

Double-blind prospective study comparing two automated sperm analyzers versus manual semen assessment

J. Lammers · C. Spingart · P. Barrière · M. Jean · T. Fréour

Received: 30 July 2013 / Accepted: 1 November 2013 / Published online: 16 November 2013
© Springer Science+Business Media New York 2013

Abstract

Purpose Despite controversy regarding its clinical value, male fertility investigation mainly relies on semen analysis. Even though reference guidelines are available, manual sperm analysis still suffers from analytical variability, thus questioning the interest of automated sperm analysis systems. The aim of this study is to compare automated computerized semen analysis systems (SQA-V GOLD and CASA CEROS) to the conventional manual method in terms of accuracy and precision.

Methods We included 250 men in this double-blind prospective study. The SQA-V GOLD (Medical Electronic Systems) and CEROS, CASA system (Hamilton Thorne) were compared to the standard manual assessment based on the WHO 5th Edition. The main outcome measures were sperm concentration, total sperm number, total motility, progressive motility, non-progressive motility, morphology, motile sperm concentration (MSC) and progressively motile sperm concentration (PMSC) with the three methods.

Results Statistical analysis of the test results from the automated systems and the manual method demonstrated no significant differences for most of the semen parameters. The Spearman coefficients of rank correlation (ρ) for CASA and the SQA-V GOLD automated systems vs. the manual method were: Sperm concentration (0.95 and 0.95), total sperm number (0.95 and 0.95), MSC (0.94 and 0.96) and PMSC (0.94

and 0.93) correspondingly. Concerning sperm morphology, both automated systems demonstrated high specificity (Sp) and negative predictive values (NPV), despite significantly different medians (CASA: 83.7 % for Sp and 95.2 % for NPV, SQA-V: 97.9 % for Sp and 92.5 %). The highest precision (lowest 95 % confidence interval for duplicate tests) for all semen variables was found in the SQA-V GOLD.

Conclusions The advantages of using automated semen analysers are: Standardization, speed (lower turnaround time), precision, reduced potential for human error, automated data recording and less need for highly skilled professionals to run the systems. The disadvantages of using automated systems are: notably the problem with testing some atypical samples and the inability to perform an assessment of morphology abnormalities. Based on the results of this study, the SQA-V Gold demonstrated better agreement vs. the manual method. In conclusion, automated semen analyzers can be used for routine semen analysis providing rapid clinically acceptable results with higher precision, and positively impacting laboratory standardization.

Keywords Male infertility · Semen analysis · SQA-V GOLD · CASA · WHO 5th manual for sperm analysis

Introduction

More than 15 % of couples trying to conceive will confront infertility issues during their reproductive life. Causes of men infertility are numerous; genetic, environmental exposure or lifestyle. Semen analysis is one of the first diagnostic tools used to quickly and effectively evaluate the male infertility factor, even if its real clinical value is still questionable [1, 2]. Most laboratories manually assess semen under a microscope. However, conventional manual semen analysis has limitations, as illustrated by studies demonstrating inter-lab variability engendering discordant results that can cause inappropriate diagnosis or delayed treatment of infertile couples [3]. Keel

Capsule The SQA-V GOLD and CASA automated systems were compared to manual semen analysis. Both systems can be considered reliable analytical tools for routine semen analysis, providing high quality results with improved laboratory standardization.

J. Lammers · C. Spingart · P. Barrière · M. Jean · T. Fréour
Service de Médecine et Biologie de la Reproduction,
University Hospital of Nantes, 44093 Nantes Cedex, France

T. Fréour (✉)
Hôpital Mère et Enfant, CHU de Nantes, 38 boulevard Jean Monnet,
44093 Nantes Cedex, France
e-mail: thomas.freour@chu-nantes.fr

et al. (2000) analyzed the results of the American Association of Bioanalysts national proficiency testing programme in andrology [4]. The majority of the labs (79 %) performed manual semen assessment. A wide inter-lab variation in concentration and percent normal morphology was found in the reports. However, it is possible to achieve good inter-laboratory agreement in semen analysis, provided the technicians undergo regular training and participate in an ongoing program of external quality and internal quality control [5, 6]. Recently, the World Health Organization published the 5th edition manual for sperm analysis (WHO 5th Edition) [7]. These guidelines encompass the results of studies conducted at laboratories worldwide. The objective of these published guidelines is to promote standardization of semen analysis through adherence to technical procedures, to emphasize the use of newly established reference values and to establish in-depth guidance for acceptable differences for duplicate measurements [8]. The introduction of automated sperm analyzers into the market over the last two decades demonstrated that they could be considered as an alternative approach to routine manual semen analysis that can promote laboratory standardization. Initially these systems demonstrated difficulty with accurately reporting sperm concentration due to the presence of round cells or debris. New systems are capable of measuring sperm motility and kinematics and the accuracy of reporting sperm concentration in the presence of debris or round cells has improved. These systems are pre-calibrated vs. manual analysis with strict conformity to WHO 5th Edition and follow regular quality control protocols [9]. Several studies have demonstrated that automated sperm analyzers could provide an accurate and highly correlated alternative to manual sperm analysis [9–17]. Basically, there are two categories of automated sperm analyzers on the market which can be characterized by their detection technology. The SQA-V GOLD is a fully automated system, which is based on the detection of electro-optical signals generated by moving spermatozoa and interpreted by proprietary algorithms. This signal processing for sperm motility is coupled with spectrophotometry technology to determine sperm concentration [9, 12, 15]. The Computer Assisted Sperm Analysis (CASA) systems are based on another principle: capturing microscopic images and image processing to detect motile and immotile spermatozoa through acquisition of rapid and successive frames [18–20].

A prospective double-blind study of this scope, comparing semen analysis results obtained manually and through two automated technologies, and based on WHO 5th criteria has yet to be published. The present study encompasses all of these characteristics, focusing on the analytical performance of each method and the potential for integrating automated sperm analysis into laboratory practice. In light of the move to promote higher levels of accuracy in semen analysis through stricter adherence to WHO 5th Edition, the rationale of the present study is to evaluate the precision (95 % confidence

intervals) and agreement of automated and manual semen analysis based on these guidelines and to present the advantages and disadvantages of each method.

Materials and methods

Sperm collection and pre-analytical workup

All men presenting at our andrology laboratory between February and April 2011 for semen analysis as a part of an infertility workup were prospectively included in the study. Untreated semen specimens with ejaculate volumes of >2.5 ml were included in the trial. The samples were split in order to assess the three methods: SQA-V GOLD, CASA and manual evaluation. Ejaculates were collected by masturbation at the laboratory after 2–5 days of sexual abstinence. After liquefaction for 30–45 min at room temperature, viscosity and pH (pH strips, Merck, Darmstadt, Germany) were evaluated. White blood cells ($>1 \times 10^6/\text{ml}$) were detected using QwickCheck™ Test Strips (Medical Electronic Systems). The volume of the ejaculate was then measured with a graduated pipette. To prepare samples for manual morphology examination, a 0.5 to 1 ml aliquot of the ejaculate was washed in 3 ml of Ferticult® (Fertipro, Belgium) and centrifuged for 10 min at 600 g (2,000 rpm). The pellet was then re-suspended in Ferticult® and 2 smears were prepared, air dried and stained following the conventional Shorr procedure [7].

Sperm analysis

Sperm analysis, including sperm concentration, total sperm number, % progressive motility (PR), % non-progressive motility (NP), total motility (PR+NP), motile sperm concentration (MSC), progressively motile sperm concentration (PMSC) and normal morphology (WHO 5th manual strict criteria, 4 % threshold), was performed simultaneously and independently by two operators manually, on the SQA-V GOLD (Sperm Quality Analyzer, Medical Electronic Systems, Los Angeles, CA, USA) and on the Computer-Assisted Sperm Analysis (CASA, CEROS Sperm Analyzer version 12.2 L, Hamilton Thorne, Beverly, MA, USA).

Manual sperm analysis was performed by two independent and highly trained technicians following the WHO 5th Edition [7]. These two operators performed manual sperm analysis simultaneously on 2 separate microscopes, and reported their results on two separate forms, in order to avoid any communication that would have questioned the double-blind assessment. The intra-operator agreement of the duplicate measurements was within the acceptable WHO 5th range or sample testing was repeated. Sperm motility was evaluated at room temperature by counting at least 200 spermatozoa in duplicate under a phase contrast microscope with magnification $\times 400$

(Zeiss Axioskop40). Sperm concentration was assessed by counting a minimum of 200 spermatozoa in duplicate using a Thoma counting chamber after dilution in a fixative solution. Normal morphology was assessed following slide preparation using Shorr stained smears. Proficiency and quality control challenges for concentration, motility and morphology are performed by all technicians on a routine basis. An external accreditation entity provides known challenges which include video and stained slides on a bi-monthly basis and technicians run fresh sperm on a monthly basis to comply with internal laboratory quality control protocols.

Automated sperm analysis for all semen parameters was performed in duplicate on the SQA-V GOLD following the manufacturer's guidelines. After entering patient and sample information, a disposable testing capillary (Medical Electronic Systems) is filled with 0.5 ml of undiluted and homogeneously mixed sperm sample was inserted into the SQA-V GOLD measurement chamber, and tested twice. To effectively run the SQA-V, the operator does not have to be highly trained. This is due to the fact that the automated system is easy to use and thus can be readily integrated into the laboratory for routine analysis. Quality controls are run daily on the SQA-V prior to semen testing with the manufacturer's quality control kit.

CASA analysis was performed in duplicate according to the manufacturer's instructions and WHO 5th Edition recommendations to assess a minimum of 1,000 cells. Operators need to be trained to use this automated system but not necessarily highly trained in semen analysis. The settings used in the laboratory are; 60 Hz frames per seconds and 30 frames for image capture. For cell detection the minimum contrast is 80 and the minimum cell size is 3 pixels. Progressive motility settings are 25.0 μ /s for path velocity (VAP) and 80.0 % for straightness (STR). Slow motility settings are 5.0 μ /s for VAP cut-off and 11.0 μ /s for VSL cut-off and the default (if <5 motile cells) sperm cell size is set to 6 pixels with cell intensity at 150. A 7 μ L sample of sperm was loaded into 2 separate disposable analysis chambers with a depth of 20 μ m (Leja Products, Nieuw-Vennep, Netherlands). The chambers were then placed on the plate of the HTM-CEROS for analysis. Sperm cells with an average path velocity (VAP) >5 μ m/s were classified as progressively motile cells. Stained smears were prepared for morphology analysis and assessed under 1,000 \times magnification using Dimensions software (Hamilton Thorne). Quality controls were run per the manufacturer's recommendations.

Statistical analysis

To compare the different methodologies, a gold standard was required. In spite of its known limitations, the manual method was selected as it is the standard laboratory practice and is recommended by the WHO 5th Edition. Bearing in mind that semen variables do not generally follow a Gaussian distribution, the **accuracy** (degree of the measurements proximity) of

the two automated analysers vs. the manual gold standard (true value) was assessed using multiple statistical approaches: Comparison medians and 95 % confidence intervals for medians, analysis of significant systematic differences, rank correlation between the semen variables reported by the automated systems vs. the gold standard and Receiver Operating Characteristic (ROC) analysis (where applicable). Specifically, ROC analysis was performed to calculate both specificity (Sp) and negative predictive value (NPV) of the SQA-V GOLD and CASA normal morphology vs. the manual gold standard. This statistical approach was used because, when applying WHO 5th Edition criteria, the narrow range of normal morphology results accompanied by the high variability of manual morphology assessment, minimizes the validity of other statistical approaches.

Mountain Plot is one of the statistical methods utilized to analyse the results of the present study (MedCalc software version 11.1.1.0 Belgium) for the following reasons: (1) Two or three laboratory assays can be compared with this method; (2) It is easier to find the central 95 % of the data (even if the data is not distributed normally); (3) different distributions can be more readily compared. A mountain plot (or 'folded empirical cumulative distribution plot') is created by computing a percentile for each ranked difference between the new and the reference method. The Mountain Plot presents, in a percentile scale, the differences in results reported by the two methods vs. the gold standard. In this study, the Mountain Plots method was used to demonstrate the percentile distribution of the SQA-V GOLD and CASA differences vs. the manual method.

Precision, also referred to as reproducibility or repeatability, demonstrates reliability as it is the degree to which repeated measurements show the same results under unchanged conditions. The precision of each method was investigated by calculating the coefficients of variation (CV) between two operators and between the duplicate measurements of each automated method. The use of multiple statistical approaches to assess accuracy and precision results in reliable conclusions concerning the level of agreement between the different automated methods and the gold standard. Additionally, the pros and cons of each method can be demonstrated. Statistical analysis calculations were performed utilizing the MedCalc software version 11.1.1.0 (Belgium).

Results

A total of 250 men with a mean age of 33.7 \pm 5.7 years (mean \pm SD) were included in this study. Abstinence of intercourse prior to sperm analysis was 4.1 \pm 3.1 days. The delay between semen collection and initiating sperm analysis was 37.7 min \pm 12.4. The turnaround time (TAT) for manual semen analysis is approximately 30–40 min (20 min more are

required to prepare stained smears for morphology assessment). The TAT for the CASA evaluation of concentration and motility is approximately 10 min with an additional 30 min required for morphological automated evaluation (40 min more are required to prepare stained smears for morphology assessment). The SQA-V completes the testing of normozoospermic samples in approximately 1 min and requires approximately 3–4 min for low quality samples.

Of the 250 semen samples, 31 demonstrated severe oligozoospermic ($<5 \times 10^6/\text{ml}$), 48 were mild to moderate oligozoospermic ($5\text{--}15 \times 10^6/\text{ml}$), and the rest were normal ($>15 \times 10^6/\text{ml}$). Four samples were excluded due to testing issues. Tables 1, 2, 3 and 4 illustrate the median values and 95 % confidence intervals for the medians of the semen parameters tested in each method. The 95 % confidence intervals for repeatability are presented as well. The overall results (Table 1) are sub-divided into three groups (Tables 2, 3 and 4) based on the range of sperm concentration: Severe oligozoospermia, mild to moderate oligozoospermia and normozoospermic group according to WHO 5th Edition. In the sub-groups, three main semen variables: sperm concentration, total motility and normal morphology are compared. Medians and 95 % confidence intervals could not be reliably expressed for progressively motile sperm in the severe and mild to moderate oligospermic groups due to high statistical counting errors which were the result of the low number of progressively motile sperm in the samples and the low number of representative samples in these groups.

The statistical analysis demonstrated that the median values for normal morphology reported by the SQA-V GOLD and CASA were significantly different vs. the manual results (Table 1). The median values of the other semen variables were in close alignment for all three methods demonstrating no significant systematic discrepancies. The highest 95 % confidence interval values were found in the manual between-operator variability and CASA duplicates. The duplicate readings of the SQA-V GOLD analyser demonstrated the lowest 95 % confidence intervals for all semen variables (Table 1).

In the severe oligozoospermic group (Table 2, $n=31$), the median value of sperm concentration reported by CASA was significantly higher than the median of the manual assessment. The medians of total motility reported by CASA and the SQA-V GOLD were significantly lower ($p < 0.05$) than those of the manual assessment. Note that the CASA median for sperm concentration was more than 4 times higher and the median for total motility more than 2 times lower than the median for manual assessment. Normal morphology values reported by the CASA and the SQA-V GOLD are in line with the manual results. The highest repeatability was shown by the SQA-V GOLD analyser and the lowest repeatability was shown by the CASA.

In the mild to moderate oligozoospermic group (Table 3, $n=48$), the median values of sperm concentration reported by

Table 1 Overall medians and 95 % confidence intervals ($n=246$)

Semen parameter	Median					95 % confidence interval for repeatability					
	Manual: Operator 1	Manual: Operator 2	Manual average	95 % CI for manual median	CASA	95 % CI for CASA median	SQA-V GOLD	95 % CI for SQA-V median	Between-operator	CASA duplicates	SQA-V GOLD duplicates
Sperm Concentration, $\times 10^6/\text{ml}$	28.0	27.4	28.1	23.4–34.1	32.0	28.4–38.9	32.6	26.4–36.3	6.9	6.0	1.7
Total Sperm Number, $\times 10^6/\text{ejaculate}$	107.5	107.1	110.6	91.2–124.9	128.1	110.0–143.9	119.7	98.8–137.1	25.6	22.4	6.8
Total Motility (PR+NP), %	57.0	59.5	58.3	55.7–60.2	57.0	51.8–63.7	54.7	52.4–57.5	10.5	12.1	3.9
Progressive Motility (PR), %	40.0	40.8	40.6	38.7–43.0	40.8	36.9–44.6	40.1	37.1–42.4	8.4	5.4	3.4
Non-progressive Motility (NP), %	17.0	18.0	17.3	16.5–18.0	17.3	15.9–18.5	16.0	14.6–17.2	7.8	9.6	2.4
Normal Morphology, %	7.0	7.0	7.0	6.4–7.9	5.0*	4.5–6.0	10.6*	9.6–11.1	3.5	3.6	1.4
Motile Sperm Concentration (MSC), $\times 10^6/\text{ml}$	15.8	16.2	15.6	12.8–20.2	18.2	14.1–22.7	19.2	15.9–22.9	5.5	8.1	1.8
Progressively Motile Sperm Concentration (PMSC), $\times 10^6/\text{ml}$	12.4	13.0	13.3	9.9–16.5	13.9	11.0–17.7	15.3	11.9–17.8	4.1	4.4	1.6

*Significantly different than manual results ($p < 0.05$)

Table 2 Medians and 95 % confidence intervals in the severe oligozoospermic group (sperm concentration $<5 \times 10^6/\text{ml}$; $n=31$)

Semen parameter	Median										
	Manual: Operator 1	Manual: Operator 2	Manual average	95 % CI for manual median	CASA	95 % CI for CASA median	SQA-V GOLD	95 % CI for SQA-V median	Between-operator	CASA duplicates	SQA-V GOLD duplicates
Sperm Concentration, $\times 10^6/\text{ml}$	1.4	1.6	1.6	1.1–2.5	7.4*	5.6–9.2	2.0	1.0–4.9	0.7	3.7	0.4
Total Motility (PR+NP), %	45.2	47.0	48	38.5–55.8	22.0*	10.1–30.0	38.9*	12.9–44.8	9.4	16.8	7.6
Normal Morphology, %	3.0	3.0	3.0	2.3–4.5	3.0	0.6–6.0	2.8	2.65–3.8	3.3	3.3	1.2

*Significantly different than manual results ($p < 0.05$)

Table 3 Medians and 95 % confidence intervals in the mild to moderate oligozoospermic group (sperm concentration $5–15 \times 10^6/\text{ml}$; $n=48$)

Semen parameter	Median										
	Manual: Operator 1	Manual: Operator 2	Manual average	95 % CI for manual median	CASA	95 % CI for CASA median	SQA-V GOLD	95 % CI for SQA-V median	Between-operator	CASA duplicates	SQA-V GOLD duplicates
Sperm Concentration, $\times 10^6/\text{ml}$	8.1	8.8	8.9	7.4–10.7	14.2*	11.9–15.9	10.8*	9.2–13.6	2.5	3.4	1.7
Total Motility (PR+NP), %	49.0	53.0	50.5	47.1–54.5	39.3*	31.2–47.0	49.6	45.4–52.3	12.3	14.8	6.2
Normal Morphology, %	5.0	4.0	4.0	3.3–5.4	3.0	2.0–4.5	4.5	3.5–5.8	2.7	3.1	1.0

*Significantly different than manual results ($p < 0.05$)

Table 4 Medians and 95 % confidence intervals in the normozoospermic group (sperm concentration $> 15 \times 10^6/\text{ml}$; $n = 167$)

Semen parameter	Median			95 % confidence interval for repeatability							
	Manual: Operator 1	Manual: Operator 2	Manual average	95 % CI for manual median	CASA	95 % CI for CASA median	SQA-V GOLD	95 % CI for SQA-V median	Between-operator	CASA duplicates	SQA-V GOLD duplicates
Sperm Concentration, $\times 10^6/\text{ml}$	42.7	44.6	43.6	37.9–49.7	49.3	44.0–60.0	48.1	42.5–52.7	9.4	7.1	1.9
Total Motility (PR+NP), %	61.0	62.4	62.0	59.8–64.3	74.3*	66.0–78.0	58.7	56.6–60.8	10.1	10.3	3.1
Normal Morphology, %	7.8	8.0	8.0	7.3–8.8	6.0*	5.0–6.5	12.0*	10.9–13.5	3.7	3.8	1.5

*Significantly different than manual results ($p < 0.05$)

CASA and SQA-V GOLD were significantly higher and total motility reported by CASA was significantly lower ($p < 0.05$) than the median values of the manual assessment. Normal morphology values reported by the CASA and SQA-V GOLD were not significantly different vs. the manual results. Between the two automated analyzers, the magnitude of discrepancies in the severe and mild to moderate oligozoospermic groups was much greater in the CASA system. Again, the highest repeatability in this group was shown by the SQA-V GOLD and the lowest by the CASA.

In the normozoospermic group (Table 4, $n = 167$), the median values of sperm concentration and total motility reported by CASA and SQA-V GOLD were not significantly different than manual results. However, total motility reported by the CASA and normal morphology reported by the CASA and the SQA-V GOLD were significantly different vs. the manual results. In 49 cases (19.9 %) CASA considered normal manual clinical results as abnormal vs. the SQA-V which considered 13 cases (5.3 %) of normal manual assessment to be abnormal. In these cases, manual analysis is considered the standard. The highest repeatability in this group was demonstrated by the SQA-V GOLD analyser.

In order to estimate the clinical impact of morphology discrepancies, ROC analysis was used to calculate morphology specificity and negative predictive value for the SQA-V GOLD and CASA versus manual assessment, resulting in the following outcome: $Sp_{SQA} 97.9 \%$, $NPV_{SQA} 92.5 \%$ and $Sp_{CASA} 83.7 \%$, $NPV_{CASA} 95.2 \%$ respectively. Due to the small number of positive cases, i.e. samples with $< 4 \%$ normal morphology, Sensitivity and Positive Predictive Value (PPV) could not be reliably calculated. The SQA-V GOLD and CASA Spearman's coefficients of rank correlation (ρ) vs. manual analysis results were as follows: Sperm concentration 0.95 and 0.95; total sperm number 0.95 and 0.95; MSC 0.94 and 0.96; and PMSC 0.94 and 0.93. The overall data distribution percentile mountain plots of the SQA-V GOLD and CASA differences vs. the manual method, demonstrated that the plot median points are close to zero (Fig. 1a–d). This indicates an absence of systematic discrepancies between the compared methods. However, the distribution plots of CASA motility, MSC and PMSC showed asymmetric tails at higher levels and larger plot areas than the SQA-V GOLD. This indicates that CASA has a tendency to report larger random motility discrepancies vs. the manual method as seen in the normozoospermic group (Fig. 1b–d).

Discussion

Though the relationship between sperm assessment and accurately diagnosing male infertility or predicting conception remains ambiguous [2, 21–23], semen analysis is the first diagnostic tool used to determine a contributing male factor

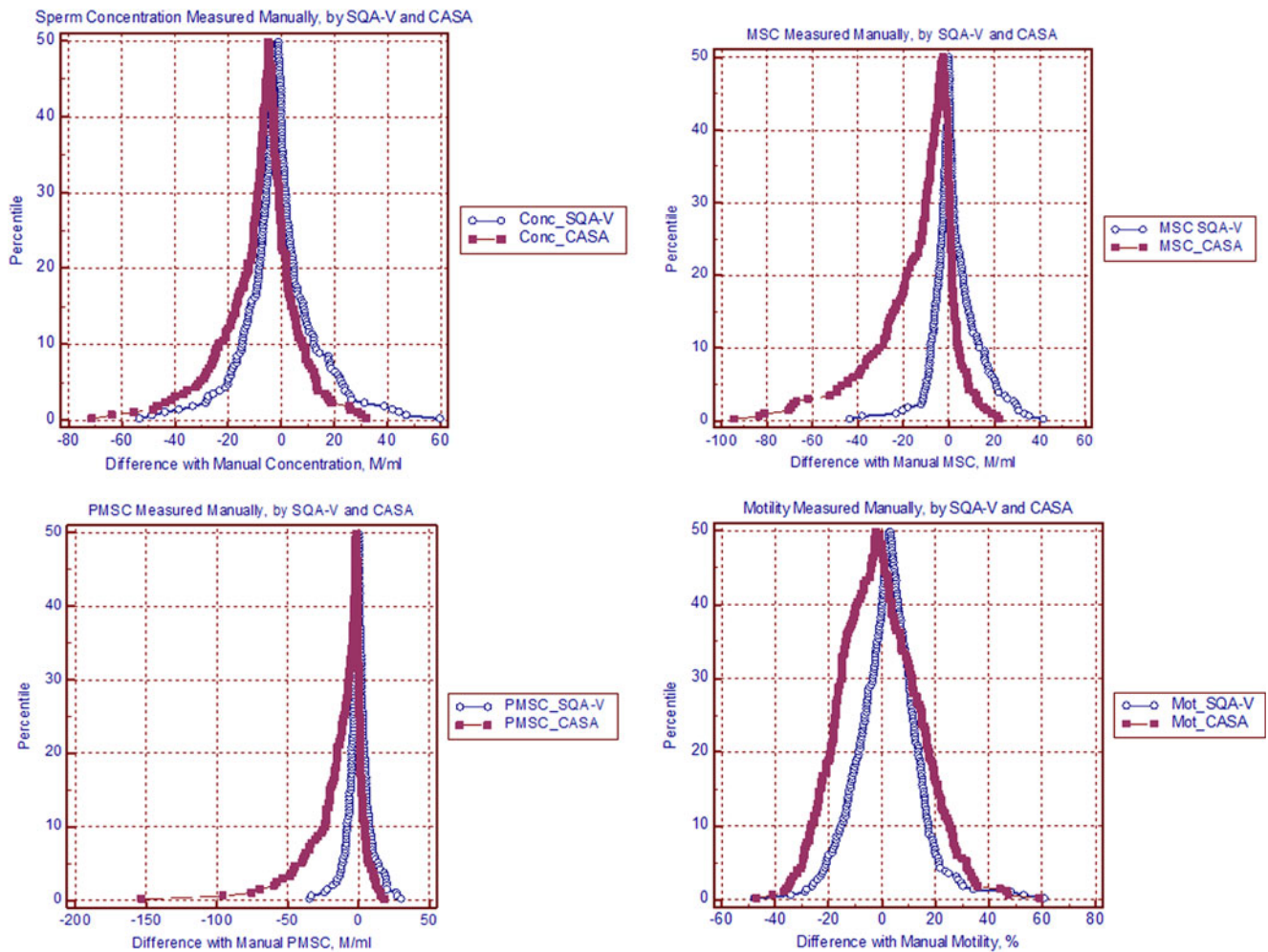


Fig. 1 Mountain Plots showing the percentile distribution of the CASA and SQA-V GOLD differences vs. the manual method for Concentration (a), Motile Sperm Concentration (b), Progressively Motile Sperm Concentration (c) and percent Motility (d)

in an infertile couple. Manual semen analysis, based largely on the evaluation of sperm concentration, motility and morphology, is known to suffer from lack of standardization, repeatability and accuracy [4].

In an attempt to address the problems with manual semen analysis, the WHO 5th has made continuous efforts to promote standardizing semen analysis, particularly with the recently published WHO 5th Edition manual for semen examination [7]. In this manual, automated semen analysis as a clinical laboratory practice is presented in a special section devoted to describing CASA semen assessment methodology.

Two basic types of automated sperm analysis systems have been developed and evaluated over the last 20 years. These systems have demonstrated accuracy, precision and objectivity resulting in an opportunity for laboratories to standardize semen analysis [9, 13].

In the present study, both the SQA-V GOLD and CASA demonstrated a high correlation with no significant discrepancies in most of the semen parameters versus the manual method which adhered to WHO 5th protocol compliance and

standardized sample handling conditions (Table 1). We found that the morphology results of the two automated systems are in line with the manual results in the two most clinically significant groups, severe and mild to moderate oligozoospermia. The significant differences in morphology are seen in the normozoospermic group only, though the medians of the data reported by all three methods are well above the 4 % reference value cut-off. These differences are quite minor when considering that the assessment of this parameter is highly variable due to the complexity and subjectivity of manual morphology and the WHO 5th normal morphology cut-off of 4 % [24]. Low normal morphology median values coupled with the low precision of the manual and CASA assessment (95 % confidence intervals are: Manual 3.7 and CASA 3.8 vs. SQA-V GOLD 1.5) resulted in the significant differences that were reported. Additionally, differences in technologies (electro-optical signal versus image processing) may contribute to the morphology discrepancies between the two automated systems. For example, the SQA-V GOLD determines morphology based on the detection of electro-optical signal

patterns produced by moving sperm cells. These signals are then interpreted by algorithms and % normal morphology is reported. CASA and manual morphology are determined using computerized and visual interpretation of cell images which are highly dependent on the quality of the cell staining preparation and the assessment of the shape of the cell. To optimally characterize the clinical performance of each method by ROC analysis, specificity and negative predictive values were calculated for normal morphology. Both the SQA-V GOLD and CASA demonstrated very high specificity and negative predictive values for normal morphology versus manual assessment. It is therefore concluded that both automated systems are clinically acceptable for assessing normal morphology, as the risk of inappropriate classification of the semen sample, and eventually subsequent delay in treatment, is very low.

Concerning the significant differences in sperm concentration reported by the different methods in the severe and mild to moderate oligozoospermic groups, it is known that this parameter depends on the characteristics of the counting chamber used for manual assessment [25, 26] and the accuracy of the automated CASA settings [27]. The low number of spermatozoa in these groups may have also contributed to some level of error in the manual assessment as well, despite strict fulfillment of WHO 5th recommendations in terms of sperm dilution. Several papers reported lower accuracy at very low sperm concentration levels using CASA systems [27–29]. In the severe oligozoospermic group, the CASA showed a much higher median for sperm concentration: Greater than 4 times the manual results (CASA: 7.6 vs. Manual: 1.75×10^6 /ml) whereas the SQA-V GOLD was in line with the manual results. These differences of results with the two analyses induce a difference of conclusions and some semen samples can be classified as moderate oligozoospermic group instead of severe (5×10^6 /ml cutoff). This is important from a clinical perspective since severe oligozoospermic men are typically referred for IVF/ICSI treatment directly while mild to moderate oligozoospermic cases could initially be candidates for intra-uterine insemination.

The manual motility results were significantly higher than the automated system's results primarily in the severe oligozoospermic group, whereas in the other groups they were in line, except for a slight difference in CASA motility in the mild to moderate oligozoospermic group. It is known that manual assessment of sperm motility is frequently overestimated [7]. It is possible that the limited number of spermatozoa manually counted in the severe oligozoospermic group (< 200 sperm cells), has an impact on the accuracy of the motility count.

Mountain plots showing differences between the automated and manual results demonstrated that the SQA-V GOLD and CASA plots were comparable for sperm concentration. The CASA mountain plots for MSC and PMSC tended to be

slightly wider, indicating an overestimation of these parameters. In our opinion, this tendency towards CASA overestimating MSC and PMSC is low and does not question the diagnostic accuracy of the system.

The SQA-V GOLD showed the best precision (lowest 95 % confidence interval) between the three methods. The high precision of the SQA-V GOLD may be due to the use of a larger, more representative sample volume. This larger volume is also a disadvantage because of the high minimum volume required. In contrast to the manual and CASA sample volumes of 10–50 μ l, the SQA-V GOLD technology analyses a more representative raw semen aliquot of 0.5 ml, which probably improves precision.

Concerning the pros and cons of the three semen assessment methods compared in this study, any technology that analyses live, motile cells, has inherent limitations. In both automated systems, the presence of debris in the raw semen may impact test results unless it is taken into consideration. This limitation can be overcome by visualizing the sample on-screen or in the microscope. Test strips (QwikCheck™ Test Strips, Medical Electronic Systems) can also be used to detect high concentrations of leucocytes in the sperm prior to testing. If detected previously by the operator, the automated analyzers can be adjusted to compensate for this. However, leucocytes are only part of the debris which may interfere with accurate analyses of the sperm parameters when using automated systems.

It is important to note that the term “automated” used in this study includes both the SQA-V GOLD which is a fully automated sperm analyzer and the HTM-Ceros CASA which relies on operator interface for determining the number of fields analyzed and for a variety of settings [18, 30–32]. CASA systems can be pre-set to the type of counting chamber and time interval between sperm collection and assay. The ability to adjust the CASA system can impact standardization between laboratories. Holt et al. compared 5 different CASA systems, including Hamilton Thorne HTM-2000, and concluded that emphasis should be placed on operator training and standardized specimen handling rather than on technical improvement and that variability within the CASA system was considerably greater than between CASA systems [33].

Positive reasons for using automated semen analysers are: Standardization, speed, precision, automated data recording, less human error and less need for highly skilled professionals to perform semen analysis. The negative aspect of using automated semen analysers, notably the problem with testing atypical samples and the inability to perform an automated morphology differential, can be addressed by integrating some manual assessment in these situations.

In conclusion, this double blind prospective study, based on WHO 5th Edition guidelines comparing three clinically distinguishable groups of semen quality analyzed using three separate and distinct methods, may be the first of its kind.

The two automated analysers, the SQA-V GOLD and CASA demonstrated acceptable agreement with the manual method. Between the two, the SQA-V GOLD demonstrated better agreement with the manual method, showing the best precision and clinically acceptable results. Our conclusion is that the automated semen analyzers can be used for routine semen analysis providing rapid clinically acceptable results with higher precision, and positively impacting laboratory standardization.

Disclosure All the authors disclose financial interests.

Conflict of interest All the authors confirm that they have nothing to disclose.

Grant, support None.

References

- De Jonge C. Semen analysis: looking for an upgrade in class. *Fertil Steril.* 2012;97:260–6.
- Guzick DS et al. Sperm morphology, motility, and concentration in fertile and infertile men. *N Engl J Med.* 2001;345:1388–93.
- Auger J et al. Intra- and inter-individual variability in human sperm concentration, motility and vitality assessment during a workshop involving ten laboratories. *Hum Reprod.* 2000;15:2360–8.
- Keel BA et al. Results of the American Association of Bioanalysts national proficiency testing programme in andrology. *Hum Reprod.* 2000;15:680–6.
- Filimberti E et al. High variability in results of semen analysis in Andrology Laboratories in Tuscany (Italy): the experience of an external quality control (EQC) programme. *Andrology.* 2013;1:401–7.
- Mallidis C et al. Ten years' experience with an external quality control program for semen analysis. *Fertil Steril.* 2012;98:611–e4.
- WHO. Laboratory manual for the examination and processing of human semen. 5th ed. Geneva: World Health Organization; 2010.
- Cooper TG et al. World Health Organization reference values for human semen characteristics. *Hum Reprod Update.* 2010;16:231–45.
- Agarwal A, Sharma RK. Automation is the key to standardized semen analysis using the automated SQA-V sperm quality analyzer. *Fertil Steril.* 2007;87:156–62.
- Hirano Y et al. Relationships between sperm motility characteristics assessed by the computer-aided sperm analysis (CASA) and fertilization rates in vitro. *J Assist Reprod Genet.* 2001;18:213–8.
- Fuse H et al. Assessment of sperm quality analyzer II B: comparison with manual semen analysis and CASA. *Arch Androl.* 2005;51:65–7.
- Akashi T et al. Usefulness of sperm quality analyzer-V (SQA-V) for the assessment of sperm quality in infertile men. *Arch Androl.* 2005;51:437–42.
- Tomlinson MJ et al. Validation of a novel computer-assisted sperm analysis (CASA) system using multitarget-tracking algorithms. *Fertil Steril.* 2010;93:1911–20.
- Iguer-Ouada M, Versteegen JP. Validation of the sperm quality analyzer (SQA) for dog sperm analysis. *Theriogenology.* 2001;55:1143–58.
- Schieferstein G, Hook-Vervier B, Schwarz M. Sperm motility index. *Arch Androl.* 1998;40:43–8.
- Martinez C et al. Sperm motility index: a quick screening parameter from sperm quality analyzer-IIB to rule out oligo- and asthenozoospermia in male fertility study. *Hum Reprod.* 2000;15:1727–33.
- Yeung CH, Cooper TG, Nieschlag E. A technique for standardization and quality control of subjective sperm motility assessments in semen analysis. *Fertil Steril.* 1997;67:1156–8.
- Davis RO, Boyers SP. The role of digital image analysis in reproductive biology and medicine. *Arch Pathol Lab Med.* 1992;16:351–63.
- Jagoe JR et al. Sperm morphology by image analysis compared with subjective assessment. *Br J Urol.* 1987;60:457–62.
- Moruzzi JF et al. Quantification and classification of human sperm morphology by computer-assisted image analysis. *Fertil Steril.* 1988;50:142–52.
- Esbert M et al. Impact of sperm DNA fragmentation on the outcome of IVF with own or donated oocytes. *Reprod BioMed Online.* 2011;23:704–10.
- Lewis SE. Is sperm evaluation useful in predicting human fertility? *Reproduction.* 2007;134:31–40.
- van der Merwe FH et al. The use of semen parameters to identify the subfertile male in the general population. *Gynecol Obstet Investig.* 2005;59:86–91.
- Menkveld R. Clinical significance of the low normal sperm morphology value as proposed in the fifth edition of the WHO Laboratory Manual for the Examination and Processing of Human Semen. *Asian J Androl.* 2010;12:47–58.
- Brazil C et al. Standardized methods for semen evaluation in a multicenter research study. *J Androl.* 2004;25:635–44.
- Mahmoud AM et al. The performance of 10 different methods for the estimation of sperm concentration. *Fertil Steril.* 1997;68:340–5.
- Davis RO, Katz DF. Operational standards for CASA instruments. *J Androl.* 1993;14:385–94.
- Vantman D et al. Computer-assisted semen analysis: evaluation of method and assessment of the influence of sperm concentration on linear velocity determination. *Fertil Steril.* 1988;49:510–5.
- Togni G et al. Computer-aided semen analysis: sperm concentration assessment by the Stromberg-Mika system. *Andrologia.* 1995;27:55–65.
- Johnson JE, Boone WR, Blackhurst DW. Manual versus computer-automated semen analyses. Part I. Comparison of counting chambers. *Fertil Steril.* 1996;65:150–5.
- Johnson JE, Boone WR, Blackhurst DW. Manual versus computer-automated semen analyses. Part II. Determination of the working range of a computer-automated semen analyzer. *Fertil Steril.* 1996;65:156–9.
- Johnson JE, Boone WR, Blackhurst DW. Manual versus computer-automated semen analyses. Part III. Comparison of old versus new design MicroCell Chambers. *Fertil Steril.* 1996;65:446–7.
- Holt W et al. Reproducibility of computer-aided semen analysis: comparison of five different systems used in a practical workshop. *Fertil Steril.* 1994;62:1277–82.