Asian Journal of Andrology (2010) 12: 21–25

www.nature.com/aja

Perspectives

Quality assurance and quality control in the laboratory andrology

Allan A. Pacey

School of Medicine and Biomedical Sciences, Academic Unit of Reproductive and Developmental Medicine, the University of Sheffield, the Jessop Wing, Tree Root Walk, Sheffield S10 2SF, UK

Abstract

Quality assurance (QA) and quality control (QC) are fundamental aspects of any laboratory measurement. However, in comparison with other medical disciplines, the need for QA and QC in laboratory andrology has been recognized only recently. Furthermore, there is evidence that the effort required to undertake QA and QC has not been wholly welcomed by some clinicians. Nevertheless, accrediting bodies and regulatory authorities increasingly require evidence that laboratories have effective QA and QC measures in place because both are central to the quality management processes. Following the publication of the 5th edition of the World Health Organization Laboratory Manual, existing QA and QC systems will need to be updated to take into account some of the methodological changes recommended by the manual. Three of these are discussed in this commentary; they relate to: (i) the move to infer semen volume from its weight; (ii) the re-classification of sperm motility grades from four to three; and (iii) the publication of a lower reference limit for morphology of 4% (with a corresponding 95% confidence interval of 3%–4%). The importance of QA and QC in all laboratory tests, including up and coming new tests to assess sperm DNA integrity, is discussed. The need for adequate initial training and continuing professional development programmes to support laboratory scientists performing andrology is also described.

Asian Journal of Andrology (2010) 12: 21–25. doi: 10.1038/aja.2009.16

Keywords: andrology, health care, quality assurance, quality control, semen analysis

1 Introduction

Quality assurance (QA) and quality control (QC) are essential aspects of any laboratory testing process [1]. Both ensure that the data generated by the laboratory are consistent from one day to the next and that the results from one laboratory can be compared with those generated by others. However, whereas QA and QC approaches in many areas of laboratory medicine (e.g., clinical biochemistry and haematology) were relatively well developed in the 1970s and 1980s [2, 3], their application to laboratory andrology has been a relatively recent development.

Early editions of the World Health Organization (WHO) Laboratory Manual for the Examination of Human Semen and Sperm–Cervical Mucus Interaction [4–6] made almost no reference to the need for QA or QC as a part of semen analysis. Yet, a subsequent series of publications [7–9] made the observation that significant disagreement could occur between analysing laboratories, hence highlighting the need for improved QA and QC. It was, therefore, a significant step forward when the 4th edition of the WHO manual [10] published detailed guidelines for QA and QC of semen analysis for the first time.

The increased emphasis placed on QA and QC procedures in the 4th edition [10] was not welcomed by everyone in the field. In 2005, for example, Jequier wrote, 'it would appear that semen analysis is for the most part carried out satisfactorily all over the world' and therefore 'the energy expended in running QA schemes in relation to semen analysis might now be a waste of time' [11]. Although this position was not supported by other authors [12–14], the views of Jequier are still shared by many of those with responsibility for setting budgets and allocating

Correspondence to: Dr Allan A. Pacey, School of Medicine and Biomedical Sciences, Academic Unit of Reproductive and Developmental Medicine, the University of Sheffield, Level 4, the Jessop Wing, Royal Hallamshire Hospital, Sheffield S10 2SF, UK.
Fax: +44-114-226-1074 E-mail: A.Pacey@Sheffield.ac.uk
Received: 7 January 2009 Revised: 8 January 2009
Accepted: 6 February 2009

resources. However, Pacey [14] argued that what such views fail to recognize is that QA and QC are now firmly embedded in ISO-based quality systems [15], besides being requirements of national and regional legislation, such as the European Commission on Setting Standards of Quality and Safety for the Donation, Procurement, Testing, Processing, Preservation, Storage and Distribution of Human Tissues and Cells [16]. As such, QA and QC in andrology are here to stay.

2 Existing challenges for andrology QA and QC

In order for effective QA and QC in the andrology laboratory to take place, the heterogenous nature of the analyte (i.e., human semen) must be taken into account. Unfortunately, the QA and QC of many laboratory measurements made during semen analysis (e.g., of sperm morphology and motility) are not as straightforward as those undertaken in other pathological disciplines. For example, in endocrinology, the measurement of blood hormone levels is done alongside control standards and is performed using large automated analysers. Although the assessment of seminal plasma biochemistry could be comparable in this regard (notwithstanding the difficulties encountered in accurately dispensing aliquots of a viscous, non-homogeneous fluid), such tests are now rarely undertaken outside the research setting, and many consider them to be of limited clinical value.

Therefore, the bulk of measurements made on human semen rely on the visual identification of spermatozoa by a human observer who counts them (sperm concentration), evaluates their behaviour (sperm motility) or attempts to estimate (or measure) the size and shape of cellular features (sperm morphology). Despite the definitions and guidelines in place [10], the fact that these measurements are made by humans implies that they are invariably prone to error. Moreover, they are usually undertaken with only a small sub-population of spermatozoa (often fewer than 200 in duplicate of a potential population of millions), from which the final results are mathematically extrapolated. Nor are there control standards that can be referred to during this process (e.g., a bottle of spermatozoa with normal morphology is unavailable). Moreover, there may be problems with initial staff training, maintenance of competence and the variation imposed by choice of method, all of which mean that from the very beginning QA and QC of semen analysis present a unique challenge in laboratory medicine.

It is unclear how many andrology laboratories around the world have attempted to embrace the principles of QA and QC as outlined in the 4th edition of the WHO manual [10]. Surveys of laboratory practice suggest that there has been relatively poor implementation of other methodologies. For example, in the USA, 6% and 5% of laboratories fail to report data for sperm concentration and motility, respectively, as part of semen analysis, and 15% fail to report sperm morphology [17]—this is in spite of the fact that each of these measures is listed as a mandatory test in the 4th edition. Moreover, in the assessment of sperm morphology, 83% of US [17] and 69% of UK [18] laboratories that report it do so on the basis of their observation of fewer than 100 spermatozoa. Although on the surface this may seem trivial, it does indicate that the rationale behind the WHO recommendations has not been fully accepted. It also suggests that laboratory scientists do not fully understand the statistical basis on which accurate measurements are made [19]. This is a crucial point to bear in mind when considering how some of the changes in the 5th edition of the WHO manual will have an impact on andrology QA and QC.

3 How will the new manual change andrology QA and QC?

The 5th edition of the WHO manual has been updated in many ways from the 4th edition [10]. With regard to QA and QC, there are three areas in which the revisions will have an obvious and immediate impact: (i) the measurement of semen volume to facilitate the calculation of total sperm number per ejaculate (see Section 2.3.4); (ii) the measurement of sperm motility (see Section 2.5); and (iii) the measurement of sperm morphology (see Section 2.12), now including a reference range for the fertile population (see Appendix 1). Each point is discussed in more detail below.

3.1 Measurement of semen volume

To date, the significance of measuring semen volume has probably been under-recognized by scientists and has not been given adequate thought with regard to QA and QC. However, the recognition that total sperm number per ejaculate may be a better predictor of pregnancy than sperm concentration (numbers per unit volume) implies that it should now be taken more seriously. Laboratory scientists have traditionally employed a variety of methods to measure semen volume, but the majority of the scientists have probably relied on methods based on decanting liquefied semen into a measuring device. However, the 5th edition recommends that it is preferable to infer volume from the sample weight, because the specific density of semen is sufficiently close to that of water (at 1 g per mL). This means that men providing samples for analysis will need to be given a pre-weighed container in order to collect their sample, which may cause problems for some laboratories. For example, the laboratory may not have a suitable balance available for this purpose (although this

would seem unlikely given that it is a standard piece of laboratory equipment). Additionally, there are still many laboratories that obtain samples from patients in containers provided by other health-care professionals, who may not recognize the importance of pre-weighing. The latter situation constitutes poor QA. Therefore, the requirement to work with pre-weighed specimen containers will not only improve the accuracy of semen volume measurements but also enable laboratories to take charge of and control the process of sample delivery by mandating the use of pre-weighed containers issued directly by the laboratory.

With regard to QC processes, the measurement of weight is arguably one of the easiest to control given that the science of measurement (metrology) is a discipline in its own right. There are three international organizations (the General Conference on Weights and Measures, the International Committee for Weights and Measures, and the International Bureau of Weights and Measures) created by the Convention du Mètre (1875) to oversee the measurement of weight to a standard kilogram. Andrology laboratories with appropriate service contracts in place will, in accordance with national laws and international conventions, have their laboratory balance regularly checked against the International Prototype Kilogram stored at the International Bureau of Weights and Measures (Sèvres, Paris, France). Additionally, the laboratories will fulfil QA by undertaking regular QC checks of the weighing process. For example, this could be performed through repetitive weighing of known volumes of semen (or water) dispensed into specimen containers from pipettes that have themselves been calibrated by weighing known volumes of water measured by a balance calibrated with the International Prototype Kilogram. Control charts should be used to monitor whether corrective action is required. This is QC at its simplest.

3.2 Measurement of sperm motility

To date, sperm motility has been reported by classifying spermatozoa into one of four movement patterns (a to d) and reporting the proportion of spermatozoa in each group as a percentage of the whole number observed. The four-grade system was underpinned by observations that only sperm with specific motility patterns were able to penetrate the cervical mucus (and therefore enter the female reproductive tract) [20, 21]. The advantage of measuring sperm motility in this way has been further supported by the use of computer-assisted sperm analysis (CASA), which has positively correlated the concentration of sperm swimming at > 25 μm per second with the probability of conception [22].

Despite this background, the 5th edition of the WHO manual proposes a simplified classification system to include only the following three grades:

(i) progressively motile (PM);
(ii) non-progressively motile (NP); and
(iii) immotile (IM).

This system is based on the argument that laboratory scientists cannot accurately define whether sperm are moving at above or below 25 μm per second [23], the threshold for two of the progressively motile grades defined in the 4th edition [10]. This view is also shared by delegates attending training courses in the UK organized by the Association of Biomedical Andrologists (http://www.aba.uk.net), who reported that the assessment of sperm as grade a or b was the most problematic of the practical techniques in semen analysis (personal observation). Perhaps this is unsurprising given that a human observer is being asked to replicate by eye the measurements first made using high-speed cinematographic analysis [20, 21] or commercially available CASA systems [22].

Perhaps a better question is to ask why, 20 years after CASA machines became available commercially, is this technology not more widely used by andrology laboratories for the measurement of sperm motility? There are several reasons: they can be expensive, they have their limitations, they require careful set-up, and they can be used for measurements of sperm motility only over a defined range of concentrations [24]. However, they are arguably easier for QC purposes than for a human observer because the same segments of videotape (or DVD) can be repeatedly run through them at regular intervals [25]. Although laboratory scientists can also observe videotapes or DVDs to check their performance (indeed, this is usually how external quality assurance specimens for sperm motility are distributed), they are not accustomed to looking at a TV screen in order to assess sperm motility and generally find it difficult. Therefore, of all the QC techniques in use in laboratory andrology, the QC of sperm motility is one of the most difficult to execute successfully. The proposed definition change in the revised manual will probably reduce frustration and instantly improve QC results. However, whether it improves the diagnostic power of semen analysis remains to be seen.

3.3 Measurement of sperm morphology

The most controversial aspect of the 4th edition of the WHO manual [10] was the lack of a reference range for sperm morphology in Appendix 1A. This caused confusion among laboratory scientists and clinicians alike, but, as stated in a footnote, there was a clear lack of multi-centre population-based studies on which to base a decision. Thankfully, those studies have now been carried out [26] and, using data obtained from semen samples of 4 500 men in 14 countries, reference intervals have been compiled for about 1 500 men whose partners had

a time to pregnancy of less than 12 months. The manual includes the 5th percentile (and 95% confidence interval) as the lower reference limit of the fertile population. The section also includes values for total motility (PR + NP), progressive motility (PR) and the value for total sperm number per ejaculate, as discussed above.

It will be a surprise to many that the data indicate that the lower reference limit for sperm morphology should be 4%, with a 95% confidence interval between 3% and 4%. This is much lower than the 14%–15% cutoff values observed during *in vitro* fertilization (IVF) [27, 28] and the 9% seen following intrauterine insemination (IUI) [29]. At 4%, such a low reference value will make QA and QC more difficult. This is, in part, because the measurement of sperm morphology has a binomial distribution (i.e., sperm are normal or abnormal) and the statistical confidence with which we can estimate the 'true' value of morphology from a sample is dependent not only on the number of spermatozoa observed but also on how close the true value is to zero [19]. Statistically, laboratories that choose to count fewer than 100 spermatozoa to evaluate sperm morphology [17, 18] will generate results with such wide confidence intervals that they will be almost meaningless for identifying men with sperm morphology above or below the 4% threshold. Even when following the recommendation that 400 spermatozoa per semen sample (two counts of 200 in duplicate) be assessed, making assertions at values close to this threshold will be a challenge.

4 QA and QC of other laboratory tests

Although the above discussion concentrates on the QA and QC implications of the main laboratory measures of sperm concentration, motility and morphology, the revised manual lists many more mandatory and optional tests. Although there is no space here to consider each one, it is worth recognizing the importance of QA and QC in any new tests that enter routine use.

The tests with perhaps the most potential for widespread introduction into andrology laboratories are those that examine the integrity of sperm DNA [30]. A recent systematic review and meta-analysis [31] has shown that, in 11 studies involving 1 549 cycles of IVF or intracytoplasmic sperm injection (ICSI), 640 pregnancies and 122 pregnancy losses, sperm DNA damage is predictive of pregnancy loss (odds ratio 2.48; 95% confidence interval 1.52–4.04). These results provide a clinical indication that such tests may be useful before IVF or ICSI. However, although many papers have been published on laboratory techniques to measure DNA damage, few have considered QA and QC aspects of the tests they are conducting. This needs to be addressed.

5 Final remarks

Although the implementation of comprehensive QA and QC programmes in the andrology laboratory will certainly be aided by the revised QA and QC chapter in the 5th edition, other factors will also be important in this endeavour. First, there must be a general realization within the profession that QA and QC are not optional parts of the laboratory repertoire but inherent parts of the modern medical laboratory, as they are in other disciplines [14]. This may be achieved by the publication of the revised manual itself, but also through articles, such as this one, that raise awareness of the issue and promote debate. Second, there must be increasing emphasis on investigating QA and QC by accrediting bodies and regulatory authorities as part of their surveillance of andrology laboratories. In countries where such organizations do not exist, they need to be developed. Finally, and crucially, appropriate and accessible basic training and continuing professional development must be available for laboratory scientists engaged in andrology to allow them to develop and maintain their skills. It has been shown that attending training courses can significantly improve the performance of individual scientists [32, 33]; yet, a common complaint is that too few such courses exist. Moreover, there is no consensus as to which teaching methods are the most effective. These aspects, too, warrant our attention.

Acknowledgment

I would like to thank Dr Julia Tindall (University of Bristol, Bristol, UK) for statistical help and Mr Robert Burgess (University of Manchester, UK) for proofreading.

References

1 Burnett D. Understanding Accreditation in Laboratory Medicine. London: ACB Venture Publications; 1996.
2 Whitehead TP, Browning DM, Gregory A. A comparative survey of the results of analyses of blood serum in clinical chemistry laboratories in the United Kingdom. J Clin Pathol 1973; 26: 435–45.
3 Bullock DG, Wilde CE. External quality assessment of urinary pregnancy oestrogen assay: further experience in the United Kingdom. Ann Clin Biochem 1985; 22: 273–82.
4 World Health Organisation. WHO Laboratory Manual for the Examination of Human Semen and Sperm-Cervical Mucus Interaction, 1st edn. Singapore: Press Concern; 1980.
5 World Health Organisation. WHO Laboratory Manual for the Examination of Human Semen and Sperm-Cervical Mucus Interaction, 2nd edn. Cambridge: Cambridge University Press; 1987. p67.
6 World Health Organisation. WHO Laboratory Manual for the Examination of Human Semen and Sperm-Cervical Mucus Interaction, 3rd edn. Cambridge: Cambridge University Press; 1992. p107.

7 Neuwinger J, Bejre H, Nieschlag E. External quality control in the andrology laboratory: an experimental multicenter trial. Fertil Steril 1990; 54: 308–14.
8 Matson PL. External quality assessment for semen analysis and sperm antibody detection: results of a pilot scheme. Hum Reprod 1995; 10: 620–5.
9 Cooper TG, Atkinson A, Nieschlag E. Experience with external quality control in spermatology. Hum Reprod 1999; 14: 765–9.
10 World Health Organisation. WHO Laboratory Manual for the Examination of Human Semen and Sperm-Cervical Mucus Interaction, 4th edn. Cambridge: Cambridge University Press; 1999. p128.
11 Jequier AM. Is quality assurance in semen analysis still really necessary? A clinician's viewpoint. Hum Reprod 2005; 20: 2039–42.
12 Holt WV. Is quality assurance in semen analysis still really necessary? A spermatologist's viewpoint. Hum Reprod 2005; 20: 2983–6.
13 Castilla JA, Alvarez C, Aguilar J, González-Varea C, Gonzalvo, MC, *et al.* Influence of analytical and biological variation on the clinical interpretation of seminal parameters. Hum Reprod 2005; 21: 847–51.
14 Pacey AA. Is quality assurance in semen analysis still really necessary? A view from the Andrology Laboratory. Hum Reprod 2006; 21: 1105–9.
15 Clinical Pathology Accreditation. Standards for the Medical Laboratory. Sheffield: Clinical Pathology Accreditation Ltd; 2007. p57.
16 Directive 2004/23/EC of the European Parliament and of the Council of 31 March 2004 on setting standards of quality and safety for the donation, procurement, testing, processing, preservation, storage and distribution of human tissues and cells. Official J Eur Union 2004; 47: L102. 48–58.
17 Keel BA, Sternbridge TW, Pineda G, Serafy NT. Lack of standardisation in performance of the semen analysis among laboratories in the United States. Fertil Steril 2002; 78: 603–8.
18 Riddell D, Pacey A, Whittington K. Lack of compliance in UK Andrology laboratories to World Health Organisation recommendations for sperm morphology assessment. Hum Reprod 2005; 20: 3441–5.
19 Kuster CE, Singer RS, Althouse GC. Determining sample size for the morphological assessment of sperm. Theriogenology 2004; 61: 691–703.
20 Aitken RJ, Sutton M, Warner P, Richardson DW. Relationship between the movement characteristics of human spermatozoa and their ability to penetrate cervical mucus and zona-free hamster oocytes. J Reprod Fertil 1985; 73: 441–9.
21 Katz DF, Morales P, Samuels SJ, Overstreet JW. Mechanisms of filtration of morphologically abnormal human sperm by cervical mucus. Fertil Steril 1990; 54: 513–6.
22 Larsen L, Scheike T, Jensen TK, Bonde JP, Ernst E, *et al*. Computer-assisted semen analysis parameters as predictors for fertility of men from the general population. Hum Reprod 2000; 15: 1562–7.
23 Cooper TG, Yeung CH. Computer-aided evaluation of assessment of 'grade a' spermatozoa by experienced technicians. Fertil Steril 2006; 85: 220–4.
24 Mortimer D, Aitken RJ, Mortimer ST, Pacey AA. Workshop report: clinical CASA—the quest for consensus. Reprod Fertil Dev 1995; 7: 951–9.
25 Clements S, Cooke ID, Barratt CL. Implementing comprehensive quality control in the andrology laboratory. Hum Reprod 1995; 10: 2096–106.
26 Cooper TG, Noonan E, von Eckardstein S, Auger J, Baker HW, *et al*. World Health Organization reference values for human semen characteristics. Hum Reprod Update 2009 Nov 24. [Epub ahead of print].
27 Kruger TF, Menkweld R, Stander FS, Lombard CJ. Sperm morphological features as a prognostic factor in IVF. Fertil Steril 1986; 46: 1118–23.
28 Grow DR, Oehninger S, Seltman HJ, Toner JP, Swanson RJ, *et al*. Sperm morphology as diagnosed by strict criteria: probing the impact of teratozoospermia on fertilization rate and pregnancy outcome in a large *in vitro* fertilization population. Fertil Steril 1994; 62: 559–65.
29 Guzick DS, Overstreet JW, Factor-Litvak P, Brazil CK, Nakajima ST, *et al*. Sperm morphology, motility, and concentration in fertile and infertile men. N Engl J Med 2001; 345: 1388–93.
30 Agarwal A, Said TM. Role of sperm chromatin abnormalities and DNA damage in male infertility. Hum Reprod Update 2003; 9: 331–45.
31 Zini A, Boman JM, Belzile E, Ciampi A. Sperm DNA damage is associated with an increased risk of pregnancy loss after IVF and ICSI: systematic review and meta-analysis. Hum Reprod 2008; 23: 2663–8.
32 Bjorndahl L, Barratt CL, Fraser LR, Kvist U, Mortimer D. ESHRE basic semen analysis courses 1995–1999: immediate beneficial of standardized training. Hum Reprod 2002; 17: 1299–305.
33 Franken DR, Kruger TF. Lessons learned from a sperm morphology quality control programme. Andrologia 2006; 38: 225–9.