Automation is the key to standardized semen analysis using the automated SQA-V sperm quality analyzer

Ashok Agarwal, Ph.D., H.C.L.D.,[a,b] *and Rakesh K. Sharma, Ph.D.*[a,b]

[a] Reproductive Research Center, Glickman Urological Institute; and [b] Department of Obstetrics-Gynecology, Cleveland Clinic Foundation, Cleveland, Ohio

Objective: To evaluate the performance of the automated semen quality analyzer system for assessing sperm quality.

Design: Double-blind prospective study.

Setting: Tertiary care hospital.

Patient(s): Fifty healthy men donated semen samples.

Intervention(s): None.

Main Outcome Measure(s): Precision, accuracy and agreement between automated and manual semen analysis methods was assessed for sperm concentration, motility, morphology, and known concentrations of latex bead quality control media.

Result(s): A good agreement was seen between the results of sperm concentration reported by the SQA-V automated analyzer (Spermalite/SQA-V; Medical Electronic Systems Ltd, Caesarea Industrial Park, Israel) and those obtained manually. A similar linearity was seen when the SQA-V results were compared with the manual data and also when the manual results of individual operators were compared with each other. The automated assessment of morphology showed high sensitivity (89.9%) for identifying percent normal morphology, and the precision of the SQA-V was considerably higher when compared with the manual method. The interoperator variability for manual assessment was significant. The automated analysis was quick compared with the manual method.

Conclusion(s): The SQA-V can be used interchangeably with manual semen analysis methods for examining sperm concentration and motility. The automated SQA-V analyzer is more precise and shows the ability to accurately classify normal versus abnormal sperm morphology. (Fertil Steril® 2007;87:156–62. ©2007 by American Society for Reproductive Medicine.)

Key Words: Male infertility, semen analysis, sperm concentration, motility, SQA-V

Semen analysis is the first tool a medical practitioner uses to assess the male factor in an infertility workup. Conventional microscopic examination of semen is prone to high variability and lack of standardization. Reporting an accurate manual semen analysis is fraught with difficulty because of a variety of factors (1–4). Human errors or inconsistencies influencing the accuracy of semen analysis are most often associated with counting, statistical errors, poor sample handling, lack of consistent adherence to protocols, and technician stress. This is further compounded by instrument variation and deterioration, varying protocols, and the nature of a time sensitive biologic sample.

It is a challenge to perform a single accurate assessment of the basic semen parameters such as sperm concentration, motility, and morphology. Because of the factors listed previously and the subjective nature of manual analysis, repeatability is poor and interoperator variability is high. The World Health Organization (WHO) manual (4) has attempted to standardize semen analysis and promote consistency and accuracy by recommending that 200 spermatozoa be counted in duplicate to enhance the repeatability by increasing the sample size. Even if this recommendation is followed for manual semen analysis, WHO permits a 20% difference between duplicate sperm counts.

Accuracy and precision can only be achieved by eliminating human error, adhering to an effective and standardized protocol, and assessing a very large sample size. It is evident that automation is a key factor to address all of these objectives. New automated methods for semen analysis are of clinical interest if the automated system is proven to accurately report semen analysis parameters (5–9).

The objectives of our study were:

1. to compare the sperm concentration, motility, and morphology results obtained manually by two independent operators to those run on a new automated sperm quality analyzer;
2. to assess the performance of the two methods for sensitivity, specificity, between-method, and interoperator agreement; and

3. to evaluate precision and accuracy of the automated and manual methods using both semen samples and a quality control material. Additionally, this study will evaluate the extent to which the SQA-V (Spermalite/SQA-V; Medical Electronic Systems Ltd, Caesarea Industrial Park, Israel) presents as a precise, rapid, and cost-effective alternative to existing automated semen analyzers that require extensive professional skill to operate effectively.

MATERIALS AND METHODS
Collection and Evaluation of Semen Samples

After approval from the institutional review board, semen samples from 50 healthy men were collected by masturbation after 3–5 days of sexual abstinence. After liquefaction, samples were evaluated for sperm concentration, motility, and morphology manually and using the SQA-V automated sperm quality analyzer (Spermalite/SQA-V; Medical Electronic Systems Ltd, Caesarea Industrial Park, Israel).

All manual assessments were performed in duplicate by two independent operators according to WHO guidelines (4). Sperm concentration was assessed using a Makler counting chamber (Seifi-Medical Instruments, Haifa, Israel) under a phase contrast microscope (Olympus BH2, Lake Success, NY) (magnification x200). Completely liquefied, nondiluted semen samples were heated at 55°C for 5 minutes to thoroughly mix and immobilize the spermatozoa.

A 5-μL aliquot was loaded onto the Makler chamber according to the manufacturer's instructions. Motility was scored under a phase contrast microscope using a standard glass slide and a simple grading system as defined by the WHO (4). Motile sperm concentration was calculated from the sperm concentration and motility results.

For quality control of sperm concentration, two known concentrations (45 $\pm$ 6.3 and 22 $\pm$ 3.1 million/mL) of QwikCheck beads (Medical Electronic Systems, Ltd) were analyzed using a phase contrast microscope and a Makler counting chamber. Morphology slides were prepared by air-drying smears and Diff-Quik staining (Baxter Healthcare Corporation, Inc, McGaw Parl, IL). The slides were scored for normal morphology by two independent operators according to WHO guidelines (10). The results of the two operators were averaged.

Automated Analysis of Semen and QwikCheck Beads

Automated semen analysis was conducted by an independent third operator in parallel to manual assesment using the SQA-V/SPERMALITE sperm quality analyzer (Medical Electronic Systems, Ltd, Caesarea Industrial Park, Israel) for sperm concentration, motility, and normal morphology. The SQA-V technology is based on the principle of electrooptical signal processing in combination with built-in computer algorithms (Fig. 1A).

Two independent channels, one measuring sperm concentration and the other measuring motile sperm concentration,

FIGURE 1

Diagram showing (**A**) the SQA-V Sperm Quality Analyzer with visualization system and (**B**) the SQA-V measurement capillary and the slide adapter used in the visualization compartment.
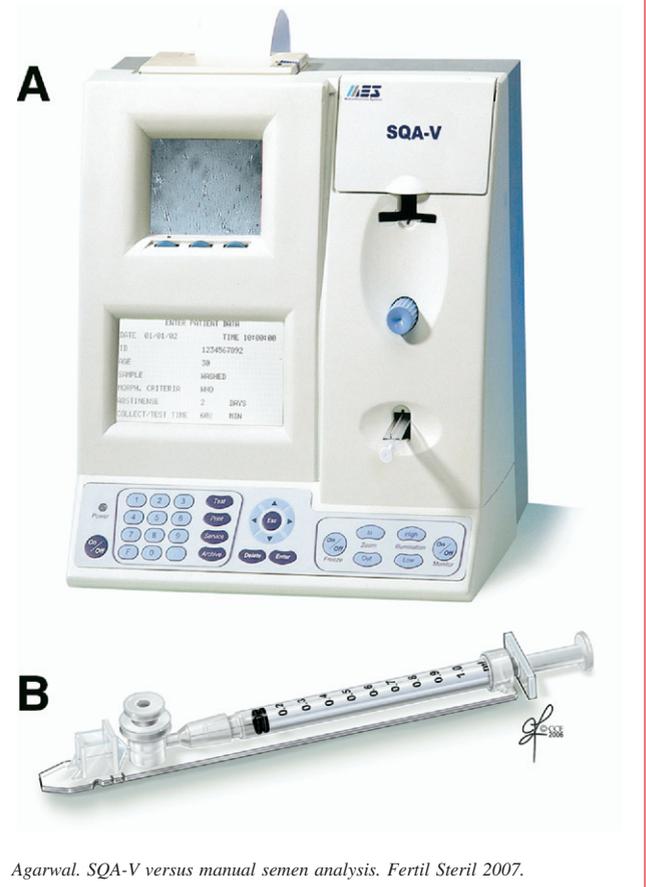
Agarwal. SQA-V versus manual semen analysis. Fertil Steril 2007.

transmit analogue signals for analysis. Following the SQA-V manufacturer's user guide instructions, a disposable testing capillary was filled with a fully liquefied, nondiluted, thoroughly mixed semen sample and run on the SQA-V at room temperature (Fig. 1B). For quality control, the same two known concentrations of QwikCheck beads used for manual assessment were also run on the SQA-V according to the manufacturer's instructions.

Statistical Analysis

Passing-Bablok regression analysis (11) was used to quantify the agreement between the two methods in this study because this is commonly used in studies requiring clinical analysis (12–15). We selected the above test based on the following considerations: [1] to quantify the agreement (or accuracy) between 2 methods, [2] to apply a more robust algorithm to address outliers, and [3] to optimally address the variance between methods that is commonly experienced in semen analysis studies.

The agreement between the two operators performing manual semen analysis was also examined. A logarithmic (log) transformation was applied to address deviations over the range of measurements (16, 17). The study results were further analyzed and mean values, SD, slope, intercept, and 95% confidence intervals were compared. Precision for both automated and manual methods was determined by comparing duplicate measurements run on each semen parameter and comparing the averaged coefficients of variation (CV). Precision was analyzed in the same manner using QwikCheck beads quality control material based on five replicate readings.

The SQA-V reports percent normal morphology without noting specific abnormalities. Therefore, and for the purposes of this study, morphology was graded as "normal" or "abnormal" for both the automated and manual assessment based on WHO (3rd edition) morphology criteria (10). The sensitivity and specificity of the automated morphology results versus manual data were compared. Interoperator sensitivity and specificity for morphology results were also evaluated.

RESULTS
Sperm Concentration

Table 1 summarizes the statistical assessment of sperm concentration and motility. Passing-Bablok regression plots for sperm concentration (after log transformation) are shown in Figures 2A–C. The mean ± SDs for the manual and automated sperm concentration are similar. There is good agreement when comparing the sperm concentration results obtained using the automated sperm quality analyzer to those obtained manually. Slight differences are observed in the intercept, slope, and 95% confidence interval. This is true for the interoperator results as well. The comparison between the SQA-V and the second operator shows the best intercept and slope (0.028 and 0.987, respectively). No significant linear deviation or statistically significant differences are seen for each operator when automated versus manual results were compared for sperm concentration. This is true when comparing the manual results of the two operators as well.

Percent Motility

The statistical evaluation of sperm motility summarized in Table 1 shows that the mean values of both manual operators are slightly higher than the mean values of the automated system; however, the standard deviations are similar. When the automated motility readings were compared with the test results obtained manually by two operators, a good agreement was seen and only marginal differences were found in the intercept, slope, and 95% confidence intervals of the slopes. This was true for the interoperator results as well. A comparison between the SQA-V and the second operator identified the best intercept and slope (0.111 and 0.90, respectively). No significant linear deviation was seen when the SQA-V motility results were compared with the first

TABLE 1
Comparison of automated SQA-V to first and second operator's results for sperm concentration and motility.

Method	Mean ± SD	Comparison	Intercept	Slope	95% CI of slope	Linearity
Concentration						
SQA-V	1.7334 ± 0.3165	SQA-V vs. first operator	0.1359	0.9412	0.8049–1.0833	$P>.10$[a]
First operator	1.6910 ± 0.3091	SQA-V vs. second operator	0.0281	0.9870	0.9010–1.0625	$P>.10$[a]
Second operator	1.7274 ± 0.2998	Second vs. first operator	0.1120	0.9500	0.8842–1.0366	$P>.10$[a]
Motility						
SQA-V	1.7122 ± 0.0990	SQA-V vs. first operator	0.6337	0.6283	0.4286–1.0000	$P>.10$[a]
First operator	1.7598 ± 0.1258	SQA-V vs. second operator	0.1110	0.9000	0.6316–1.3333	$P<.05$
Second operator	1.7922 ± 0.0933	Second vs. first operator	0.6300	0.6667	0.4839–0.9000	$P<.05$

Note: Passing-Bablok regression analysis was used; results were log transformed; SD = standard deviation; CI = confidence interval.
[a] No significant linear deviation.

Agarwal. SQA-V versus manual semen analysis. Fertil Steril 2007.

Passing-Bablok regression plots comparing automated (SQA-V) versus manual readings. (**A**) SQA-V analyzer versus first operator (mean ± SD values: 1.733 ± 0.316 and 1.691 ± 0.309 respectively), (**B**) SQA-V analyzer versus 2nd operator (mean ± SD values: 1.733 ± 0.316 and 1.727 ± 0.299, respectively), and (**C**) manual results of the first and second operator sperm concentration readings (mean ± SD values: 1.691 ± 0.309 and 1.727 ± 0.299 respectively. No significant linear deviation was observed (*P*>.10). (**D**) Accuracy and precision of the SQA-V analyzer and two independent operators using known concentration of control beads.
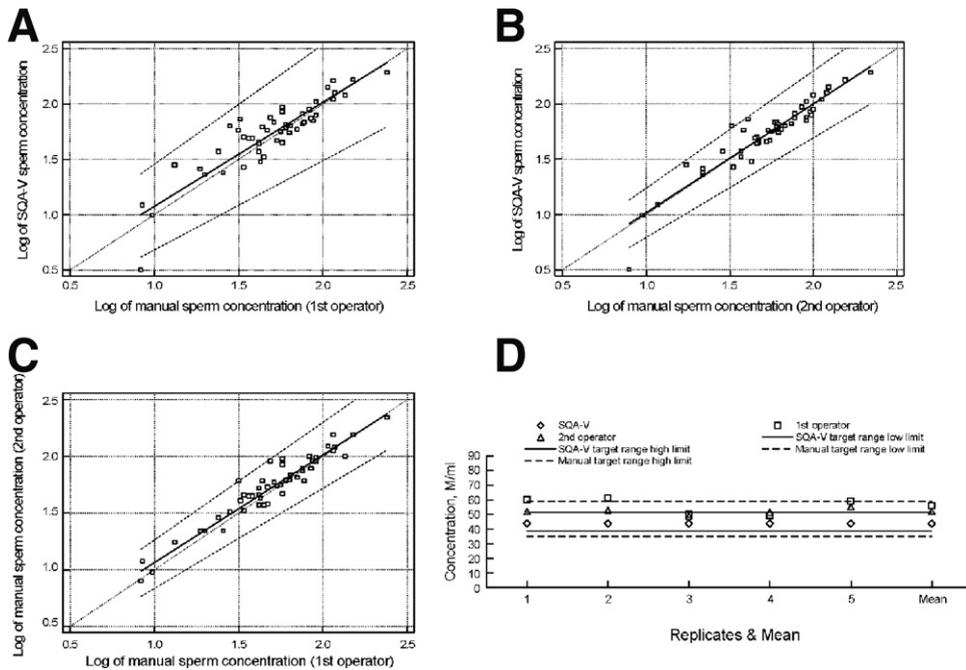
Agarwal. SQA-V versus manual semen analysis. Fertil Steril 2007.

operator, whereas a significant linear deviation was observed between the SQA-V and the second operator. A significant linear deviation was found between the two operators as well (Table 1).

Sperm Morphology

For the automated versus manual morphology comparison, the manual results of the two operators was averaged and qualifed as normal or abnormal based on WHO (3rd edition) criteria for assessing percent normal morphology (10). The morpholgy readings of the first and second operators were compared based on the same WHO criteria (Table 2). Morphology results were categorized as true-positive, true-negative, false-positive, and false-negative. The results for automated percent normal morphology reported by the SQA-V showed a sensitivity of 89.9% and a specificity of 50% when compared with the averaged manual results.

Sperm morphology results comparing the automated SQA-V device versus manual results and the results between two operators.

Comparison	Sensitivity (%)	Specificity (%)	FP + FN cases	Agreement (kappa)
SQA-V vs. manual	88.90	50.00	11	0.416[a]
First vs. second operator	78.80	58.80	14	0.376[b]

Note: FP = false positive; FN = false negative.
[a] Moderate agreement.
[b] Fair agreement.

Agarwal. SQA-V versus manual semen analysis. Fertil Steril 2007.

Interoperator sensitivity was 78.8% with 58.8% specificity. A better agreement was seen between the SQA-V and the manual method than between the first and second operators (Table 2). A lower number of false-positive ± false-negative cases were reported by SQA-V than between the two operators.

Precision

Table 3 shows precision data for automated and manual assessment of sperm concentration, motility, motile sperm concentration, normal morphology, and the two concentrations of control beads. CVs for the SQA-V were considerably lower than those obtained by the manual method. Figure 2D shows the precision of the SQA-V versus the manual method when control samples of known bead concentration were examined. The CVs of the SQA-V was zero compared with the CVs obtained by the manual method that ranged from 4.4% to 10.4% (Table 3). Replicates of control beads run on the SQA-V were equal and evenly spread within the target range compared with manual replicates, which were unequal and in some cases exceeded the target range (Fig. 2D).

DISCUSSION

Lack of standardization in semen analysis has been discussed in a number of publications (18–20). In their study, Kvist and Bjorndahl (18) noted that semen analysis has not gained the attention or benefited from the technological advancements that have impacted modern medicine and promoted good laboratory practices. The results of the American Association of Bioanalyists national proficiency testing

program showed that CVs in sperm counts ranged from 24% to 138%, with computer-assisted sperm analysis (CASA) displaying a lower overall CV (53% ± 8%) compared with manual methods (80% ± 9%) (19).

A wide variation in normal morphology results was reported in the same publication. The study concluded that an urgent need to improve the quality of semen analysis was in order. Similar conclusions were reached in the Brazil et al. (20) study. This study recommended the need to improve the quality of semen analysis and showed the critical requirement to standardized protocols and techniques through automation.

An automated system such as CASA partially addresses the need to improve quality and standardize protocols and techniques. Sidhu et al. (21) showed that CASA systems are reliable for sperm counts greater than 20×10^6/mL but post-thaw motility is generally underestimated. Knuth and Nieschlag (22) noted that sperm concentration can be overestimated and motility underestimated by CASA systems when specimens are contaminated with debris. Davis and Katz (23) reviewed CASA technology and found that in very low or highly concentrated specimens, counts and percent motility are not accurate.

Furthermore, the accuracy of reporting concentration is dependent on the number of frames analyzed and impacted by the presence or absence of debris in the specimen. Additionally, the study found that the ability to standardize semen analysis using CASA is impacted by the varying algorithms resident in different CASA systems (23).

In this study, we used semen from 50 healthy men with presumably normal semen specimens as well as the QwikCheck beads at 45 and 22×10^6/mL; however, in a clinical setting a majority of the samples are abnormal with poor sperm count and/poor motility. The accurate assessment of low sperm counts, poor motility, and increased abnormal forms is therefore critical. The dynamic range of the SQA-V in both fresh and washed samples is $2-400 \times 10^6$/mL. There are two systems available on the SQA-V: automated and visualization, which allows user the flexibility to analyze all types of semen samples. In addition, if the sample is of low quality, the sample is tested for an additional 2 minutes. Furthermore, if the sample has an extremely low number of motile and immotile cells as seen in post vasectomy sample, both automated and the visualization system of the SQA-V can be used with a very high accuracy for identifying motile and nonmotile cells. However, it is imperative that the manufacturers' protocol is strictly followed. In addition, in this version of the analyzer, the user has the opportunity to document test results by capturing and archiving a video clip of the postvasectomy sample using SQV-software.

Based on the publications reviewed it is evident that precision, accuracy, and standardization are still issues impacting manual and CASA. We evaluated a new system, the

SQA-V analyzer in our study. It is an automated system, the basic technology is different from CASA technology (signal processing as compared to image processing) and it requires no subjective calibration. This study compared the SQA-V with the manual results of two independent operators.

Our study shows that sperm concentration results obtained by the SQA-V are in agreement with manual results, and our findings are similar to those reported earlier by Akashi et al. (24) and Fuse et al. (25). We found no significant linear deviation when quantifying the agreement between the two methods. We can therefore conclude that for analyzing sperm concentration, the SQA-V can be used interchangeably with manual analysis. The advantages of the SQA-V are speed, objectivity, and highly repeatable results.

Scoring of motility manually is prone to overestimation. In our study, a significant linear deviation is seen between the SQA-V and the second operator as well as between the two operators. This may be attributed to the well documented fact that manual assessment of motility is subjective and generally overestimated because of the attraction of the eye to movement (4). This is further compounded when the sample has a highly motile sperm concentration. The SQA-V shows better motility statistics when compared to manual motility results. In addition, compared with the manual method, the results of the SQA-V are objective and rapid (a few minutes versus over 30 minutes).

The SQA-V only provides percent normal morphology results without quantifying specific abnormalities. As such, it is limited when compared with manual methodology where morphological defects need to be identified and quantified. Statistically, the agreement between the percent normal morphology readings of the SQA-V versus manual data is moderate by Altman classification (26). The interoperator agreement of normal morphology assessment is only fair. The SQA-V shows high sensitivity to accurately detect abnormal morpholgy and greater precision and speed compared with the manual method for determining percent normal morpholgy. Therefore, although limited, the SQA-V is useful as a screening tool for distinguishing between samples with normal versus abnormal morphology.

Precision is shown by repeatability. When comparing methods, it is relevant to assess the repeatability of each method. But, if one method shows considerable variability, the agreement between the two methods is destined to be poor even if the new method is perfect (17). In this study, we showed that the precision of the SQA-V is considerably higher compared with the manual method; CVs of sperm concentration assessed by the SQA-V and manually by first and by the second operator were 1.4%, 6.0%, and 5.1% respectively (Table 3). It was also found that the CV of the SQA-V in assessing sperm concentration (1.4%) was lower than the CVs of six other methods reported in the literature (27). Prathalingam et al. (27) compared 3 novel methods for assessing sperm concentration (flow cytometry, image analysis, and a fluorescent plate reader) with conventional

methods (hemacytometer, spectrophotometer, and Microcell counting chamber, Conception Technologies, San Diego, CA). The flow cytometry results showed the lowest CV of this study (2.3%), with the plate reader showing the highest CV (20.0%). Considering the results of our study, one might expect that the agreement between the automated and manual method would be less than optimum because of considerable differences in precision of the two methods. However, even when this is taken into consideration, good agreement between the automated and manual results is seen. In conclusion, this study shows that the SQA-V technology provides more precise, objective, and timely semen analysis compared with the manual method. Therefore, we conclude that the SQA-V qualifies as an automated system for standardizing semen analysis.

REFERENCES

1. Neuwinger J, Behre H, Nieschlag E. External quality control in the andrology laboratory: an experimental multicenter trial. Fertil Steril 1990;54:308–14.
2. Matson PL. External quality control assessment for semen analysis and sperm antibody detection: results of a pilot scheme. Hum Reprod 1995;10:620–5.
3. Keel BA, Stembridge TW, Pineda G, Serafy NT Sr. Lack of standardization in performance of the semen analysis among laboratories in the United States. Fertil Steril 2002;78:603–8.
4. World Health Organization. Laboratory manual for the examination of human semen and sperm-cervical mucus interaction. 4th ed. Cambridge University Press; New York, 1999.
5. Garrett C, Liu DY, Clarke GN, Rushford DD, Baker HW. Automated semen analysis: "zona pellucida preferred" sperm morphometry and straight-line velocity are related to pregnancy rate in subfertile couples. Hum Reprod 2003;18:1643–9.
6. Nafisi VR, Moradi MH, Nasr-Esfahani MH. A template matching algorithm for sperm tracking and classification. Physiol Meas 2005;26: 639–51.
7. Tsuji T, Okada H, Fujisawa M, Hamaguchi Y, Kamidono S. Automated sperm concentration analysis with a new flow cytometry-based device, S-FCM. Am J Clin Pathol 2002;117:401–8.
8. Misell LM, Holochwost D, Boban D, Santi N, Shefi S, Hellerstein MK, et al. A stable isotope-mass spectrometric method for measuring human spermatogenesis kinetics in vivo. J Urol 2006;175:242–6.
9. Cordelli E, Eleuteri P, Leter G, Rescia M, Spano M. Flow cytometry applications in the evaluation of sperm quality: semen analysis, sperm function and DNA integrity. Contraception 2005;72:273–9.
10. World Health Organization. Laboratory manual for the examination of human semen and sperm-cervical mucus interaction. 3rd ed. Cambridge University Press, 1992.
11. Passing H, Bablok W. A new biometrical procedure for testing the equality of measurements from two different analytical methods. Application of linear regression procedures for method comparison studies in clinical chemistry, part I. J Clin Chem Clin Biochem 1983; 21:709–20.
12. Holdt B, Lehmann JK, Schuff-Werner P. Comparative evaluation of two newly developed devices for capillary viscometry. Clin Hemorheol Microcirc 2005;33:379–87.
13. Solnica B, Naskalski JW. Quality control of SMBG in clinical practice. Scand J Clin Lab Invest Suppl 2005;240:80–5.
14. Hubner U, Pape G, Lotz J, Bertsch T, Kraft M, Geisel J, Herrmann W. Evaluation of alkaline phosphatase activity assay for the Reflotron system after substitution of the comparison method. Clin Lab 2005;51: 49–53.
15. Aulesa C, Mainar I, Prieto M, Cobos N, Galimany R. Use of the Advia 120 hematology analyzer in the differential cytologic analysis of bio-

logical fluids (cerebrospinal, peritoneal, pleural, pericardial, synovial, and others). Lab Hematol 2003;9:214–24.

16. Altman DG, Bland JM. Measurement in medicine: the analysis of method comparison studies. The Statistician 1983;32:307–17.

17. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986;(i):307–10.

18. Kvist U, Bjorndahl L. ESHRE monographs, manual on basic semen analysis. Oxford University Press; Oxford, UK, 2002:25.

19. Keel BA, Quinn P, Schmidt CF, Serafy NT Jr, Sarafy NT Sr, Schalue TK. Results of the American Association of Bioanalists national proficiency testing programme in andrology. Hum Reprod 2000;15:680–6.

20. Brazil C, Swan SH, Drobnis EZ, Liu F, Wang C, Redmon JB, et al. Standardized methods for semen evaluation in a multicenter research study. J Androl 2004;25:635–44.

21. Sidhu RS, Sharma RK, Lee JC, Agarwal A. Accuracy of computer-assisted semen analysis in prefreeze and post-thaw specimens with high and low sperm counts and motility. Urology 1998;51:306–12.

22. Knuth UA, Nieschlag E. Comparison of computerized semen analysis with the conventional procedure in 322 patients. Fertil Steril 1988;49: 881–5.

23. Davis RO, Katz DF. Operational standards for CASA instruments. J Androl 1993;14:385–94.

24. Akashi T, Mizuno I, Okumura A, Fuse H. Usefulness of sperm quality analyzer-CV (SQA-V) for the assessment of sperm quality in infertile men. Arch Androl 2005;51:437–42.

25. Fuse H, Akashi T, Nozaki R, Nishio R, Mizuno I. Assessment off sperm quality analyzer II B: comparison with manual semen analysis and CASA. Arch Androl 2005;51:65–7.

26. Altman DG. Practical statistics for medical research. New York: Chapman and Hall, 1991:611.

27. Prathalingam NS, Holt WW, Revell SG, Jones S, Watson PF. The precision and accuracy of six different methods to determine sperm concentration. J Androl 2006;27:257–62.