

ORIGINAL ARTICLE

Computer-aided sperm analysis, the new key player in routine sperm assessment

Benoit Schubert  | Mélanie Badiou | André Force

Eurofins Biomnis – Institut Rhonalpin
IVF Center, Clinique du Val d'Ouest –
Medicentre, Ecully, France

Correspondence

Benoit Schubert, Eurofins Biomnis, Clinique
du Val d'Ouest – Médicentre 39, chemin de
la Vernique, 69 130 Ecully, France.
Email: BenoitSchubert@eurofins.com

Abstract

For sperm analysis, important inter-laboratory variations have been observed in manual analyses. In this study, a computer-aided sperm analysis (CASA) system was assessed versus manual technique, and specific software modifications were operated to fit the David's classification already used in the laboratory. Four parameters were studied (concentration, motility, vitality and morphology), and at least 30 semen samples from 30 different patients have been tested. Manual and automated analyses were compared using a least-squares regression line analysis, Student's *t* test, Bland–Altman plots and Passing–Bablok regressions. Repeatability was also assessed, and coefficients of variation (CV) were calculated. Both manual and automated methods gave similar results for sperm concentration ($n = 150$), motility ($n = 30$), vitality ($n = 90$) and morphology ($n = 90$). Repeatability always showed a decrease in the CV with automated analysis; for example in normal range of sperm values, CV for manual and CASA analyses were, respectively, 9.0% versus 4.4% for sperm concentration, 5.2% versus 4.1% for motility, 7.3% versus 4.2% for vitality and 11.4% versus 4.1% for morphology. All parameters were comparable between automated and manual analysis, and repeatability measures confirm the more reliable values of the SCA compared to those of manual analysis.

KEYWORDS

computer-aided sperm analysis, manual sperm analysis, sperm analysis

1 | INTRODUCTION

When evaluating couple infertility, sperm analysis remains of paramount importance, despite criticisms regarding measurement uncertainties and their relevance when male infertility diagnoses are based on these examinations alone (De Jonge, 2012). Even if some may consider that these analyses have no more clinical value today than they did 25 or 30 years ago (De Jonge, 2012; Tomlinson, 2016), these analyses are nonetheless carried out daily in all medical laboratories, even if prescribers often prefer experienced laboratories (Freour, Delvigne, & Barrière, 2010). Actually, important inter-laboratory variations have been observed and multi-centric studies on male subjects, have demonstrated coefficient of variation (CV)

ranging from 21% to 80% (reviewed in Brito et al., 2016) for sperm concentration, one of the least subjective parameters. External Quality Assessment (EQA) programmes are a means to better evaluate these variations. For instance, it has been very recently reported among 121 participating laboratories in Belgium, over a 15-year period, a median CV of 19.2% for sperm concentration determination, and in the same period, a median CV of 79.4% for sperm morphology characterisation (Punjabi et al., 2016). For the French EQA programme, when the CVs of the various semen parameters are not reported, limits of acceptability for normal sperm shape (NS) percentage are set between 13.2 and 23.9 (average: 18.5% of NS; 221 laboratories), or between 8.7 and 18.2 (average: 13.5% of NS; 201

laboratories) for, respectively, the two samples sent at the end of 2015 (Biologie Prospective: 2015-2).

The use of automation is widely spread in medical laboratories, but it is rarely found when it comes to spermatology. Since 1985, computer-aided sperm analysis (CASA) systems have been released on the market but despite their ongoing technical evolutions (Amann & Waberserski, 2014), results have proved rather discordant (Hu, Lu, Shao, Huang, & Lü, 2013), even incomparable to those obtained via the manual technique (Vested et al., 2011); also, regarding morphology testing, their software programs were often calibrated according to scales different from what was standardly used in France. Thus, very different new interpretation standards were to be learned by laboratory professionals and prescribers (Mortimer, Horst, & Mortimer, 2015). Over the years, these automated systems were usually geared towards animal health and research, more precisely, with a view to obtaining numerous mobility parameters and assess various physiological states of sperm. Up until recently, such systems were rarely used in medical laboratories due to their prohibitive costs as compared to the services rendered when routinely used. New machines are now available, which are better suited to the new certification standards (ISO 15189) with their numerous quality control parameters, and to reduce inter-laboratory and intralaboratory variations. In this article, we shall review and assess the Sperm Class Analyser (SCA, Microptic, Spain) versus the manual technique of reference, and more specifically the automated morphology testing, technically adapted to David's classification.

2 | MATERIAL AND METHODS

2.1 | Material

The automated analyses were carried out on a Sperm Class Analyser (SCA) version 5.4.0.0. The system was made of a NIKON® Eclipse Ci microscope with phase contrast, equipped with a BASLER® 1024 digital camera, a PRIOR® OptiScan II motorised stage system, a heating platform and a SCA image acquisition software (Microptic SL). A software filter was also used in order to take into account only spermatozoa with heads of a certain size and associated mid-piece. Thus, isolated heads and cell debris were disregarded. In addition, for each parameter, a quick visual scanning of the fields taken into account was also carried out by the operator in order to set aside inconsistent elements.

2.2 | Patients

For each parameter, at least 30 semen samples from 30 different patients have been tested. Informed consent was obtained for the patients included in the study. All procedures performed in this study were in accordance with the ethical standards of Eurofins Biomnis and with the 1964 Helsinki declaration and its later amendments. Normality thresholds for concentration, motility and vitality described here-below, are those of the WHO (5th edition, 2010). For morphology analyses, interpretations were made

according to David's modified classification (Auger, Jouannet, & Eustache, 2016), as it has been used effectively for many years in our laboratory and by most of the French andrology laboratories, although WHO 5 classification is most often used in other countries. It was decided to use David's criteria as it was well understood by the clinicians and moreover gave good indications for Assisted Reproductive Technologies in our facility as the whole sperm morphology (head, mid-piece and flagellum) is taken into account. After liquefaction (30–45 min) at 35°C, analyses were performed by sequences for motility, concentration, vitality and morphology; hence each parameter was analysed one by one and separated from the others to take great care to have the minimum time elapsed between manual and automated analyses.

2.3 | Concentration

A total of 150 sperm samples were tested covering five levels of sperm concentrations: <1 Million/ml (M/ml); 1–5; 5–15; 15–80 and >80 M/ml (normality threshold: >15 M/ml).

For manual analyses, a dilution was always used. Sperm was homogenised and diluted in a 2% formol-saline. Dilution factors then depended on eye estimations of concentrations under microscopes (mainly dilutions of 1/20 estimated from 15 to 40 M/ml). Two dilutions were performed and read using Thoma cell counting chamber (reading of the entire square at 400× magnification). When a 10% deviation was observed between the two readings, a third one was performed. Results would then correspond to the average of the two readings which were closest.

For automated analyses, whole nondiluted homogenised semen samples (2.5 µl) were read on counting chambers (LEJA 10) and analysed by SCA with phase contrast, at 100× magnification. For concentrations >80 M/ml, a dilution was required. Whole sperm was then diluted at ½ or maximum ¼ of seminal fluids.

Concentrations were calculated on 200 spermatozoa. Repeatability was assessed on 1 normal sample and 1 abnormal sample read 10 times each by the same person for both manual and automated analyses.

2.4 | Motility

Motility was assessed on 30 sperm samples of both normal and pathological motility (motility threshold ≥ 32% of sperm with progressive motility). Manual analysis was performed first (5 min), then the automated analysis (<5 min); during the manual analysis, the semen was replaced at 35°C. This short sequence of analysis is unable to bias the motility as this parameter is supposed to be assessed within 1 hr (WHO, 2010).

For manual testing, after homogenisation, 2 × 10 µl were placed between a slide and a coverslip. Motility was assessed on both these samples at a 400× magnification, and on 2 × 100 spermatozoa. When a deviation >10% was observed between the two readings, a third semen slide was performed. Results would then correspond to the average of the two readings which were closest.

TABLE 1 Measurement intervals for normal sperm assessment in SCA

	Minimum	Maximum
Head		
Length (µm)	4.65	6.50
Width (µm)	2.58	4.05
Area (µm ²)	8.79	19.09
Ellipticity	1.10	1.79
Elongation	0.03	0.29
Roughness	0.33	0.98
Regularity	0.69	1.51
Acrosome (%)	28.00	70.00
Mid-piece		
Distance (µm)	0	4.7
Angle (degree)	0	60
Flagellum		
Length (µm)	35	70

For automated analyses, motility was read simultaneously with concentration on the same sample.

For repeatability assessment, three levels of motility were read 10 times each by the same person and by the automation, and progressive motility was selected for this parameter.

2.5 | Vitality

Given that sperm concentration may affect vitality assessment and in particular, on the automated analyses, we took 90 semen samples divided into three categories according to their concentration: <15; 15–100 and >100 M/ml were tested covering normal and pathological vitality (normality threshold > 58% of live spermatozoa). Manual analysis was performed first with a dry smear (which stopped the reaction within less than 1 min) to be read later; then, the SCA staining was realised and read immediately.

For manual analyses, eosin-nigrosin staining kit (VitalScreen, Fertipro) was used. One eosin drop was added to the 50 µl of whole spermatozoa. After 30 s, 2 nigrosin drops were added, smear tests were then performed and left to dry in open air. Two readings at a ×400 magnification were carried out on 100 spermatozoa: pink stained spermatozoa were dead cells, whereas those left unstained (white) were living cells. When a 10% deviation was observed between the two readings, then a third one was performed. Results would then correspond to the average of the two readings which were closest.

For automated analyses, kits of sperm vitality under fluorescence (Fluovit[®], Microptic) were used to obtain high contrast images. About 10 µl of sperm was mixed with 1 µl of trihydrochloride trihydrate. The mixture was first incubated for 5 min at 35°C. Then, 1 µl of propidium iodide was added. A second incubation was carried out away from the light and at room temperature. Once incubation was over, 10 µl of the mixture was placed between slide and coverslip, and reading was performed on 200 spermatozoa under fluorescence (Filters for

DAPI: excitation filters at 330–380 nm, dichroic mirrors at 400 nm and barrier filters at 420 nm) at 200× magnification. Red fluorescent spermatozoa were dead cells and those in blue were live cells.

Repeatability was assessed on 1 normal sample and 1 abnormal sample each read 10 times by the same person for manual and automated analyses.

2.6 | Morphology

Thin smears were carried out on fresh spermatozoa, and air-dried at room temperature for at least 1 hr. Smears were then fixed and stained using Sperm Stain kit (Microptic), and then rinsed under running water. Tests were run on 90 sperm samples, divided into three levels of morphology (30 sperm samples per level): <10% of Normal Shape (NS) spermatozoa, 10%–20% of NS spermatozoa and >20% of NS spermatozoa (normality threshold >23% of NS according to David's modified classification (Auger et al., 2016)).

For manual analyses, sperm morphology analyses were performed in oil immersion at 1,000× magnification. Two readings of 100 spermatozoa were carried out to determine the percentage of NS and the distribution of anomalies according to David's modified classification (Auger et al., 2016). When a 10% deviation was observed on the most represented population (normal or abnormal shapes), a third reading was performed. The lowest percentage of NS was issued.

For automated analyses, standardisation was required. Such standardisation was based on 40 semen samples with at least a 30% concentration of NS spermatozoa (Normal > 23%) initially assessed via manual analysis; among these 40 semen samples, 1,000 normal spermatozoa were tested via SCA at 600× dry magnification and their morphological parameters were recorded (data not shown). Through the software filter, the SCA identified spermatozoa of sizes between 5 and 40 µm² (from microcephalic to macrocephalic spermatozoa) with attached flagella. Parameters identifying normal spermatozoa were thus established, and spermatozoa outside of these values were considered as abnormal by the SCA and ranked according to David's criteria to calculate the percentage of NS and Multiple Anomalies Index (MAI): ratio of total number of anomalies on the number of abnormal spermatozoa. Measuring intervals for normal shape spermatozoa were reported in Table 1. Head ellipticity is determined by the ratio of length on width; head elongation is determined by the following ratio (length – width)/(length + width); head roughness is determined using the following ratio: (4 × π × surface area)/(perimeter)²; and head regularity using the following ratio: (π × length × width)/(4 × surface area). For the mid-piece (MP), a maximum space between head and MP of 4.7 µm was considered acceptable.

After automated analyses, captured fields were checked by laboratory technicians. In some instances, they have had to delete 'wrong analyses' due to debris being picked up by mistake as spermatozoa. Two readings of 100 spermatozoa were performed. When a deviation >10% was observed on the most represented population (normal or abnormal shapes), a third reading was performed.

Repeatability was assessed on four semen samples of four different levels of NS, each read 10 times by the same person both for manual and for automated analyses.

2.7 | Statistics

Manual and automated methods were compared using a least-squares regression line analysis. Firstly, we verified that the studied population followed a normal distribution by calculating for both analyses, the asymmetry and the flattening coefficients. Secondly, the Student's *t* test was run to make sure that the slope of the straight line was close to 1 and that the X-axis origin was close to 0. These *t* tests results were compared to a theoretical *T* value with a 5% risk factor. All calculations to determine the straight-line equation and the Student's *t* test were carried out using the Excel add-in programme, Analysis ToolPak (Windows). Repeatability was assessed by calculating the coefficient of variation (CV) on 10 readings. Bland–Altman plots and Passing–Bablok regression analyses were performed using XLSTAT statistical and data analysis solution version 2019.1.1. (Addinsoft, 2019).

3 | RESULTS

3.1 | Concentration

Population distributions were normal for both the manual and the automated analyses (coefficient of asymmetry/standard error and flattening coefficient/standard error of 0.50 and 1.89, respectively, for population tested via manual analyses, and of 0.47 and 1.69, respectively, for those tested via automated analyses; that is, coefficients between tolerance intervals of -2 and $+2$). According to the calculation of the equation of the least-squares regression line (Figure 1), the slope and the X-axis test results showed calculated *t* values (respectively 4.9 and 2.17 in absolute value) superior to the theoretical *t* values (1.96) with a 5% risk factor for both. These two analyses thus cannot be considered as similar. The differences graph allowed to better grasp the values at stake after tracking limits were determined: 4 discordant values were observed among the 150 values, outside of these limits (shown as X on Figure 2): for 1 point, a high dilution (1/40) was carried out for manual analysis only and concentrations of 45.1 and 30.8 M/ml were observed for, respectively, manual and automated techniques. For all three other points, a 1/20 dilution was required for manual analyses only and concentrations were, respectively, for manual and automated analyses of 36.1 and 23.4; 34.1 and 23.6; 14.9 and 10.3 M/ml. Bland–Altman plots showed a random pattern and the bias coefficient did not differ with 0 [95% CI: -1.509 to 0.900] (Figure 3). However, there 8 was points beyond the bias limits (5% of the samples) which were all superior to 100 M/ml (means 101 to 230 M/ml). The Passing–Bablok regression results showed that SCA analysis reached an acceptable agreement with manual analysis (slope [95% CI]: 0.959 to 1.000; intercept [95% CI]: -0.016 to 0.025 ; Figure 4).

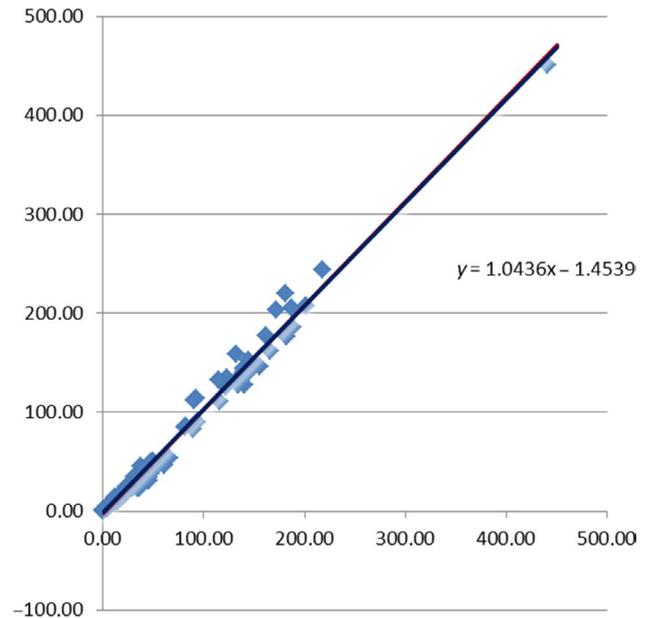


FIGURE 1 Sperm concentration. Least-squares regression line ($n = 150$)

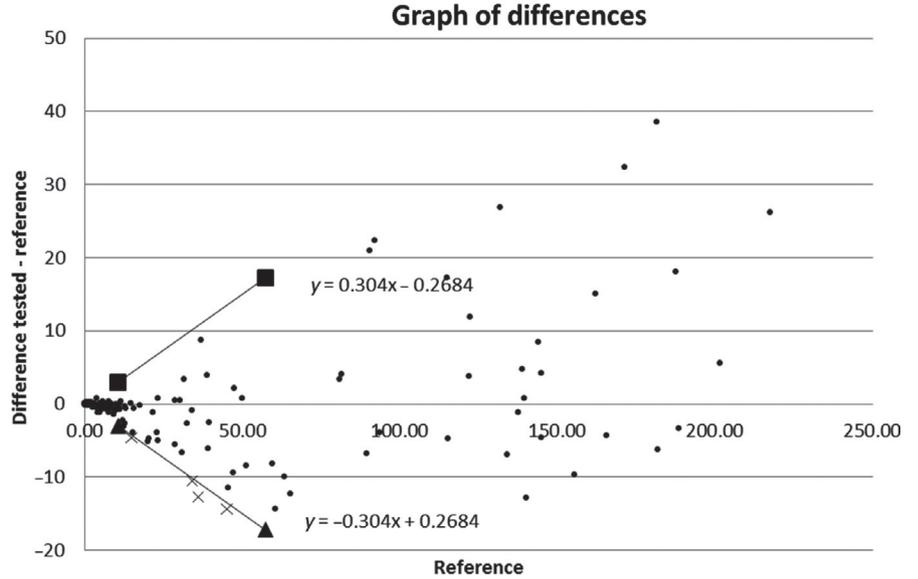
3.2 | Motility

Population distributions were normal for both manual and automated analyses (coefficient of asymmetry/standard error and flattening coefficient/standard error of -0.28 and 0.13 , respectively, for population tested via manual analyses, and of -0.24 and 0.25 , respectively, for those tested via automated analyses). For both manual and automated analyses, 30 mobility analyses were assessed covering a wide range of values for this parameter from 8% to 69% of spermatozoa with progressive motility (normal progressive motility threshold $>32\%$, WHO, 2010). According to the calculation of the least-squares regression line (Figure 5), the slope and the X-axis test results showed calculated *t* values (respectively 0.90 and 1.28) inferior to the theoretical *t* value (2.0) with a 5% risk factor for both. Both methods were indeed comparable. Although one observation (3% of the samples) was beyond the bias limit, Bland–Altman analysis showed a random pattern and the bias coefficient did not differ with 0 [95% CI: -2.500 to 0.182] (Figure 6). The Passing–Bablok regression results showed that SCA analysis reached an acceptable agreement with manual analysis (slope [95% CI]: 0.846 to 1.057; intercept [95% CI]: -0.929 to 6.071 ; Figure 7).

3.3 | Vitality

Population distributions were normal for both manual and automated analyses (coefficient of asymmetry/standard error and flattening coefficient/standard error of -1.06 and 1.90 , respectively, for population tested via manual analyses, and of -1.00 and 1.78 , respectively, for those tested via automated analyses). A total of 90 vitality analyses were assessed covering a wide range of values for this parameter from 23% to 93% of live spermatozoa (normal vitality threshold

FIGURE 2 Sperm concentration. Differences graph and tracking limits determination. For better representation, the point (441.0 and 450.7) included in the tracking limits and was removed in the differences graph. Solid bars with squares and triangles represent the tracking limits. ●, Values within the tracking limits; X, Values outside of the tracking limits



	Level 1	Level 2
Standard deviation for the method of reference	0.50	5.30
Standard deviation for the tested method	0.86	2.17
Tracking limits	2.98	17.18
-(negative) Tracking limits	-2.98	-17.18
Control target values	10.70	57.40

For better representation, the point (441.0 and 450.7) included in the tracking limits, was removed in the differences graph. Solid bars with squares and triangles represent the tracking limits.

- ▲ ——— ▲ ■ ——— ■ : Tracking limits
- Values within the tracking limits
- X Values outside of the tracking limits

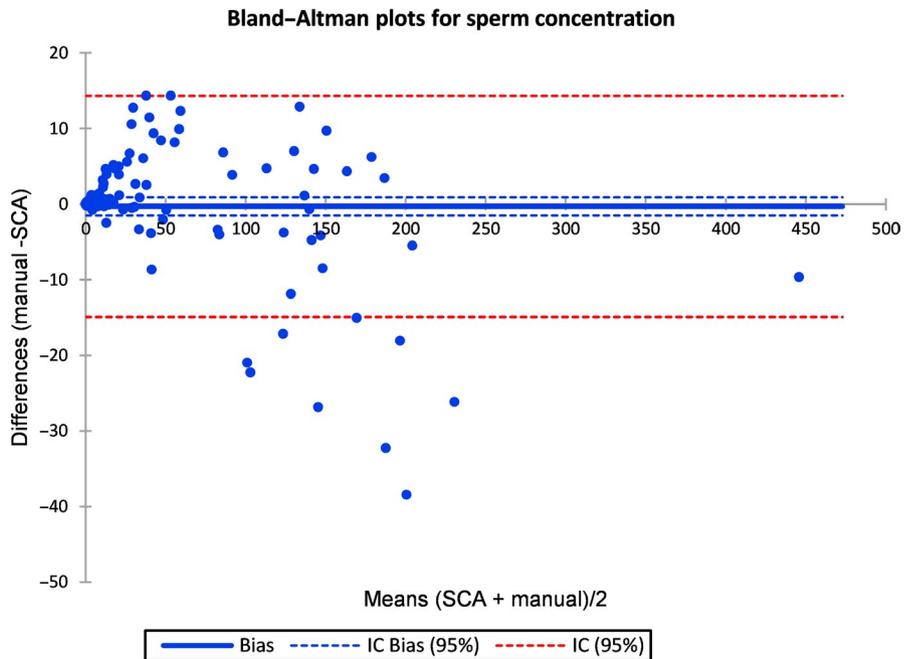


FIGURE 3 Sperm concentration. Bland-Altman plots

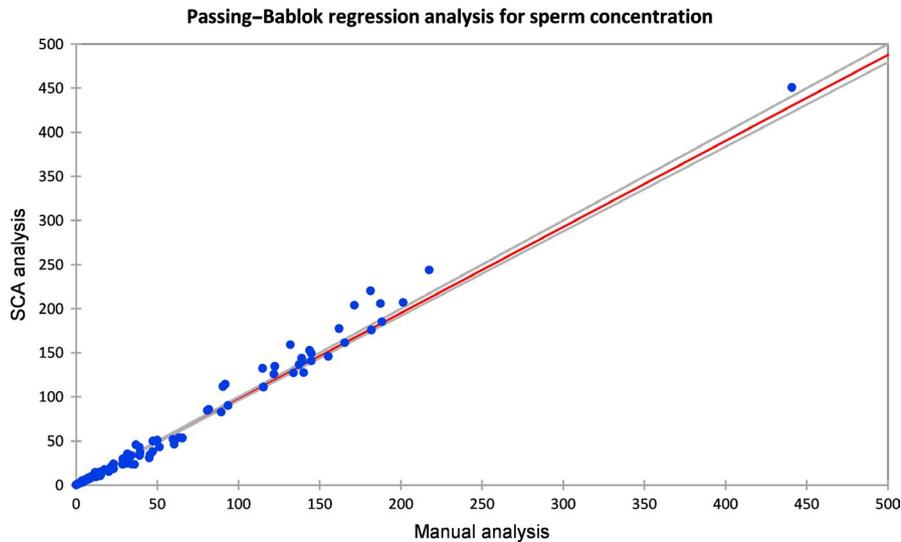


FIGURE 4 Sperm concentration. Passing–Bablok regression

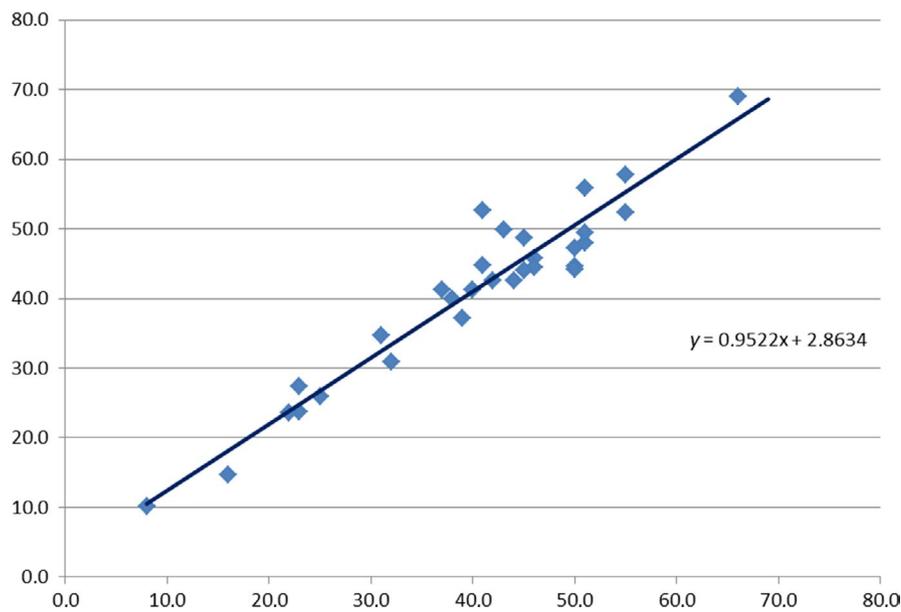


FIGURE 5 Sperm motility. Least-squares regression line ($n = 30$)

$\geq 58\%$, WHO, 2010). According to the calculation of the least-squares regression line (Figure 8), the slope and the X-axis test results showed calculated t values (respectively 0.86 and 1.12) inferior to the theoretical t value (≥ 1.96). Both methods were indeed comparable. Although two observations (2% of the samples) were beyond the bias limit, Bland–Altman analysis showed a random pattern and the bias coefficient did not differ with 0 [95% CI: -1.137 to 0.115] (Figure 9). The Passing–Bablok regression results showed that SCA analysis reached an acceptable agreement with manual analysis (slope [95% CI]: 0.968 to 1.083 ; intercept [95% CI]: -5.292 to 3.166 ; Figure 10).

3.4 | Morphology

Population distributions were normal for both manual and automated analyses (coefficient of asymmetry/standard error and flattening coefficient/standard error of 0.57 and 0.25, respectively, for population tested via manual analyses, and of 0.35 and -0.65 ,

respectively, for those tested via automated analyses). Ninety morphology analyses were assessed covering a wide range of values for this parameter from 0% to 42% of NS spermatozoa (normal morphology threshold $\geq 23\%$, Auger et al., 2016). According to the calculation of the least-squares regression line (Figure 11), the slope and the X-axis test results showed calculated t values (respectively 0.57 and 1.11) both inferior to the theoretical t value (≥ 1.96). Both methods were indeed comparable. Bland–Altman analysis showed a random pattern and the bias coefficient did not differ with 0 [95% CI: -1.550 to 0.127] (Figure 12). However, seven observations (8% of the samples) were outside the bias limits: three had both values under, or both above the normal threshold and four had values surrounding and near this threshold (means = 26.0; 27.5; 17.5; 26). The Passing–Bablok regression results showed that SCA analysis reached an acceptable agreement with manual analysis (slope [95% CI]: 1.000 to 1.172 ; intercept [95% CI]: -1.559 to 0.000 ; Figure 13).

FIGURE 6 Sperm motility. Bland-Altman plots

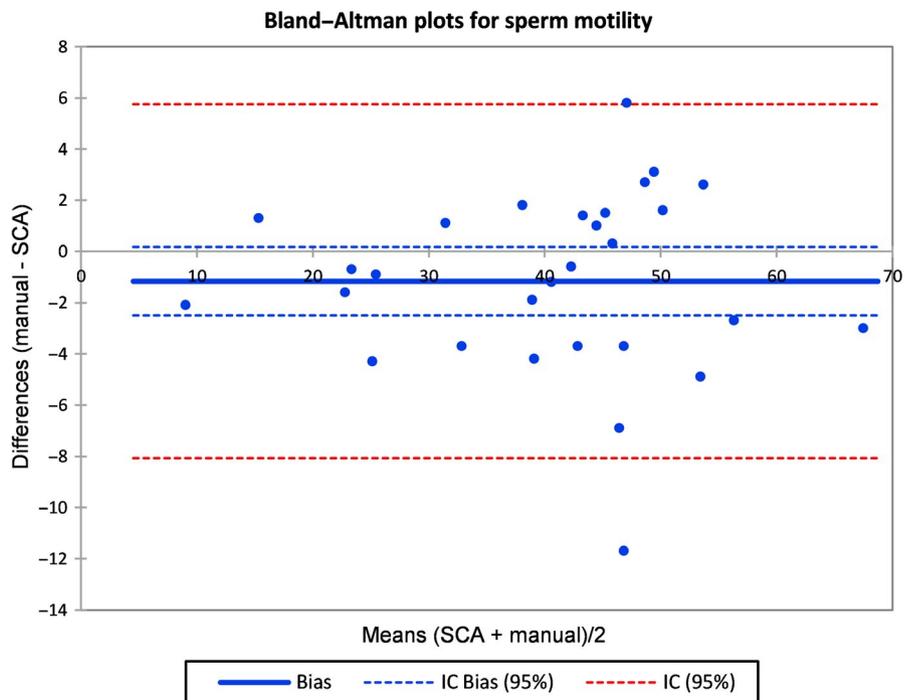
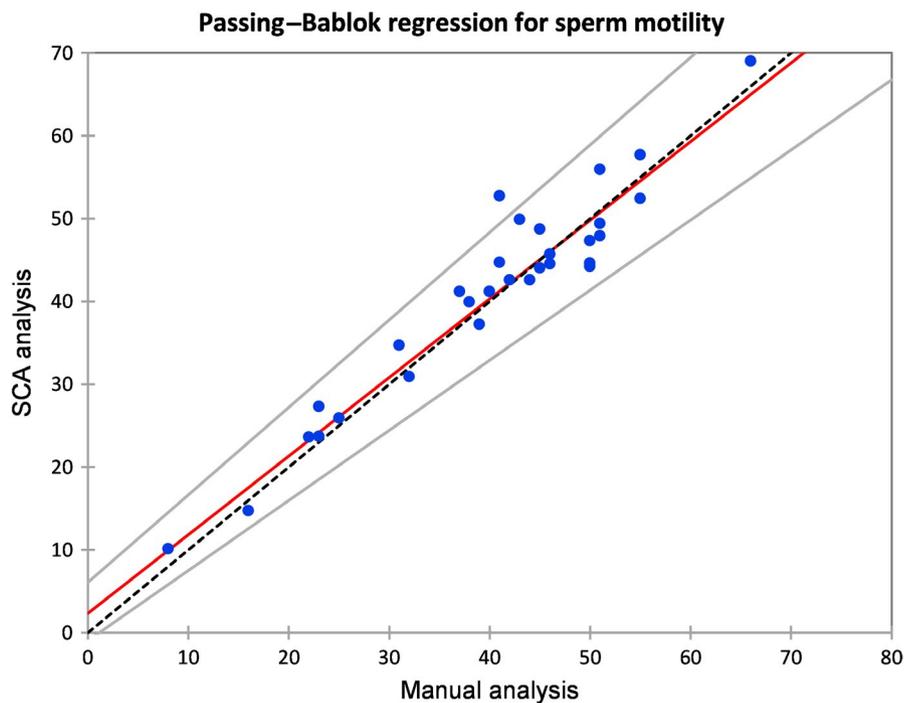


FIGURE 7 Sperm motility. Passing-Bablok regression



Repeatability results hereafter are shown, respectively, for manual and automated analyses. For concentration, CV of, respectively, 8.5% and 5.9% were observed for abnormal concentrations (oligozoospermia: 7.3 M/ml), and of, respectively, 9.0% and 4.4% and for normal sperm concentrations (36.5 M/ml). Motility was assessed on 3 populations with CV of, respectively, 14.4% and 7.0% for asthenozoospermic spermatozoa (progressive motility: 15%); and of, respectively, 7.8% and 7.3% for sperm motility close to normal threshold (progressive motility 30%); and of, respectively, 5.2% and 4.1% for normal sperm motility

(progressive motility 50%). Vitality was assessed on two populations with CV of, respectively, 7.4% and 3.4% for abnormal sperm vitality (50%) and of, respectively, 7.3% and 2.4% for normal vitality (65%). Lastly, sperm morphology was assessed on four populations with CV of, respectively, 31.0% and 14.5% for severe teratozoospermia (5% of NS); of, respectively, 16.5% and 6.9% for moderate teratozoospermia (15% of NS); of, respectively, 12.1% and 5.0% for sperm morphology close to normal threshold (20% of NS) and of, respectively, 11.4% and 4.1% for normal sperm morphology (30% of NS).

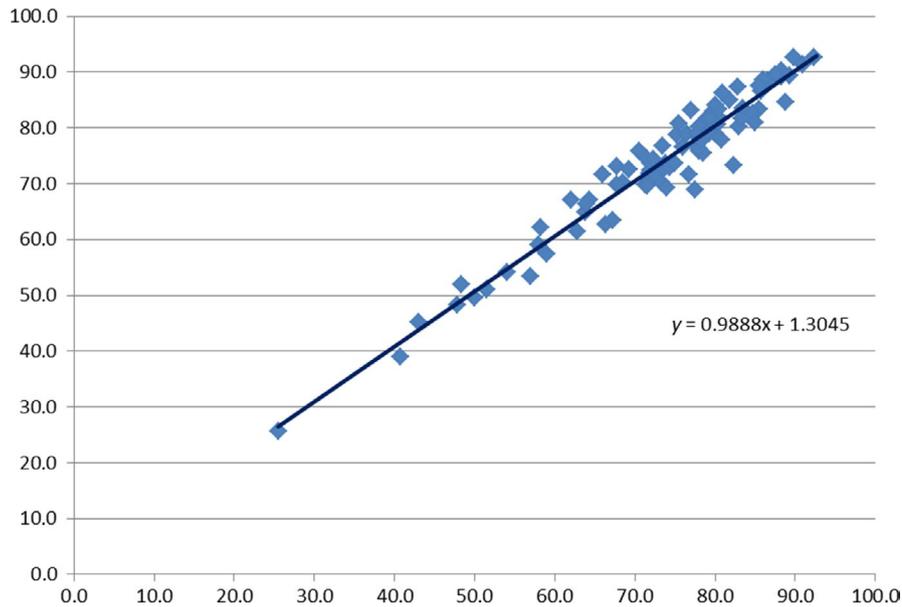


FIGURE 8 Sperm vitality. Least-squares regression line ($n = 90$)

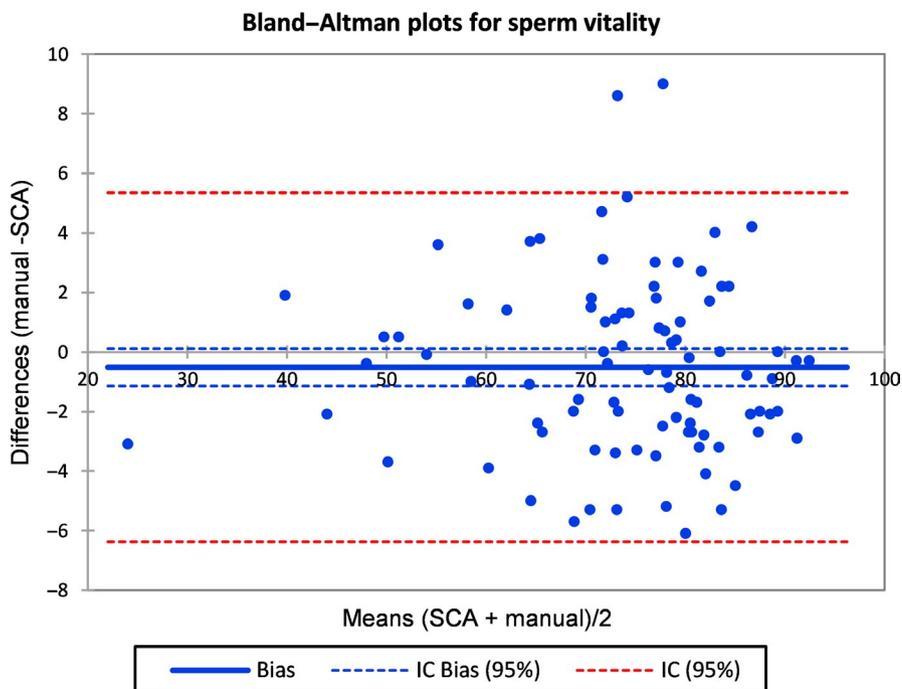


FIGURE 9 Sperm vitality. Bland-Altman plots

Finally, as an example of the variations obtained among the samples with both techniques studied, the data from 12 patients are provided in Table 2. Small changes were observed between the analyses and their conclusions were the same whatever the technique used.

4 | DISCUSSION

In this comparative study between manual (gold standard) and automated sperm analyses on the main parameters, results were very similar between the two techniques. This comparative study was carried out on a large number of patients to cover all possible

normal or pathological status for each parameter, and also to tally the great heterogeneity between semen samples from a single patient, a difficulty inherent to sperm analysis. For sperm concentration, the Student's *t* test used did not allow for comparison of the manual and automated techniques, as opposed to motility, vitality and morphology parameters. This statistical test is very sensitive and the four discordant values among the 150 samples which were tested, were all situated in the lower part of the differences graph. The additional systematic dilution required when using the manual technique necessarily induced a technical bias, whereas on the SCA system it was nondiluted whole spermatozoa which was tested. The volume of diluent being relatively greater than that of spermatozoa, even a slight pipetting error would modify results significantly, as

FIGURE 10 Sperm vitality. Passing–Bablok regression

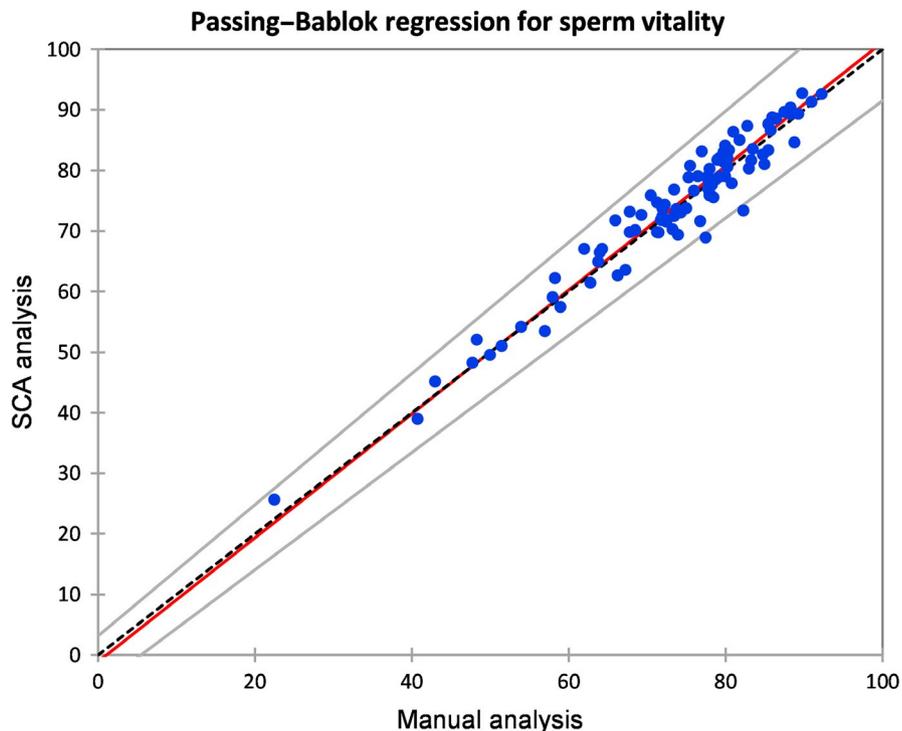
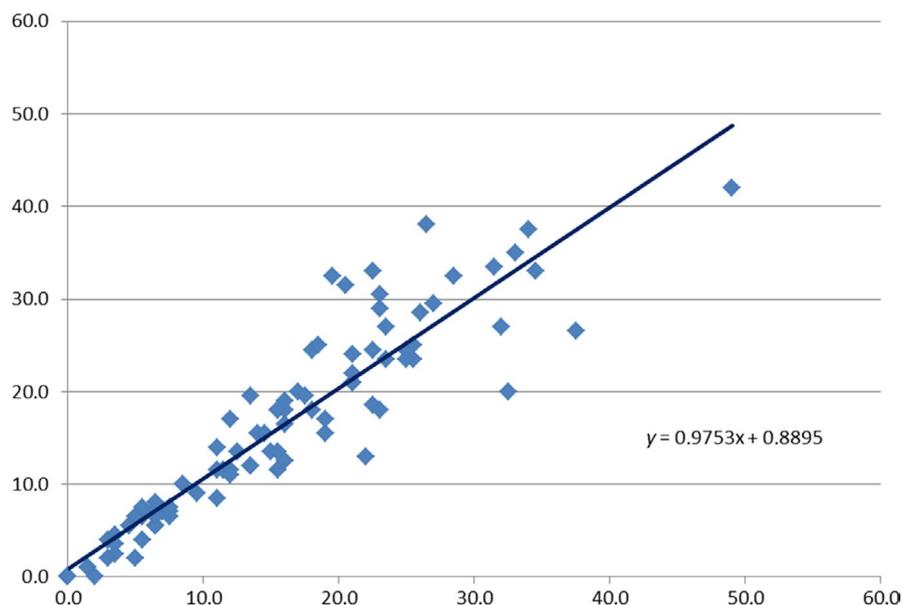


FIGURE 11 Sperm morphology (David's classification modified by Auger et al., 2016). Least-squares regression line ($n = 90$)



per recently reminded by Brito et al. (2016). For all 4 discordant points, dilutions of 1/40 or 1/20 were used for manual analyses. Concentrations were all found to be superior to those of the automated technique, and the greater the dilution, the greater the concentration. All these elements confirmed the bias effect of dilution. However, for each case, analysis conclusions were the same regardless of the technique used: normal sperm concentration for the first 3 points, and oligozoospermia for the last one. There was neither interpretation error, nor any clinical consequences for each of these 4 points on either analysis. Bland–Altman plots and Passing–Bablok regression confirmed that the comparison of these

two methods appeared acceptable, although eight values were outside the bias limits for the Bland–Altman plots. These values were not a matter of concern as they were all clearly considered as normal with both techniques (101 to 230 M/ml; normal threshold >15 M/ml). Overall, except for these values, we also came to the conclusion that for sperm concentration determination, manual and automated analyses were comparable. Moreover, this was true for a wide range of concentrations, here estimated between 0.1 and 440 M/ml. This span was considerably widened as compared to the works of Johnson (Johnson et al., Part II, 1996) in which it had been restricted to concentrations between 20 and 149 M/ml

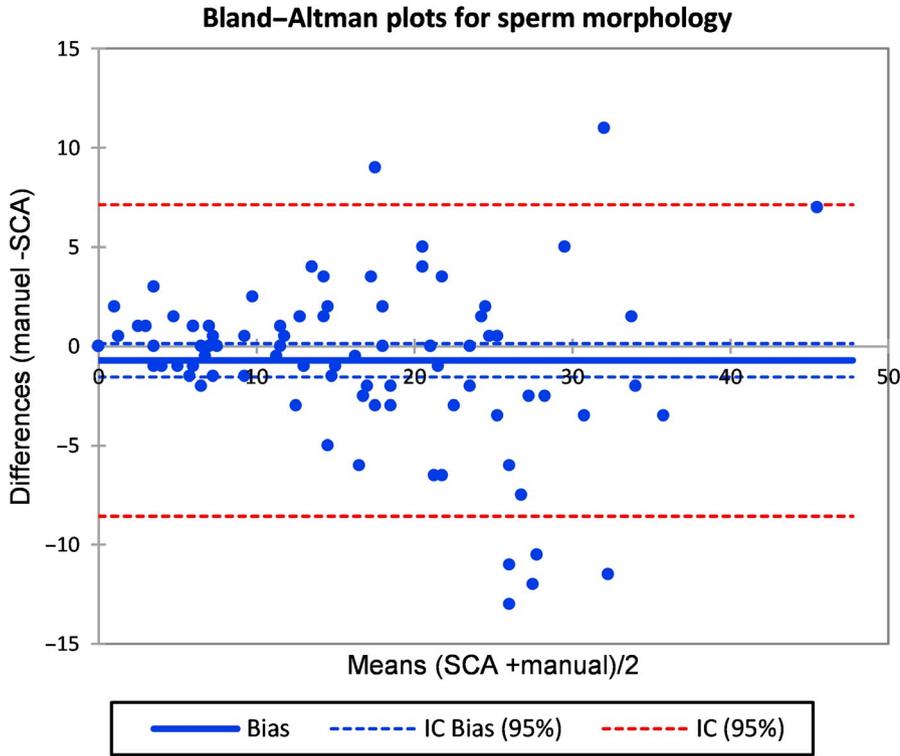


FIGURE 12 Sperm morphology (David's classification modified by Auger et al., 2016). Bland-Altman plots

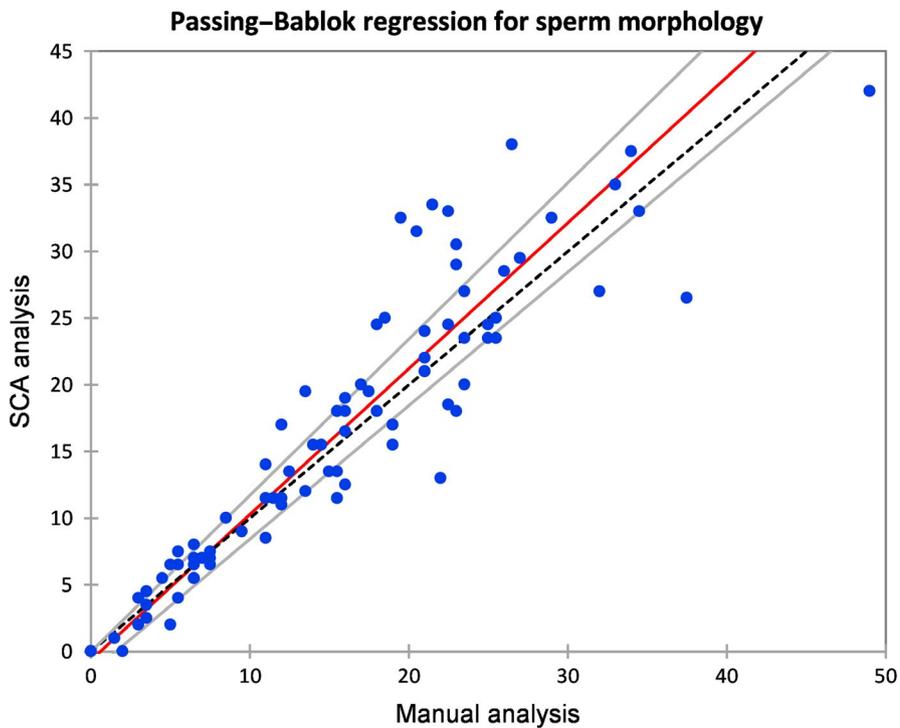


FIGURE 13 Sperm morphology (David's classification modified by Auger et al., 2016). Passing-Bablok regression

on another automated system. Additional recent developments were taken into account in our comparative study; for instance, a software filter was used allowing identification of cellular debris from spermatozoa of equal sizes. Thus, as opposed to what was found in Dearing, Kilburn, and Lindsay (2014) in which a SCA was also used, isolated heads and cellular debris were discarded thanks to the combined actions of the software filter and a brief visual inspection of the fields involved.

Regarding concentration determination, the comparison between manual and automated analyses via other CASA systems have sometimes shown significant variations in the past. Each time, these variations proved to be due to technical errors such as inadequate inclusion of cellular debris, unsuitable use of dilutions, nonapproved analytical chambers (reviewed by Brito et al., 2016). These error factors were under control in our study, as we took single-use calibrated chambers and doubled checked the tested population thanks

TABLE 2 Examples of manual and automated (SCA) sperm analyses from 12 patients

Analysis		Concentration (M/ml)	Concl.	Progressive motility (%)	Concl.	Vitality (%)	Concl.	Morphol. (% NS)	Concl.
1	Manual	8.4	Ab	8.0	Ab	22.5	Ab	6	Ab
	SCA	8.0	Ab	10.1	Ab	25.6	Ab	7	Ab
2	Manual	5.4	Ab	23.0	Ab	64.3	N	5	Ab
	SCA	4.6	Ab	23.7	Ab	67.0	N	6	Ab
3	Manual	11.6	Ab	25.0	Ab	54.0	Ab	16	Ab
	SCA	11.9	Ab	25.9	Ab	54.1	Ab	13	Ab
4	Manual	12.1	Ab	23.0	Ab	73.5	N	3	Ab
	SCA	9.9	Ab	27.3	Ab	72.4	N	4	Ab
5	Manual	21.8	N	51.0	N	47.8	N	32	N
	SCA	20.6	N	47.9	N	48.2	N	27	N
6	Manual	50.0	N	46.0	N	83.5	N	25	N
	SCA	50.8	N	44.5	N	83.5	N	24	N
7	Manual	47.5	N	45.0	N	85.8	N	12	Ab
	SCA	49.6	N	44.0	N	86.6	N	11	Ab
8	Manual	23.0	N	66.0	N	71.5	N	26	N
	SCA	19.1	N	69.0	N	69.7	N	25	N
9	Manual	139.0	N	41.0	N	86.0	N	7	Ab
	SCA	143.8	N	44.7	N	88.7	N	8	Ab
10	Manual	122.0	N	55.0	N	69.3	N	12	Ab
	SCA	125.8	N	52.4	N	72.6	N	12	Ab
11	Manual	89.5	N	43.0	N	78.3	N	11	Ab
	SCA	82.7	N	49.9	N	77.6	N	14	Ab
12	Manual	201.5	N	45.0	N	75.0	N	16	Ab
	SCA	207.0	N	48.7	N	73.7	N	18	Ab

to the combined actions of the software filter and visual inspections. Another study reported noncomparable concentration and motility results when using a CASA CRISMAS system (Vested et al., 2011). Again, the cellular debris issue was raised along with a software fault as comparable results had been found between manual and automated analyses on a previous version of the said software (Larsen et al., 2000). On the contrary, a study carried out on an earlier version of the SCA than the one we have used, showed a highly consistent coefficient of determination ($r^2 = .95$; 250 samples) between manual and automated analyses for concentration determination (Dearing et al., 2014). Another study using an undetermined type of CASA system also found a high correlation coefficient for concentration and motility (respectively $r = .84$ and $.62$; Kose, Karakoc Sokmensuer, Demir, Bozdog, & Gunalp, 2014). This correlation was even stronger with two other CASA systems (SQA-V Gold, Electronic Medical Systems and CASA CEROS, Hamilton) when tested versus manual analyses. All correlation coefficients were superior to 0.94 for both concentration and motility determination (Lammers, Splingart, Barrière, Jean, & Fréour, 2014). When using a CASA prototype with a customised software including a path tracking algorithm (Tomlinson et al., 2010), Tomlinson reported a high rate of convergence, strongly significant for concentration determination ($R^2 = .94$; $p < .001$ on 287 samples)

and for progressive motility assessment ($r = .85$; $p < .0001$). Authors nonetheless reported a disagreement regarding concentration interpretation (normal or oligozoospermic) for 5 out of the 287 tested semen samples (2%). In our study, we used an off the shelf CASA system, and no discordance was observed in result interpretation even if four discordant points were reported out of 150 semen samples (3%). In the same way, sperm motility and sperm viability results appeared very comparable with both techniques as measured by least square regressions, Bland-Altman plots and Passing-Bablok regressions.

As for morphology testing, very few comparative studies between manual and automated analyses are available. Significant statistical differences were reported only on normal values and according to the WHO classification (2010): 7% versus 5% and 7% versus 10.6%, respectively, for CASA CEROS and SQA-V Gold ($p < .05$). Given the very high levels of specificity ($>83\%$) and the negative predictive value ($>92\%$) reported for each automated system as compared to manual analyses, values obtained via automated systems were considered clinically satisfying in order to estimate normal morphology (Lammers et al., 2014). In our study, except for minor points near the normal threshold, both methods of analysis appeared comparable for sperm morphology as measured by least square regression, Bland-Altman plots and Passing-Bablok regression.

CASA systems are primarily meant to limit subjectivity (Amann & Waberserski, 2014; Yániz, Soler, & Santolaria, 2015) and for each of the parameters we tested, repeatability CV was indeed significantly reduced with automated analyses. Few elements of comparison are available in literature regarding repeatability assessment. However, when it comes to human sperm concentrations, repeatability CV was reported in manual analyses, ranging from 8.5% to 17.9% (10 samples read 5 times each; Bailey et al., 2007) or from 7.4% to 10.6% (each sample read 4 times; Tomlinson et al., 2010) depending on the cell types included for cell counting (Auger et al., 2000). The latter also reported a mean intra-individual CV of 15.8% for 12 participants (5 frozen samples read 3 times each). These values are really close to those we found. For automated analyses in Tomlinson et al. (2010), repeatability CV was reduced (1.0%–5.0% for 5 samples of immobilised sperm, tested 10 times). These last values are slightly inferior to those observed in our study (4.4%–5.9%), most certainly due to sperm immobilisation. Actually, with motile spermatozoa, a slight imprecision has been reported then (Brito et al., 2016). In our study, we chose to carry out experiments as close to real conditions of use of the SCA as possible, in a real andrology laboratory with fresh whole spermatozoa with motile spermatozoa.

For total motility assessment in manual analyses, mean intra-individual CV (on 12 participants read 3 times each sample) and mean inter-individual CV (17 samples tested among 12 participants) of, respectively, 26.2% and 21.8% were reported in Auger et al. (2000). In our study, intra-individual CV for progressive motility was lower (14.4% and 5.2% depending on sample characteristics) using manual analyses and even lower using automated analyses (between 7.0% and 4.1%).

For vitality assessment in manual analyses, mean intra- and inter-individual CV were, respectively, of 13.1% and 17.5% in the study of Auger et al. (2000). In our repeatability assessment, CV were lower, around 7% for manual analyses and around 3% for automated analyses.

Morphology is a parameter subjected to strong inter-laboratory variations, so much so that some consider it to be one of the most subjective parameters in sperm testing (Punjabi et al., 2016) and moreover some question its relevance (Gatimel, Moreau, Parinaud, & Léandri, 2017). Authors revealed that classification changes and standard modifications inside the successive editions of the WHO classification all played a part in keeping this parameter in blurred territory. In fact, the first edition of the WHO manual (1980) considered a normal percentage of normal shapes beyond 80%. For the second edition (1987) and the third one (1992), percentages were, respectively, of 50% and 30%. The fourth edition (1999) recommended the Tygerberg strict criteria which are restricted to sperm heads. This parameter was then considered normal above 14% of normal shapes. In 2010, the fifth edition, while keeping the same procedure and the same classification, once more changed the standard and considered as normal, sperm morphology above 4% of normal shapes (Menkveld, 2010)! Thus, standard values have changed considerably over 30 years, and trust in this parameter was necessarily undermined. Inter-laboratory variations are still

considerable today and French or foreign EQA programmes are the very expression of this trend: in 2008, the American Association of Bioanalysts (AAB) reported satisfying values around mean values of $CV \pm 3$, from 0% to 28% for one tested sample with the same coloration and interpretation methods according to strict criteria (threshold at 4%) and from 5% to 70% with the WHO criteria from 1992 (threshold at 30%; reported by Brazil, 2010). The College of American Pathology (CAP), another EQA supplier, noted 'high CV' in sperm morphology results without going as far as to giving figures (www.cap.org, 2016 report); it must be pointed out that the use of ocular micrometers was required when referring to strict criteria, when an internal enquiry revealed that less than 25% of the participating laboratories did use them (535 participants in 2016 with 54% of them referring to strict criteria). Therefore, others requested that the use of micrometers shall be documented when this classification was used (Carrell & De Jonge, 2016). Such deviations were similarly reported in the French EQA programme (Biologie Prospective) with a majority of laboratories using David's method of classification (77% of the 320 participating laboratories in 2015), and in keeping with the 3rd edition of the WHO manual (threshold 30%).

In our laboratory, we use David's method of classification (David, Bisson, Czyglick, Jouannet, & Gernigon, 1975) as per modified by Auger et al. (2016), the only method that does take into account spermatozoon in its entirety: head + mid-piece + flagellum. With this method, the threshold of normal shape spermatozoa is set at 23%. This new threshold was the one we used in our study for both manual and automated analyses. In our study, repeatability CV covered a wide range of values in manual analyses and was more limited in automated analyses (respectively from 11% to 31% and from 4% to 14%). The use of automated systems with precise and systematic assessment of spermatozoa is perfectly indicated especially to restrict such extreme variations (Auger et al., 2016; Bellastella et al., 2010; Cooper, 2016; Yániz et al., 2015).

5 | CONCLUSIONS

For andrology laboratories, CASA systems, despite their elevated costs, are a means to strengthen results especially when dealing with analyses with wide ranges of normal values. The least CV was always obtained with the SCA whatever the parameter analysed in normal and abnormal ranges. It is nevertheless paramount that laboratory technicians be trained in sperm analyses. The implementation of CASA systems must be supported by training which should be quite rapid for experimental sperm testing laboratory technicians. Human expertise is always crucial, more specifically for the parameters with very low values where automated systems are most likely pushed to their limits. Unlike commonly stated, CASA systems are no simple 'black boxes', issuing cold, bare results outside any human intervention (Dearing et al., 2014) but rather they are valuable assets providing faster and more precise results for the main parameters in routine sperm analyses.

ACKNOWLEDGEMENTS

This paper was made possible by all the lab staff of Microptic* who helped us in monitoring and adjusting the software.

CONFLICT OF INTEREST

All authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

AUTHORS' CONTRIBUTION

André Force and Benoît Schubert designed the research. Mélanie Badiou performed the analysis. Benoît Schubert wrote the paper. André Force and Benoît Schubert both critically reviewed the manuscript.

ORCID

Benoit Schubert  <https://orcid.org/0000-0002-9003-0937>

REFERENCES

- Amann, R. P., & Waberserski, D. (2014). Computer-assisted sperm analysis (CASA): Capabilities and potential developments. *Theriogenology*, 81, 5–17. <https://doi.org/10.1016/j.theriogenology.2013.09.004>
- Auger, J., Eustache, F., Ducot, B., Blandin, T., Daudin, M., Diaz, I., ... Jouannet, P. (2000). Intra and inter-variability in human sperm concentration, motility and viability assessment during a workshop involving ten laboratories. *Human Reproduction*, 15, 2360–2368.
- Auger, J., Jouannet, P., & Eustache, F. (2016). Another look at human sperm morphology. *Human Reproduction*, 31, 10–23. <https://doi.org/10.1093/humrep/dev251>
- Bailey, E., Fenning, N., Chamberlain, S., Devlin, L., Hopkisson, J., & Tomlinson, M. (2007). Validation of sperm counting methods using limits of agreement. *Journal of Andrology*, 28, 364–373. <https://doi.org/10.2164/jandrol.106.002188>
- Bellastella, G., Cooper, T. G., Battaglia, M., Ströse, A., Torres, I., Hellenkemper, B., ... Sinisi, A. A. (2010). Dimensions of human ejaculated spermatozoa in Papanicolaou-stained seminal and swim-up smears obtained from the Integrated Semen Analysis System (ISAS®). *Asian Journal of Andrology*, 12, 871–879. <https://doi.org/10.1038/aja.2010.90>
- Brazil, C. (2010). Practical semen analysis: From A to Z. *Asian Journal of Andrology*, 12, 14–20. <https://doi.org/10.1038/aja.2008.51>
- Brito, L. F. C., Althouse, G. C., Aurich, C., Chenoweth, P. J., Eilts, B. E., Love, C. C., ... Waberski, D. (2016). Andrology laboratory review: Evaluation of sperm concentration. *Theriogenology*, 85, 1507–1527. <https://doi.org/10.1016/j.theriogenology.2016.01.002>
- Carrell, D. T., & De Jonge, C. J. (2016). The troubling state of the semen analysis. *Andrology*, 4, 761–762. <https://doi.org/10.1111/andr.12257>
- Cooper, T. G. (2016). Looking forward to human sperm morphology. *Human Reproduction*, 31, 8–9. <https://doi.org/10.1093/humrep/dev253>
- David, G., Bisson, P., Czyglick, F., Jouannet, P., & Gernigon, C. (1975). Anomalies morphologiques du spermatozoïde humain. Proposition pour un système de classification. *Journal de Gynécologie Obstétrique et Biologie de la Reproduction*, 4(Suppl), 37–86.
- De Jonge, C. (2012). Semen analysis: Looking for an upgrade in class. *Fertility and Sterility*, 97, 260–266. <https://doi.org/10.1016/j.fertnstert.2011.12.045>
- Dearing, C. G., Kilburn, S., & Lindsay, K. S. (2014). Validation of the sperm class analyser CASA system for sperm counting in a busy diagnostic semen analysis laboratory. *Human Fertility*, 17, 37–44. <https://doi.org/10.3109/14647273.2013.865843>
- Freour, T., Delvigne, A., & Barrière, P. (2010). L'exploration de l'homme du couple infécond. *Journal de Gynécologie, Obstétrique et Biologie de la Reproduction*, 39, S42–S45. [https://doi.org/10.1016/S0368-2315\(10\)70030-2](https://doi.org/10.1016/S0368-2315(10)70030-2)
- Gatimel, N., Moreau, J., Parinaud, J., & Léandri, R. D. (2017). Sperm morphology: Assessment, pathophysiology, clinical relevance, and state of the art in 2017. *Andrology*, 5, 845–862. <https://doi.org/10.1111/andr.12389>
- Hu, Y. A., Lu, J. C., Shao, Y., Huang, Y. F., & Lü, N. Q. (2013). Comparison of the semen analysis results obtained from two branded computer-aided sperm analysis systems. *Andrology*, 45, 315–318. <https://doi.org/10.1111/and.12010>
- Johnson, J. E., Boone, W. R., & Blackhurst, D. W. (1996). Manual versus computer-automated semen analyses. Part II. Determination of the working range of a computer-automated semen analyzer. *Fertility and Sterility*, 65, 156–159. [https://doi.org/10.1016/S0015-0282\(16\)58044-3](https://doi.org/10.1016/S0015-0282(16)58044-3)
- Kose, M., Karakoc Sokmensuer, L., Demir, A., Bozdog, G., & Gunalp, S. (2014). Manual versus computer-automated semen analysis. *Clinical and Experimental Obstetrics & Gynecology*, 6, 662–664.
- Lammers, J., Spingart, C., Barrière, P., Jean, M., & Fréour, T. (2014). Double-blind prospective study comparing two automated sperm analyzers versus manual semen assessment. *Journal of Assisted Reproduction and Genetics*, 31, 35–43. <https://doi.org/10.1007/s10815-013-0139-2>
- Larsen, L., Scheike, T., Jensen, T. K., Bonde, J. P., Ernst, E., Hjollund, N. H., ... Giwercman, A. (2000). Computer-assisted semen analysis parameters as predictors for fertility of men from the general population. The Danish first pregnancy planner study team. *Human Reproduction*, 15, 1562–1567. <https://doi.org/10.1093/humrep/15.7.1562>
- Menkveld, R. (2010). 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 Journal of Andrology*, 12(1), 47–58. <https://doi.org/10.1038/aja.2009.14>
- Mortimer, S. T., Van der Horst, G., & Mortimer, D. (2015). The future of computer-aided sperm analysis. *Asian Journal of Andrology*, 17, 545–553. <https://doi.org/10.4103/1008-682X.154312>
- Punjabi, U., Wyns, C., Mahmoud, A., Vernelen, K., China, B., & Verheyen, G. (2016). Fifteen years of Belgian experience with external quality assessment of semen analysis. *Andrology*, 4, 1084–1093. <https://doi.org/10.1111/andr.12230>
- Tomlinson, M. J. (2016). Uncertainty of measurement and clinical value of semen analysis: Has standardization through professional guidelines helped or hindered progress? *Andrology*, 7, 763–770.
- Tomlinson, M. J., Pooley, K., Simpson, T., Newton, T., Hopkisson, J., Jayaprakasan, K., ... Pridmore, T. (2010). Validation of a novel computer-assisted sperm analysis (CASA) system using multitarget-tracking algorithms. *Fertility and Sterility*, 93, 1911–1920. <https://doi.org/10.1016/j.fertnstert.2008.12.064>
- Vested, A., Ramlau-Hansen, C. H., Bonde, J. P., Thulstrup, A. M., Kristensen, S. L., & Toft, G. (2011). A comparison of conventional and computer-assisted semen analysis (CRISMAS software) using samples from 166 young Danish men. *Asian Journal of Andrology*, 13, 453–458. <https://doi.org/10.1038/aja.2011.14>
- World Health Organization (1980). *WHO laboratory manual for the examination of human semen and semen-cervical mucus interaction* (1st ed.). Singapore City, Singapore: Press Concern.
- World Health Organization (1987). *WHO laboratory manual for the examination of human semen and semen-cervical mucus interaction* (2nd ed.). Cambridge, UK: Cambridge University Press.

- World Health Organization (1992). *WHO laboratory manual for the examination of human semen and sperm-cervical mucus interaction* (3rd ed.). Cambridge, UK: Cambridge University Press.
- World Health Organization (1999). *WHO laboratory manual for the examination of human semen and sperm-cervical mucus interaction* (4th ed.). Cambridge, UK: Cambridge University Press.
- World Health Organization (2010). *WHO laboratory manual for the examination and processing of human semen* (5th ed.). Geneva, Switzerland: WHO Press.
- Yániz, J. L., Soler, C., & Santolaria, P. (2015). Computer assisted sperm morphometry in mammals: A review. *Animal Reproduction Science*, 156, 1–12. <https://doi.org/10.1016/j.anireprosci.2015.03.002>

How to cite this article: Schubert B, Badiou M, Force A. Computer-aided sperm analysis, the new key player in routine sperm assessment. *Andrologia*. 2019;00:e13417. <https://doi.org/10.1111/and.13417>