record keeping tools. This process could determine optimal quality-control conditions for any aspect of assisted reproductive technology.

Recent successful consensus meetings held in Cairo and Vienna focused on laboratory practices (*ESHRE, 2017*; *Mortimer et al., 2018*) and highlight the need for dialogue among experts to formulate worldwide standards. We should move away from adjectives such as 'adequate', 'suitable' and 'recommended' and use quantitative parameters as well as guidelines to determine minimal standards of laboratory monitoring logs. Compliance with external quality standards such as the International Organization for Standardization have become the foundation for ensuring good working practices but can be superficial when investigating the subtleties of factors affecting success. Internal quality control and routine laboratory practices need a record-keeping overhaul. The sharing of data is the only way to improve quality-control practices internationally.

REFERENCES

Alikani, M., Schimmel, T., Willadsen, S.M. **Cytoplasmic fragmentation in activated eggs occurs in the cytokinetic phase of the cell cycle, in lieu of normal cytokinesis, and in response to cytoskeletal disorder.** Mol. Hum. Reprod. 2005; 11: 335–344

Alikani, M. **Cryostorage of human gametes and embryos: a reckoning (Editorial).** Reprod. Biomed. Online. 2018; 37: 1–3

Anifandis, G. **Temperature variations inside commercial IVF incubators.** J. Assist. Reprod. Genet. 2013; 30: 1587–1588

Anonymous**. Directive 2004/23/EC of the European Parliament and the Council on Setting Standards of quality and safety for the donation, procurement, testing, processing, preservation, storage and distribution of human tissues and cells.** Off. J. Eur. Union 2004; 102: 48–58

ASRM Practice committee**. Revised guidelines for human embryology and andrology laboratories.** Fertil. Steril. 2008; 90: S45–S59

ASRM Practice committee**. Revised minimum standards for practices offering assisted reproductive technologies: a committee opinion.** Fertil. Steril. 2014; 102: 682–686

Baak, N.A., Cantineau, A.E.P., Farquhar, C., Brison, D., 2016. **Temperature of embryo culture for assisted reproduction Cochrane Database of Systematic Reviews.** www.cochranelibrary.com

Blomfield, S., Redding, G.P., Bronlund, J.E., Stewart, B., Peek, J. **A Model for Characterisation of the Thermal Environment of Embryos in a Petri Dish During In Vitro Culture.** Cell. Mol. Bioeng. 2016; 9: 546–555

Christianson, M.S., Zhao, Y., Shoham, G., Granot, I., Safran, A., Khafagy, A., Leong, M., Shoham, Z. **Embryo catheter loading and embryo culture techniques: results of a worldwide web-based survey.** J. Assist. Reprod. Genet. 2004; 31: 1029–1036

Cohen, J., Gilligan, A., Esposito, W., Schimmel, T., Dale, B. **Ambient air and its potential effects on conception in vitro.** Hum. Reprod. 1997; 12: 1742–1749

Cohen, J. **How the embryology laboratory has changed!** Niederberger C., Pellicer A (eds). Forty years of IVF. Fertil. Steril. 2018; 110: 189–191

College American Pathologist. **Reproductive accreditation program.** https://www.cap.org/laboratory-improvement/accreditation/reproductive-accreditation-program

Cooke, S., Tyler, J.P.P., Driscoll, G. **Objective Assessments of Temperature Maintenance Using In Vitro Culture Techniques.** J. Assist. Reprod. Genet 2002; 19: 368–375

Dawson, K.J. **Quality control and quality assurance in IVF laboratories in the UK.** Hum. Reprod. 1997; 12: 2590–2591

De los Santos, M.J., Gamiz, P., C, Albert, C., Galan, A., Viloria, T., Perez, S., Romero, J.L.I., Remohí, J. **Reduced oxygen tension improves embryo quality but not clinical pregnancy rates: a randomized clinical study into ovum donation cycles.** Fertil. Steril. 2013; 100: 402–407

De los Santos, M.J., Apter, S., Coticchio, G., Debrock, S., Lundin, K., Plancha, C.E., Prados, F., Rienzi, L., Verheyen, G., Woodward, B., Vermeulen, N. **Revised guidelines for good practice in IVF laboratories.** Hum. Reprod. 2015; 31: 685–686. http://humrep.oxfordjournals.org/

Esteves, S., Agarwal, A. **Ensuring that Reproductive Laboratories Provide High Quality Service.** Bento F., Esteves S.C., Agarwal A. Quality Management in ART Clinics A Practical Guide Springer New York 2013: 129–146

Esteves, S., Agarwal, A. **Explaining How Reproductive Laboratories Work.** Bento F., Esteves S., Sandro C., Agarwal A. Quality Management in ART Clinics A Practical Guide Springer Publications New York 2013: 79–127

ESHRE Special Interest Group of Embryology and Alpha Scientists in Reproductive Medicine. **The Vienna consensus: report of an expert meeting on the development of ART laboratory performance indicators.** Reprod. Biomed. Online 2017; 35: 494–510

Fujiwara, M, Takahashi, K., Izuno, M., Duan, Y.R., Kazono, M., Kimura, F., Noda, Y. **Effect of micro-environment maintenance on embryo culture after in-vitro fertilization: comparison of top-load mini incubator and conventional front-load incubator.** J. Assist. Reprod. Genet. 2007; 1: 5–9

Gianaroli, L., Plachot, M., van Kooij, R., Al-Hasani, S., Dawson, K., DeVos, A., Magli, M.C., Mandelbaum, J., Selva, J., van Inzen, W. **ESHRE guidelines for good practice in IVF laboratories. Committee of the Special Interest Group on Embryology of the European Society of Human Reproduction and Embryology.** Hum. Reprod. 2000; 10: 2241–2246

Hoff, A., Khabani, A., Khabani, C., Hickok, L., Marshall, L. **Reduced oxygen tension helps increase the quality of blastocyst available on D5.** Fertil. Steril. 2008; 90: S350

Hong, K.H., Forman, E.J., Lee, H., Ferry, K.M., Treff, N., Scott, R. **Optimizing the temperature for embryo culture in IVF: a randomized controlled trial (RCT) comparing standard culture temperature of 37°C to the reduced more physiologic temperature of 36°C.** Fertil. Steril 2012; 93: S167

Hong, K.H., Forman, E.J., Lee, H., Upham, K.M., Treff, N.R., Scott, R.T.Jr. **Optimizing the temperature of embryo culture in IVF: a randomized controlled trial.** Fertil. Steril 2013; 100: S26

Hong, K.H., Lee, H., Forman, E.J., Upham, K.M., Scott, RT.Jr. **Examining the temperature of embryo culture in in vitro fertilization: a randomized controlled trial comparing traditional core temperature (37°C) to a more physiologic, cooler temperature.** Fertil. Steril. 2014; 102: 767–773

Human Fertilisation and Embryology Authority. 2017 **Code of Practice Edition 8.0.** https://www.hfea.gov.uk/code-of-practice/

Hreinsson, J. **Quality Management.** Vargese A., C., Sjoblom P., Jayaprakasan K A Practical Guide to Setting Up an IVF Lab, Embryo Culture Systems and Running the Unit Jaypee Publishers New Delhi 2013: 87–93

Magli, C., Van den Abbeel, E., Lundin, K., Royere, D., Van der Elst, J., Gianaroli, L. **Revised guidelines for good practice in IVF laboratories.** Hum. Reprod. 2008; 23: 1253–1262

Matson, P.L. **Internal quality control and external quality assurance in the IVF laboratory.** Hum. Reprod. 1998; 13: 156–164

Meintjes, M., Chantilis, S.J., Douglas, J.D., Rodriguez, A.J., Guerami, A.R., Bookout, D.M., Barnett, B.D., Madden, J.D. **A controlled randomized trial evaluating the effect of lowered incubator oxygen tension on live births in a predominantly blastocyst transfer program Hum.** Reprod 2009; 24: 300–307

Meseguer, M., Rubio, I., Cruz, M., Basile, N., Marcos, J., Requena, A. **Embryo incubation and selection in a time-lapse monitoring system improves pregnancy outcome compared with a standard incubator: a retrospective cohort study.** Fertil. Steril. 2012; 98: 1481–1487

Morbeck, D.E **Air quality in the assisted reproduction laboratory: a mini-review.** J Assist Reprod Genet 2015; 32: 1019–1024

Mortimer, D., Mortimer, S.T. 2015 **Quality and Risk Management in the IVF Laboratory latest.** Cambridge University Press

Mortimer, D., Cohen, J., Mortimer, S.T., Fawzy, M., McCulloh, D.H., Morbeck, D.E., Pollet-Villard, X., Mansour, R.T., Brison, D.R., Doshi, A., Harper, J.C., Swain, J.C., Gilligan, A.V. **Cairo consensus on the IVF laboratory environment and air quality: report of an expert meeting.** Reprod. Biomed. Online. 2018; 36: 658–674

Munne, S., Magli, C., Adler, A., Wright, G., DeBoer, K., Mortimer, D., Tucker, M., Cohen, J., Gianaroli, L. **Treatment-related chromosome abnormalities in human embryos.** Hum. Reprod. 1997; 12: 780–784

Munne, S., Alikani, M. **Culture-induced chromosome abnormalities: the canary in the mine.** Reprod. BioMed. Online 2011; 22: 506–508

Palmer, G.A. **The Future of IVF.** Varghese A.C., Sjoblom P., Jayaprakasan K. A Practical Guide to Setting up an IVF Lab, Embryo Culture Systems and Running the Unit Jaypee Medical Publishers New Delhi 2013: 213–229

Pool, T.B., Schoolfield, J., Han, D. **Human Embryo Culture Media Comparisons.** Smith G., D., Swain J., E., Pool T., B. Embryo Culture: Methods and Protocols Humana Press New York 2012: 367–386. doi:10.1007/978-1-61779-971-6_21

Purdy, J.M **Methods for fertilization and embryo culture in vitro.** Edwards R.G, Purdy JM. Human conception In vitro: Proceedings of the first Bourn Hall meeting Academic Press London 1982: 135–148

Schoolcraft, W., Meseguer, M. **Paving the way for a gold standard of care for infertility treatment: improving outcomes through standardization of laboratory procedures.** Reprod. Biomed. Online. 2017; 35: 391–399

Sun, X.F., Zhang, W.H., Chen, X.J., Xiao, G.H., Mai, W.Y., Wang, W.H. **Spindle dynamics in living mouse oocytes during meiotic maturation, ageing, cooling and overheating: a study by polarized light microscopy.** Zygote. 2014; 12: 241–249

Swain, J.E. **Decisions for the IVF laboratory: comparative analysis of embryo culture incubators.** Reprod. BioMed. Online 2014; 28: 535–547

Varghese, A.C., Palmer, G.A. **Clean room technology for low resource IVF units.** Esteves S.C., Varghese A.C, Worrilow K.C. Clean Room Technology in ART Clinics A Practical Guide CRC press Florida 2016: 347–354

Wale, P.L., Gardner, D.K. **The effects of chemical and physical factors on mammalian embryo culture and their importance for the practice of assisted human reproduction.** Human Reproduction Update 2016; 22: 2–22

Walker, M.W., Butler, L.M., Higdon, H.L., Boone, W.R. **Temperature variations within and between incubators-a prospective, observational study.** J. Assist. Reprod. Genet. 2013; 30: 1583–1585

Wiemer, K.L. **Setup Procedures for optimizing performance in the IVF laboratory.** Montag M.H.M., Morbeck D.E. Principles of IVF Laboratory Practice: Optimizing Performance and Outcomes Cambridge University Press 2017: 88–96

Received 11 November 2018; received in revised form 2 March 2019; accepted 6 March 2019.