

Validation of the Sperm Class Analyser CASA System for Sperm Counting in a Busy Diagnostic Semen Analysis Laboratory

CHEY DEARING¹, SALLY KILBURN² & KEVIN LINDSAY¹

¹Andrology Laboratory, Hammersmith Hospital, Imperial College NHS Trust, London, United Kingdom, W120HS2, and ²Portsmouth University, James Watson West, 2 King Richard 1st Road, Portsmouth P01 2FR

Abstract

Sperm counts have been linked to several fertility outcomes making it an essential parameter of semen analysis. It has become increasingly recognised that Computer Assisted Semen Analysis provides improved precision over manual methods but systems are seldom robustly validated for use. The objective of this work was to gather the evidence to validate or reject the Sperm Class Analyser as a tool for routine sperm counting in a busy laboratory setting. The criteria examined were comparison with the Improved Neubauer and Leja 20µm chambers, within and between field precision, linearity from a stock diluted in both semen and media, accuracy against internal and external quality material, an assessment of uneven flow effects and a ROC analysis to predict fertility in comparison with the Neubauer method. The technology is not a stand alone “black box” but rather a tool for well trained staff that allow rapid, high number sperm counting providing errors are identified and corrected. The system will produce rapid, accurate, linear, precise results with less analytical variance than manual methods that correlate well against the Improved Neubauer chamber. The system provides superior predictive potential for diagnosing fertility problems.

Keywords: *Computer Assisted Semen Analysis, CASA, Validation, Semen Analysis, Sperm Count, Sperm Class Analyser*

Introduction

Semen analysis is generally regarded as mandatory in the assessment of male infertility (Jungwirth et al., 2012, Mortimer & Mortimer, 1999, N.I.C.E, 2012, Rowe & Comhaire, 2000). In the United Kingdom (UK) the National Institute for Clinical Excellence (NICE) recommends semen analysis for all male partners of couples who have been unsuccessful with natural conception for greater than 12 months (N.I.C.E, 2012). The two major quantifiable elements of a semen ejaculate are the total number of spermatozoa present and the total volume of fluid. While sperm concentration itself is not a direct measure of testicular function, as final semen volume includes secretions from the seminal vesicles and prostate (Eliasson, 1975), the total ejaculate spermatozoa number is shown to be related to various fertility endpoints (Bonde et al., 1998, Cooper et al., 2010, Krause, 1996, Larsen et al., 2000, Merviel et al., 2010, Slama et al., 2002, Sripada et al., 2010, Vanweert et al., 2004, Zinaman et al., 2000). Additionally once this dilution effect and abstinence is accounted for, the total number of sperm in the ejaculate does correlate with testicular volume (Behre et al., 1997, Handelsman et al., 1984, W.H.O, 1987) and is indicative of the sperm production capability of the testes (MacLeod & Wang, 1979). The sperm count thus is a critical parameter in the semen analysis profile.

Sperm counts from 4500 men from 14 countries whose partner achieved a pregnancy within 12 months exhibit a large range of counts; 9×10^6 at the 2.5th percentile, 73×10^6 at the 50th percentile, 259×10^6 at the 97.5th percentile (Cooper et al., 2010). The World Health Organisation (WHO) use the 5th percentile from this data as a one sided reference range for sperm counting (W.H.O., 2010a); thus men whose sperm counts are equal to or higher than 15×10^6 (95%CI 12-16) fall within the 95% reference range for fertile men. The WHO acknowledges that the reference range can only be used as a guide as semen characteristics are not exclusive determinants of a couple's fertility; men below the range will still exhibit fertility and men above the range are not guaranteed fertility (W.H.O., 2010a). In addition to counting sperm in semen, counting separated samples prior to insemination in assisted fertility treatments (AFT) can offer predictive information. The post-wash total motile sperm count (TMC) as a predictor of intrauterine insemination (IUI) success can demonstrate high specificity for non-pregnancy (Vanweert et al., 2004). The predictive value of counting sperm for In Vitro Fertilisation (IVF) or Intra Cytoplasmic Sperm Injection (ICSI) are less clear (Kini et al., 2010). Few studies discuss the implications of analytical variance on sperm count predictors.

To count spermatozoa the WHO recommends (W.H.O., 2010b) that a minimum of 200 sperm should be counted in replicate and if the replicates are not close new dilutions should be made and the process repeated. One of the reasons for the minimum of 200 sperm can be observed with simple standard error calculations on Poisson distributions, the WHO utilise the simplified

formula ($N \pm (2 \times \sqrt{N})$) (W.H.O., 2010c)). For example If a laboratory was utilising a standard dilution of 1/20 on all samples and counted only 20 sperm on a given sample, the 95% confidence interval of the result is between 11 and 30×10^6 . Because the result covers such a large range it offers reduced clinical information. Unfortunately techniques of semen analysis have previously been shown to be poorly implemented at many routine laboratories (Keel, 2004, Keel et al., 2002, Matson, 1995, Pacey, 2006b, Riddell et al., 2005). Additionally it has been suggested that those performing the analysis fail to understand the implications of smaller sample sizes on confidence intervals (Pacey, 2006a). The proportion of UK semen analysis laboratories that are actually counting a minimum of 200 sperm twice is unknown. Because semen analysis is time consuming (W.H.O., 2010b) and because of ever present cost pressures on technician time, laboratories often admit to adopting methods that allow smaller numbers to be analysed (KS Lindsay, personal communication, May 2012). Certainly the UK National External Quality Assessment Service (UKNEQAS) laboratory variance results would suggest that the WHO sperm number recommendations are not universally followed (Pacey, 2006a).

To count spermatazoa the WHO recommends counting diluted fixed sperm using a 100 μ m deep haemocytometer chamber and gives particular reference to the improved Neubauer chamber (W.H.O, 2010). The additional recommendations (W.H.O., 2010d) for accurate counting with the improved neubauer are lengthy, and are designed to minimise error associated with the method. Like the difficulties associated with sample numbers, with what

stringency UK laboratories follow the recommendations are unknown. The WHO recognises that disposable chambers for sperm counting are also available (Brazil et al., 2004, Mahmoud et al., 1997, Seaman et al., 1996), but warn that they may generate significantly different results from the improved Neubauer haemocytometer (W.H.O, 2010). The WHO recommends that chambers other than the haemocytometer can be used as long as they are extensively validated with the following threefold approach: (1) Checking the new chambers dimension measurements (2) comparing results against the improved Neubauer and (3) obtaining satisfactory performance with external quality control. The majority of these chambers are shallower than haemocytometers and fill by capillary action. Uneven sperm distribution has been noted with disposable chambers (Douglas-Hamilton et al., 2005a, Douglas-Hamilton et al., 2005b) and although correction factors for this effect are available (Douglas-Hamilton et al., 2005b), they are not universally accepted as adequate (Björndahl & Barratt, 2005). In contrast to uneven sperm distribution, greater precision has been shown in some 20µm deep chambers when compared with the WHO recommended deeper improved Neubauer chamber (Mahmoud et al., 1997).

An obvious solution to the uncertainty of low sperm number counting and constraints on technician time is the use of Computer Assisted Semen Analysis (CASA). CASA concentration and motility data has been shown to relate to various fertility measures (Barratt et al., 1993, Check et al., 1990, Donnelly et al., 1998, Garrett et al., 2003, Irvine et al., 1994, Krause, 1995, Larsen et al., 2000, Liu et al., 1991, Shibahara et al., 2004). CASA almost exclusively uses the shallower counting chambers and any validated CASA

method will have to validate the chambers used in addition to the CASA method. Although barriers to routine diagnostic application such as overestimation due to the recognition of other particulate matter as sperm has been well documented (ESHRE, 1998), other systems and chambers have been shown to accurately count sperm (Agarwal & Sharma, 2007, Garrett et al., 2003, Tomlinson et al., 2010, Zinaman et al., 1996). The WHO recognise that CASA offers improved precision over manual methods and can now be applied to routine analyses such as sperm counts, provided both adequate quality control procedures and high measurement standards are followed (W.H.O., 2010e). CASA is particularly suited to high number analysis and can count far greater numbers than manual methods allow. In addition it theoretically removes bias inherent within the laboratory encountered from multiple operators using manual methods. As the clinical value of semen analysis with manual methods has been questioned (Alvarez et al., 2003) process validated methods that can count higher numbers of sperm with precision may be shown in time to be preferable over manual methods.

Materials and Methods

Population and Chamber Comparisons

352 patients attending the Authors Laboratory for semen analysis had their sperm count estimated by both the WHO Improved Neubauer method (Tomlinson, 2012, W.H.O., 2010b) and SCA CASA (SCA V 4.0, MICROPTIC S.L. Viladomat, 321 - 6^o 4^a 08029 Barcelona, Spain). All patient samples with sperm were directly compared with the CASA system. Samples were left standing to liquefy (maximum of one hour), and then thoroughly mixed on a vortex mixer to ensure a homogeneous sample. For Neubauer counts a 1:19 dilution of sample into 1% formal saline was made using 100µl of semen with a positive displacement pipette. The dilution was mixed well on a vortex mixer and transferred with a glass capillary into an Improved Neubauer haemocytometer. For Leja and CASA counts, samples were well mixed on a vortex mixer and transferred with a glass capillary into a chamber. All chambers were allowed to settle, and then sperm cells were counted by phase contrast light microscopy. CASA counts were performed both uncorrected and corrected. Uncorrected counts are achieved by simply taking the CASA count given by the system irrespective of counting errors. Corrected CASA counts were obtained by continuously repeating a one second video of a random single field of view from the leja 20µm slide. All the sperm visible on that single field are counted manually; dividing this result by a factor of 6 converts the result to 10⁶/ml. All CASA counts were performed at 100x magnification. Azoospermic samples and post vasectomy samples without sperm and samples with high Nucleated Cells Other than Sperm (NCOS) counts / debris were used for blank and recovery experiments.

Accuracy

Leja 20 μ m chambers (*20 μ m Leja slides, Leja; Gynotec Malden, Nieuw Vennep, The Netherlands*) were prepared with two levels of accubead quality control material (*Hamilton Thorne Biosciences, 100 Cummings Center, Suite 465C, Beverly, MA 01915*). 20 chambers were completely filled with a well mixed aliquot of accubead quality control material (10 chambers for each concentration) and left to stand for 5 minutes. Each chamber was analysed using three then five fields of analysis at 100x magnification.

External Quality Control

UK NEQAS samples used for this project are pooled semen fixed in 10% formalin prepared by the UK NEQAS Reproductive Science Scheme (*Reproductive Medicine, Andrology Laboratories, Saint Mary's Hospital, Manchester, United Kingdom, M13 9WL*). 16 UK NEQAS samples reflecting one year of returns were analysed by SCA CASA and compared to the results of five trained staff.

Precision

Single field and means of three and five fields of analysis was estimated by analysing a minimum of 10 chambers each 10 times to generate CV% for two levels of accubead quality control material, immotile sperm in semen, motile sperm in semen and motile sperm in media. Precision was also calculated for varying numbers of sperm (estimated as the mean number of sperm per motile sperm field). Precision profiles were generated for motile sperm and the effects of increasing field number were also examined.

Linearity

Linearity was estimated on both sperm in semen and swim sperm in media. A single pool of semen was made as stock solution from the samples of 10 men with high ($>80 \times 10^6/\text{ml}$) sperm counts. The pooled sample was serially diluted with pooled cell / debris free seminal plasma. Pooled swim sperm from 10 men was centrifuged, the supernatant decanted, and the pellet re-suspended in 1 ml of flushing media (*Origio 1084 1076 Origio a/s Knardrupvej 2 DK-2760 Måløv Denmark*). The concentrated sample was then serially diluted with flushing media. Neat and diluted samples were analysed by the SCA CASA system using means from 5 fields on analysis.

Chamber Bias, Dimensions and Segre-Silberberg effects

Nine specific areas were identified and counted on 10 Leja 20 μm chambers for two levels of accubead quality control material. Area 1 was counted first and repeated at the end to test for dehydration effects that occurred during the experiment. Slides had 10 single field analysis made on each area to examine if single field precision was different between areas. A further test of the SS effect was performed with 16 semen samples of immotile sperm measured for the same 9 areas as described.

Recovery Experiments

10 separate SCA CASA count estimations were made on concentrated pooled semen spiked into cell free pooled seminal plasma and sperm free seminal plasma with a high concentration of NCOS / debris. Percentage recovery was calculated for each estimation.

Clinical Decision Tables

Agreement between both the Improved Neubauer and the Leja 20 μ m manual method with uncorrected and corrected SCA CASA counts for three clinically important ranges (1-5 x 10⁶, 6-15 x 10⁶, and > 15 x 10⁶) were made. The number of comparisons, correlation within the given range, p value and the percentage agreement of results within the range were calculated for each separate clinically significant range.

ROC Analysis

1011 intrauterine insemination (IUI) preparation results from 530 patients between February 2008 until May 2012 were collated and analysed. All men who achieved a pregnancy with their partner after IUI from that or subsequent samples (191 semen samples from 102 men) were compared to all men who did not achieve a pregnancy with IUI (823 semen samples from 436 men). Separate comparisons were made using SCA CASA data from May 2012 until December 2012 (16 successful couples, 74 unsuccessful).

Statistics

Correlations were performed by Spearman's and regression lines calculated using Deming's regression for all analysis between manual counting chambers and SCA CASA. Linear regression was used for linearity experiments. Bland-Altman plots were constructed to test for bias. Significant differences between counts from 9 different areas of the Leja 20 μ m chamber were tested for by ANOVA with Dunns as a post hoc test. Significant differences between proximal, central and distal areas of the chamber were

tested by unpaired t test. The data was analysed using Prism version 4.0 (*GraphPad Version 4.01, San Diego, CA, USA, www.graphpad.com*) on a Samsung personal computer (*Samsung NC10 Samsung Electronics, [Samsung.com](http://www.samsung.com)*). Receiver Operator Characteristic graphs and data were produced using IBM SPSS (*<http://www.ibm.com/software/uk>*)

Results

Chamber Comparisons

The sperm count results of trained staff from the Andrology Laboratory at Hammersmith Hospital (combined results of four operators using both Improved Neubauer and Leja 20 μ m manual counting methods) showed a strong correlation with SCA CASA ($r^2 = 0.95$) without correcting for counting errors (see Figure 1). A strong correlation with SCA CASA was also evident when the two manual counting methods were separated, Neubauer ($r^2 = 0.94$) and Leja ($r^2 = 0.92$). However overestimation of low (<15x10⁶) sperm counts was noted and resulted in poor correlations at those concentrations: counts $\leq 5 \times 10^6$ Neubauer ($r^2=0.66$) and Leja ($r^2=0.56$), counts 5-15x10⁶ Neubauer ($r^2=0.64$) and Leja ($r^2=0.74$).

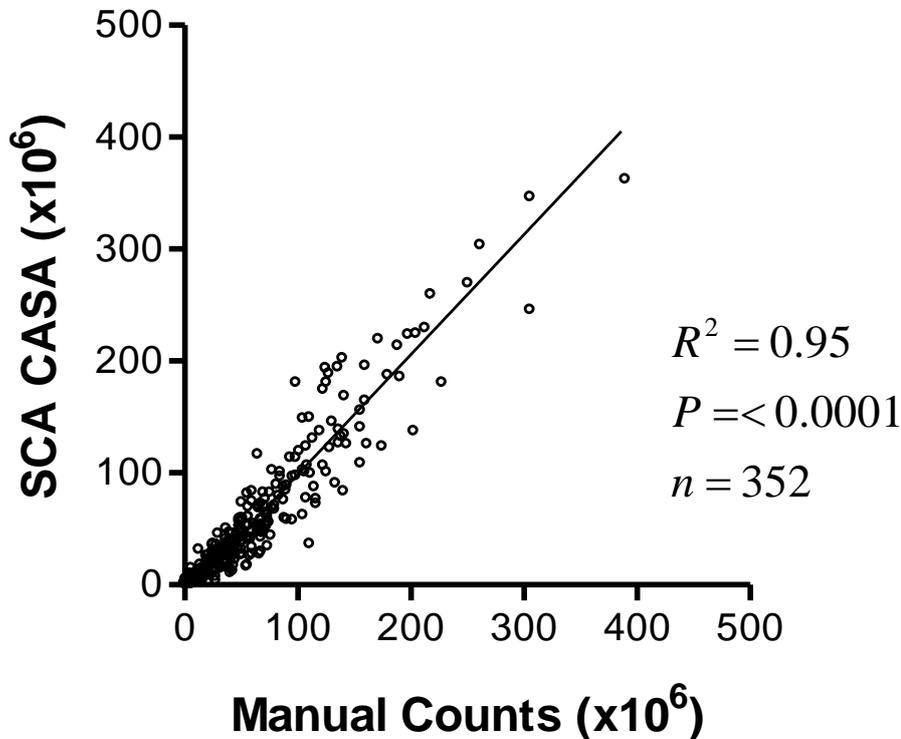


Figure 1. Correlation of Manual Counting Methods with SCA CASA. Spermatozoa counts for 352 patient's samples were measured by Neubauer or Leja 20 μ m chambers and by the SCA CASA system.

Overestimation was demonstrated on Bland-Altman plots as a moderate positive bias at low counts compared to the Improved Neubauer chamber and a large positive proportional bias compared to the Leja 20 μ m chamber. The positive bias was increased at counts $< 30 \times 10^6$ and reduced at counts $\geq 30 \times 10^6$ (see Fig 2). To correct for this bias Further tests on samples with sperm counts $< 30 \times 10^6$ demonstrated that capturing a one second video from SCA CASA and manually counting all the sperm on the computer screen correlated highly ($r^2 = 0.99$) with the manual Leja 20 μ m method. Using this method of performing counts from the CASA screen for results $< 30 \times 10^6$ combined with

uncorrected CASA counts for results $\geq 30 \times 10^6$ resulted in excellent overall correlation against the Improved Neubauer method ($r^2 = 0.98$). There was no bias observed using the one-second videos from the SCA CASA screen (see Figure 2).

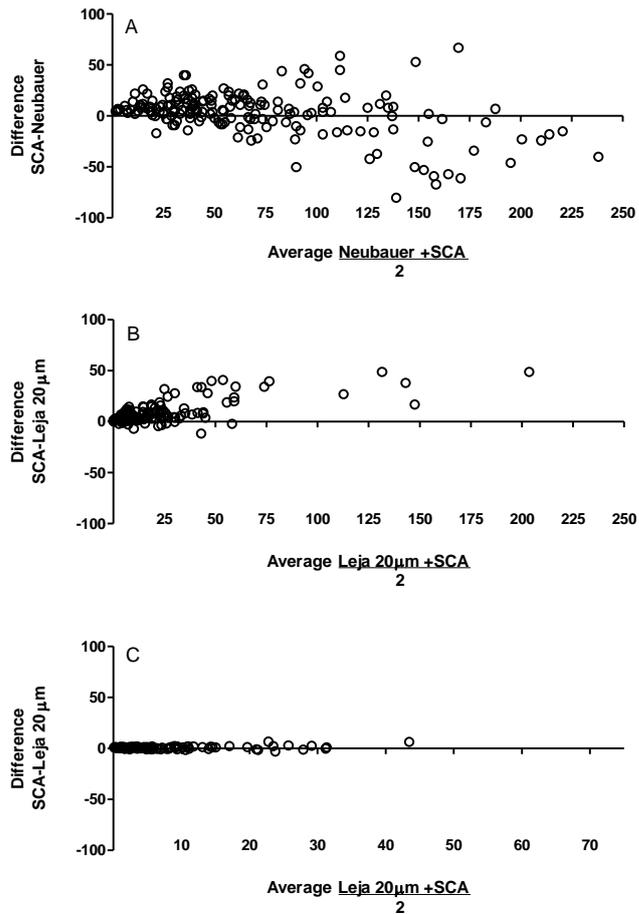


Figure 2. Bland-Altman of SCA CASA bias. (A) Difference (SCA count – Improved Neubauer Count) vs average count. CASA results are not corrected for observed counting errors. (B) Difference (SCA count – Leja Count) vs. average count. CASA results are not corrected for observed counting errors. (C) Difference (SCA screen count – Leja Manual Count) vs. average count. CASA results are corrected for observed counting errors.

Precision

Single field precision was excellent with mean CV% < 1% on latex beads and a mean CV% of 2.4% on immotile sperm. Unsurprisingly analysis of motile sperm resulted in higher single field CV% (mean of 3.6% for semen, mean of 4.5% for swim sperm). Both motile and immotile sperm in semen and swim sperm displayed a heteroscedastic precision profile with optimal precision at approximately 200 to 600 sperm per field. Multiple field precision was predictably higher than single field precision, though optimal multiple field precision was also approximately 200 to 600 sperm per field (see Figure 3). Precision decreased markedly below 100 sperm per field. Precision improves for each increase in field number included in the final analysis, though each increase yields a smaller improvement (see Table I).

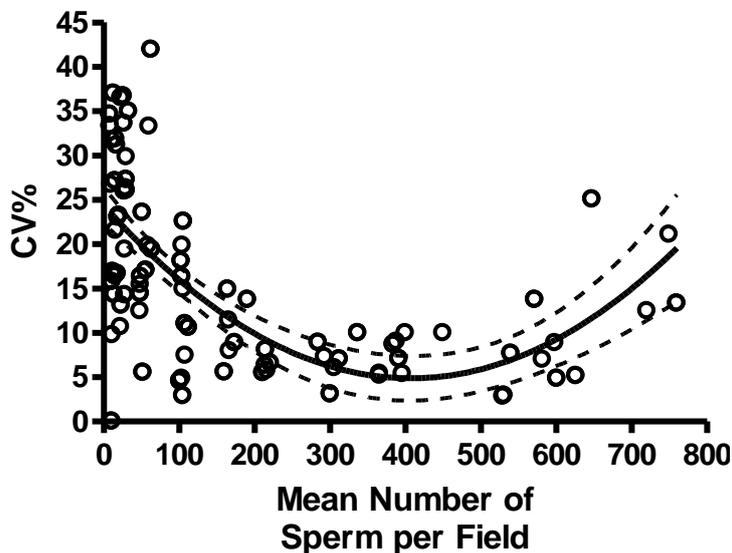


Figure 3. Motile Sperm Precision Profile. Three or five random fields were analysed for each Leja 20µm chamber by SCA CASA. Mean CV% are graphed against the mean number of sperm per field for the chamber.

Table I Effects of Field Number on Precision

Category	Field Number		
	1	3	5
Accubeads High	14.0 (13.4-14.5)	6.8 (5.9-7.7)	4.0 (3.5-4.5)
Accubeads Low	17.9 (15.8-20.0)	6.9 (3.3-10.4)	6.3 (2.9-9.7)
Immotile Sperm	18.4 (14.0-23.0)	7.6 (6.1-9.2)	5.7 (4.4-7.0)
Motile Sperm in Semen	17.2 (7.9-26.5)	7.2 (0.3-14.2)	3.0 (0.5-5.6)
Mptile Sperm in Media	15.7 (8.5-22.9)	8.2 (0.5-16.0)	6.3 (1.1-11.4)

Table I Effects of Field Number on Precision. 10 separate fields were analysed per chamber by SCA CASA to generate a multiple field CV%. The CV% was calculated for repeat measures of one, three and five fields.

Linearity

Pooled semen from 10 men with high ($>80 \times 10^6/\text{ml}$) sperm counts serially diluted with pooled cell free seminal plasma and analysed by the SCA CASA system demonstrated excellent linearity ($r^2=0.99$). Pooled swim sperm centrifuged to pellet and re-suspended in flushing media then serial diluted with flushing media also demonstrated excellent linearity ($r^2=0.95$).

External Quality Control

SCA CASA achieved the second highest correlation ($r^2 =0.97$) against one year of UKNEQAS samples when compared to five trained staff ($r^2 =0.98$, 0.93, 0.92, 0.92 and 0.72) from a laboratory that has never been a poor performer. While result bias compared to NEQAS consensus values for the Improved Neubauer was observed in two of the trained staff on Bland-Altman plots, no bias was detected on SCA CASA results.

Uneven Flow Effects

In a series of experiments testing the Segre-Silberberg (SS) effect the same 9 areas were counted on 10 Leja 20µm chambers for both levels of accubead quality control material. Distal areas had significantly ($p < 0.001$) higher latex bead counts than the central and proximal areas. This effect became insignificant once three fields of analysis including one proximal, central and distal field were included for each count. Further tests of the SS effect was performed on immotile sperm in semen measured for the same 9 slide areas demonstrated two significantly different counts using ANOVA, which again became insignificant once three fields of analysis were used to calculate a mean count.

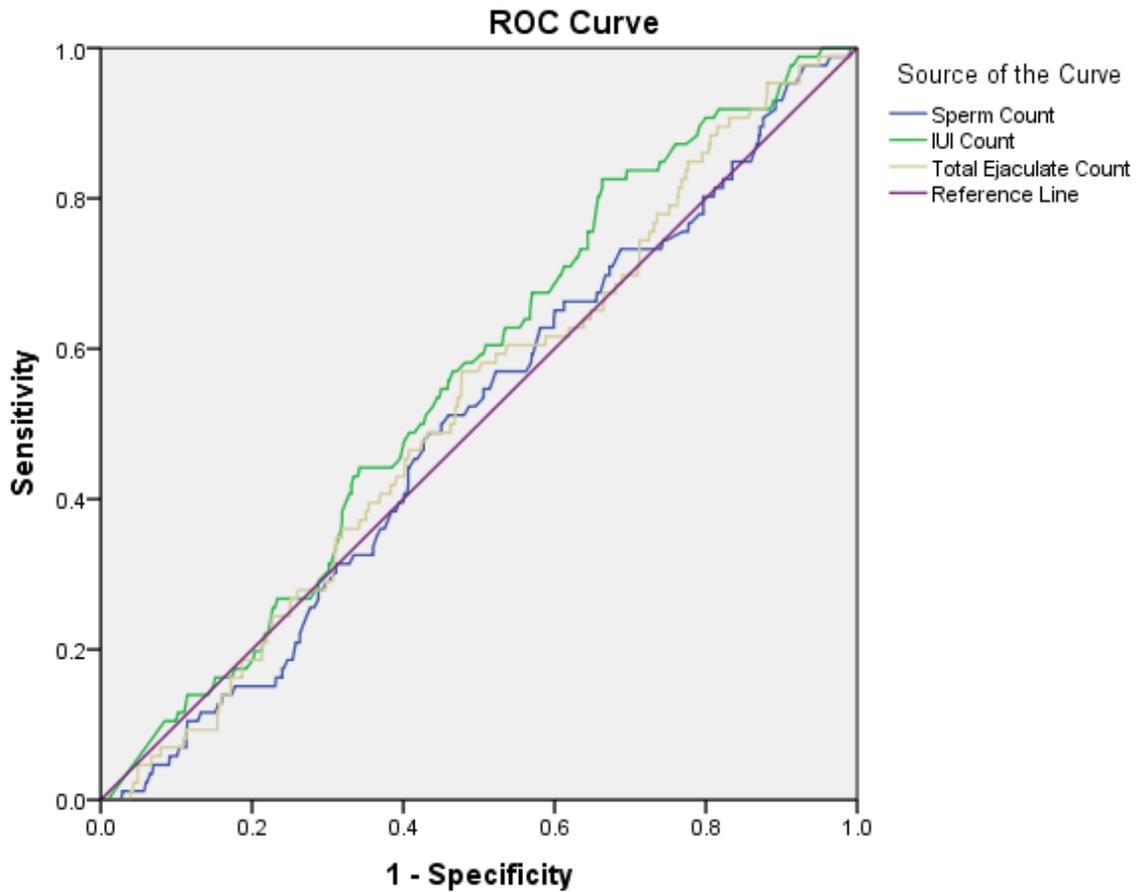
Recovery Experiments

SCA CASA count estimations from sperm spiked into cell free pooled seminal plasma and cell free seminal plasma with a high concentration of NCOS and debris demonstrated a large increase in sperm counts in samples with a high concentration of NCOS and debris (mean recovery cell free seminal plasma 98%, mean recovery NCOS and debris 176%).

The observed positive bias with uncorrected CASA misidentification of NCOS and debris as sperm, has the potential to result in clinically significant misclassification. Using the SCA CASA with no correction results in overestimations of count at concentrations below 15×10^6 . Approximately 80% of Neubauer and 33% of Leja 20µm count results between 1 and 5×10^6 were reclassified by the CASA system as being significantly higher. The SCA CASA system cannot be used without the addition of a corrective step at low

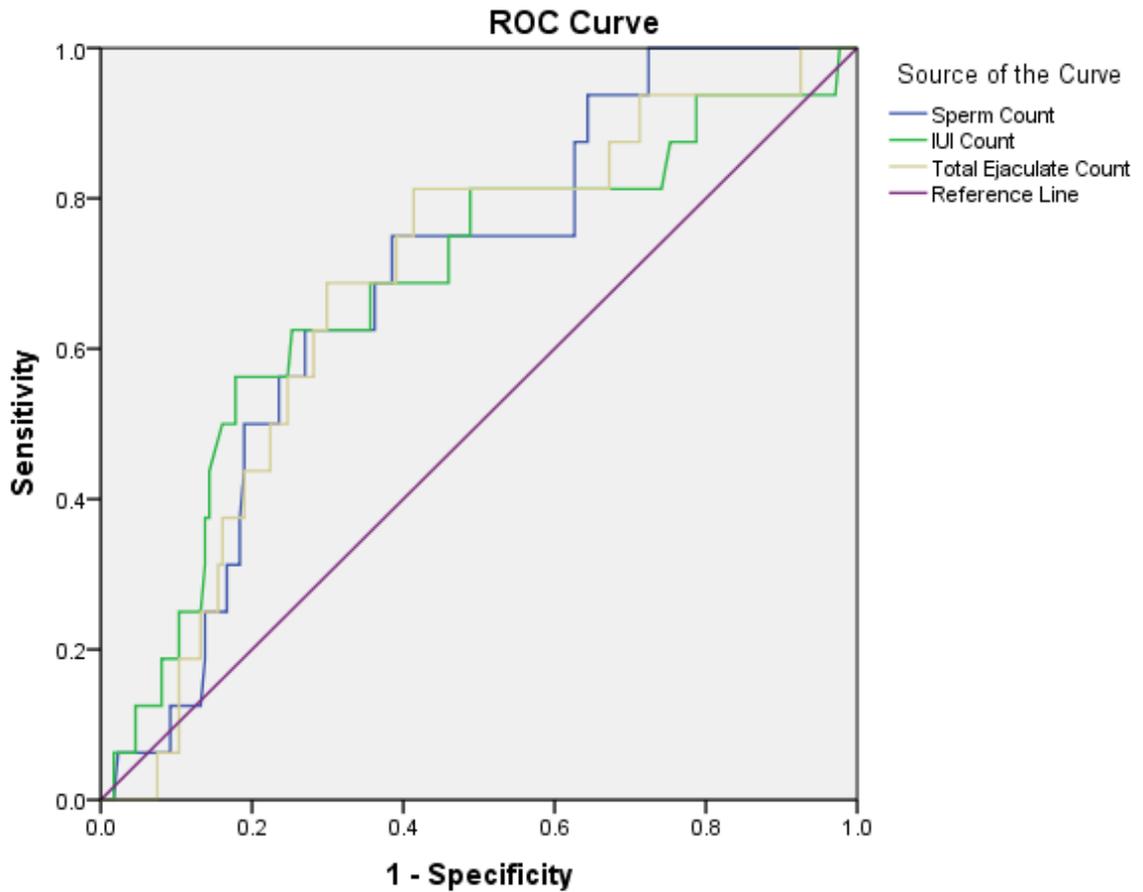
concentrations. Uncorrected SCA CASA exhibited much improved performance at concentrations above 15×10^6 . In contrast to the uncorrected SCA CASA counts, the corrected SCA CASA counts at concentrations below 15×10^6 displayed far greater agreement with the manual Leja $20 \mu\text{m}$ method. Using the manual CASA corrected count to calculate sperm concentrations $\leq 30 \times 10^6$ and the automatic (uncorrected) CASA count for sperm concentrations $> 30 \times 10^6$ results in high correlations ($r^2 = 0.98$, $p = <0.0001$) with the manual counting methods.

A comparison between using traditional Neubauer counts and CASA counts demonstrated that count data generated by CASA offered superior prediction using ROC analysis. Traditional Neubauer counts for sperm count gave an area under the curve of 0.50 (95% CI 0.44-0.56), whole ejaculate counts 0.52 (95% CI 0.46-0.58) and inseminate counts 0.56 (95% CI 0.50-0.62) (see figure 4). In comparison CASA sperm counts gave an area under the curve of 0.69 (95% CI 0.57-0.81), whole ejaculate counts 0.68 (95% CI 0.56-0.81) and inseminate counts 0.68 (95% CI 0.54-0.83) (see figure 5).



Diagonal segments are produced by ties.

Figure 4. ROC analysis of sperm count, ejaculate count and inseminate count parameters by cycle. 86 inseminates from men who achieved a pregnancy with their partner with IUI from that cycle are compared with 735 samples from men who failed to conceive with that cycle. All count data was generated from Neubauer.



Diagonal segments are produced by ties.

Figure 5. ROC analysis of sperm count, ejaculate count and inseminate count parameters by cycle. 16 inseminates from men who achieved a pregnancy with their partner with IUI from that cycle are compared with 174 samples from men who failed to conceive with that cycle. All count data was generated from SCA CASA.

Discussion

This study provided the process validation required to adopt the SCA CASA system as the main method used to count sperm in a busy diagnostic andrology laboratory. The SCA CASA system has to the author's knowledge not previously been validated for use to count human sperm. SCA CASA was compared against the Improved Neubauer chamber and performed well with the exception of low counts. A further method of counting low sperm numbers directly from the CASA screen was developed which performed well. The technology was generally well received by laboratory staff though some initial mistrust was evident. After process validation was completed, SCA CASA has provided the laboratory with a rapid precise sperm counting method that counts far greater numbers than traditional manual methods allow.

There are pros and cons of using Leja style 20µm chambers with CASA compared to manual haemocytometers when the WHO recommendations (W.H.O, 2010, W.H.O., 2010b, W.H.O., 2010d) are considered. The SCA CASA system is not advanced enough to count only whole sperm, thus all cells of similar size to sperm, including sperm without tails are counted. However several haemocytometer specific errors are also avoided. The rigorous sperm head and tail position rules for counting sperm in relation to haemocytometer chamber grid lines are unnecessary. The problems associated with coverslips such as correct thickness, poor fit, and whether or not use clamps are removed. Likewise dilution errors and positive displacement pipette errors are completely eliminated. Other sources of analytical error occurring from vortexing, filling chambers, and time frames for analysis are also at least in theory reduced. CASA certainly makes data

collection more accessible than manual counting methods, with the forthcoming necessity of uncertainty measurements for all UK diagnostic tests, this feature of CASA may be of particular interest to regulated diagnostic laboratories.

Manual semen analysis is a time consuming process that needs to be performed to a high standard with trained individuals in laboratories with robust quality assurance in order to produce clinically relevant results (Björndahl et al., 2004, Pacey, 2006a, Tomlinson et al., 1999, Tomlinson et al., 2010). However, the current cost pressures on laboratories are counterproductive to environments that produce such quality and unfortunately this is unlikely to change. Validated CASA systems have the potential to remove some sources of analytical error and dramatically increase the number of sperm analysed and thus should improve precision. CASA is more rapid than traditional semen analysis and as all technicians will be using the same automated system operator bias should be less of a problem. The training is less intense than manual methods and with automation individuals are likely less inclined or able to drift from the method. With time, increases in precision should yield more clinically relevant results and may remove some of the concerns about the clinical use of semen analysis.

Previous work has demonstrated advanced CASA systems can accurately count sperm (Garrett et al., 2003, Zinaman et al., 1996). This current work agrees with that evidence: SCA CASA was very accurate with latex bead concentrations and correlated against one year of UKNEQAS at least as well as trained andrologists without observable bias. While one year of NEQAS

data is only 16 samples and further data is required, from the initial assessment adopting SCA CASA would appear to improve NEQAS returns for our laboratory.

Uncorrected SCA CASA results correlated highly with both Improved Neubauer and Leja 20µm chambers, however low concentrations ($<30 \times 10^6$) correlations were poor due to overestimation of sperm. Overestimations at this range have clinical implications and thus uncorrected SCA CASA counts cannot be used to replace current counting methods for low sperm counts. However, capturing a one second video and manually counting sperm from the SCA CASA screen eliminates this error and results in excellent correlation at low concentration.

Poor correlations at low sperm concentrations were observed to be caused by erroneous counting of random small NCOS, cell debris and tailless sperm heads. This overestimation with CASA due to misclassification of particulate debris has been previously recognised (ESHRE, 1998, Tomlinson et al., 2010) and appears to be a common CASA problem. The positive bias was increased at results $<30 \times 10^6$ and reduced at $\geq 30 \times 10^6$. The error still occurs at higher concentrations but because of a proportional effect it appears reduced in relation to the count. There was no bias observed using the one second videos from the SCA CASA screen. Using this combined approach of reporting uncorrected counts $\geq 30 \times 10^6$ and counting from the SCA CASA screen for results $<30 \times 10^6$, bias appeared to be minimal at the critical cut off values for clinical reporting.

Precision of SCA CASA was generally excellent with optimal precision at approximately 200-600 sperm per field. This was true for both single and multiple field precision tests. Large increases in imprecision were observed at descending concentrations below 100 sperm per field, though it was demonstrated that mean results from three fields of analysis give acceptable precision for all categories tested. While increasing the field number beyond three fields does decrease imprecision the gain is not large and because of the corresponding time increase may not be warranted for routine diagnostic work.

In addition to the count overestimation errors, several other specific types of error were observed from the SCA CASA system when used with the Leja 20 μ m chambers. Infrequent counting errors occurred when SCA CASA analysed phantom motile sperm; this was always graded as WHO class b motility and invariably occurs as multiple phantom tracks. This error appears to be the result of poor focus / contrast or dirty lenses / chamber. An additional infrequent error related to poor focus / contrast was the counting of immotile sperm twice. Both of these gross errors are easily identified and removed by re-achieving Kohler illumination / Phase contrast and cleaning and optimising the microscope. Additionally sperm were occasionally not detected, however this was infrequent and while the CASA operator does have the option of adding the sperm back in manually, at 100x magnification each sperm added only increases the count by 0.2×10^6 for that particular field. Unless many sperm have been missed the error may be considered too small to warrant the time to correct. While the overestimation due to miscounting

has the potential to be a large and clinically relevant error, trained technicians should have little difficulty recognising when the error has occurred and either correcting the result or counting that particular sample with a different method. Another source of error previously described for 20 μ m chambers which relates to uneven flow of semen into the chamber and is known as the Segre-Silberberg (SS) effect (Douglas-Hamilton et al., 2005a, Douglas-Hamilton et al., 2005b). This current work found that by including one 100x magnification field from each area (distal, central and proximal) in the overall count the SS effect was not statistically significant. However the authors believe significant errors from the SS effect are still a possibility using this technology and should not be discounted. The other errors occur infrequently and are easily identified and corrected. Staff using SCA CASA will have to be specifically trained to recognise significant errors and to correct them.

The use of the method of performing manual Leja 20 μ m chamber counts from the CASA screen for results $<30 \times 10^6$ removed the majority of the overestimation bias as counts $>15 \times 10^6$ were less affected. However even occasional counts $>30 \times 10^6$ were significantly affected due to other cells and debris. It cannot be understated that significant error due to misidentification is possible at all concentrations using this technology. A recovery experiment designed to highlight this misidentification error demonstrated a maximum count increase of 76%. However the issue is somewhat alleviated by the well described negative bias between 20 μ m chambers using motile sperm and the 100 μ m Improved Neubauer chamber for fixed sperm. Proportional overestimation by SCA CASA on Leja 20 μ m counts appears to produce a result closer to that achieved with Improved Neubauer than the Leja 20 μ m.

While some overestimation does occur on all samples, the overestimation appears to be proportional to the count and results in a final count that correlates well against the Improved Nauebauer.

On the basis of this validation data, the author's laboratory adopted the SCA CASA system in May 2012 as the main system for counting sperm. It has performed well and is used for the vast majority of samples. However high viscosity samples and samples with small NCOS or debris occasionally still require manual counting. Following up pregnancy success with ROC analysis has demonstrated that the predictive potential with the new counting system has exceeded the manual methods. However it should be noted that this is not a like for like comparison because the CASA technique had been adopted as the main counting method. Further data is required but early results are indeed promising.

SCA CASA count technology is not a stand alone "black box" but rather a tool for trained andrology technicians that allow rapid, high number sperm counting. Errors will occur and individuals using the system will need specific training on how to identify and correct these errors. Providing errors are corrected SCA CASA will produce rapid, accurate, linear, precise results with less analytical variance than manual methods that correlate well against the Improved Neubauer chamber. The SCA CASA counts in comparison with manual counts on IUI data with ROC analysis suggests that the increased analytical precision of CASA may herald better predictive outcomes in time. What is required is more data from process validated CASA systems correlating CASA results with reproductive outcomes.

Acknowledgements

The authors would like to thank the staff of the Andrology Laboratory, Hammersmith Hospital, Imperial College NHS Trust, and in particular Mrs Ivy Floyd for compiling the IUI data set.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

- Agarwal, A. & Sharma, R. K. (2007). Automation is the key to standardized semen analysis using the automated SQA-V sperm quality analyzer. *Fertil Steril*, 87(1), 156-162.
- Alvarez, C., Castilla, J. A., Martinez, L., Ramírez, J. P., Vergara, F. & Gaforio, J. J. (2003). Biological variation of seminal parameters in healthy subjects. *Human Reproduction*, 18(10), 2082.
- Barratt, C. L., Tomlinson, M. J. & Cooke, I. D. (1993). Prognostic significance of computerized motility analysis for in vivo fertility. *Fertil Steril*, 60(3), 520-525.
- Behre, H., Yeung, C., Holstein, A., Weinbauer, G., Gassner, P. & Nieschlag, E. (1997). Diagnosis of male infertility and hypogonadism. *Andrology: male reproductive health and dysfunction*, 87-111.
- Björndahl, L. & Barratt, C. L. (2005). Semen analysis: setting standards for the measurement of sperm numbers. *J Androl*, 26(1), 11.
- Björndahl, L., Tomlinson, M. & Barratt, C. L. (2004). Raising standards in semen analysis: professional and personal responsibility. *J Androl*, 25(6), 862-863.
- Bonde, J. P. E., Ernst, E., Jensen, T. K., Hjollund, N. H. I., Kolstad, H., Scheike, T., Giwercman, A., Skakkebaek, N. E., Henriksen, T. B. & Olsen, J. (1998). Relation between semen quality and fertility: a population-based study of 430 first-pregnancy planners. *The Lancet*, 352(9135), 1172-1177.
- Brazil, C., Swan, S. H., Tollner, C. R., Treece, C., Drobnis, E. Z., Wang, C., Redmon, J. B. & Overstreet, J. W. (2004). Quality control of laboratory methods for semen evaluation in a multicenter research study. *Journal of Andrology*, 25(4), 645.
- Check, J. H., Bollendorf, A., Lee, M. A., Nazari, A. & Nowroozi, K. (1990). Correlation of computerized semen analysis with successful fertilization of oocytes in an in vitro fertilization program. *Arch Androl*, 24(3), 229-234.
- Cooper, T. G., Noonan, E., Von Eckardstein, S., Auger, J., Baker, H. W. G., Behre, H. M., Haugen, T. B., Kruger, T., Wang, C. & Mbizvo, M. T. (2010). *World Health*

Organization reference values for human semen characteristics. Human Reproduction Update, 163, 231-245.

Donnelly, E. T., Lewis, S. E., McNally, J. A. & Thompson, W. (1998). In vitro fertilization and pregnancy rates: the influence of sperm motility and morphology on IVF outcome. Fertil Steril, 702, 305-314.

Douglas-Hamilton, D. H., Smith, N. G., Kuster, C. E., Vermeiden, J. P. & Althouse, G. C. (2005a). Capillary-loaded particle fluid dynamics: effect on estimation of sperm concentration. J Androl, 261, 115-122.

Douglas-Hamilton, D. H., Smith, N. G., Kuster, C. E., Vermeiden, J. P. & Althouse, G. C. (2005b). Particle distribution in low-volume capillary-loaded chambers. J Androl, 261, 107-114.

Eliasson, R. (1975). Analysis of semen. Progress in infertility, 2691-713.

Eshre (1998). Guidelines on the application of CASA technology in the analysis of spermatozoa. ESHRE Andrology Special Interest Group. European Society for Human Reproduction and Embryology. Hum Reprod, 131, 142-145.

Garrett, C., Liu, D. Y., Clarke, G. N., Rushford, D. D. & Baker, H. W. (2003). Automated semen analysis: 'zona pellucida preferred' sperm morphometry and straight-line velocity are related to pregnancy rate in subfertile couples. Hum Reprod, 188, 1643-1649.

Handelsman, D. J., Conway, A. J., Boylan, L. M. & Turtle, J. R. (1984). Testicular function in potential sperm donors: normal ranges and the effects of smoking and varicocele. International Journal of Andrology, 75, 369-382.

Irvine, D. S., Macleod, I. C., Templeton, A. A., Masterton, A. & Taylor, A. (1994). A prospective clinical study of the relationship between the computer-assisted assessment of human semen quality and the achievement of pregnancy in vivo. Hum Reprod, 912, 2324-2334.

Jungwirth, A., Giwercman, A., Tournaye, H., Diemer, T., Kopa, Z., Dohle, G. & Krausz, C. (2012). European Association of Urology Guidelines on Male Infertility: The 2012 Update. European Urology.

- Keel, B. (2004). How reliable are results from the semen analysis? *Fertility and Sterility*, 821, 41.
- Keel, B. A., Stembridge, T. W., Pineda, G. & Serafy, N. T. (2002). Lack of standardization in performance of the semen analysis among laboratories in the United States. *Fertility and Sterility*, 783, 603-608.
- Kini, S., Morrell, D., Thong, K., Kopakaki, A., Hillier, S. & Irvine, D. (2010). Lack of impact of semen quality on fertilization in assisted conception. *Scottish medical journal*, 551, 20-23.
- Krause, W. (1995). Computer-assisted semen analysis systems: comparison with routine evaluation and prognostic value in male fertility and assisted reproduction. *Hum Reprod*, 10 Suppl 160-66.
- Krause, W. (1996). [Value of computer-assisted sperm analysis (CASA). Reproducibility--online documentation--prognostic value]. *Fortschr Med*, 11434, 470-473.
- Larsen, L., Scheike, T., Jensen, T. K., Bonde, J. P., Ernst, E., Hjollund, N. H., Zhou, Y., Skakkebaek, N. E. & 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. *Hum Reprod*, 157, 1562-1567.
- Liu, D. Y., Clarke, G. N. & Baker, H. W. (1991). Relationship between sperm motility assessed with the Hamilton-Thorn motility analyzer and fertilization rates in vitro. *J Androl*, 124, 231-239.
- Macleod, J. & Wang, Y. (1979). Male fertility potential in terms of semen quality: a review of the past, a study of the present. *Fertility and Sterility*, 312, 103.
- Mahmoud, A., Depoorter, B., Piens, N. & Comhaire, F. H. (1997). The performance of 10 different methods for the estimation of sperm concentration. *Fertility and Sterility*, 682, 340-345.
- Matson, P. L. (1995). External quality assessment for semen analysis and sperm antibody detection: results of a pilot scheme. *Hum Reprod*, 103, 620-625.

- Merviel, P., Heraud, M. H., Grenier, N., Lourdel, E., Sanguinet, P. & Copin, H. (2010). Predictive factors for pregnancy after intrauterine insemination (IUI): an analysis of 1038 cycles and a review of the literature. *Fertil Steril*, 931, 79-88.
- Mortimer, D. & Mortimer, S. T. (1999). *Laboratory investigation of the infertile male. A textbook of in vitro fertilization and assisted reproduction*, 2nd ed. Cambridge, Parthenon Publishing group, 53-81.
- N.I.C.E. 2012. National Institute for Clinical Excellence. CG11Fertility: Assessment and treatment for people with fertility problems [Online]. Available: <http://www.nice.org.uk>. [Accessed 30/07/2012 2012].
- Pacey, A. A. (2006a). Is quality assurance in semen analysis still really necessary? A view from the andrology laboratory. *Human Reproduction*, 215, 1105-1109.
- Pacey, A. A. (2006b). Is quality assurance in semen analysis still really necessary? A view from the andrology laboratory. *Hum Reprod*, 215, 1105-1109.
- Riddell, D., Pacey, A. & Whittington, K. (2005). Lack of compliance by UK andrology laboratories with World Health Organization recommendations for sperm morphology assessment. *Hum Reprod*, 2012, 3441-3445.
- Rowe, P. J. & Comhaire, F. H. (2000). *WHO manual for the standardized investigation, diagnosis and management of the infertile male*, Cambridge University Press.
- Seaman, E. K., Goluboff, E., Barchama, N. & Fisch, H. (1996). Accuracy of semen counting chambers as determined by the use of latex beads. *Fertility and Sterility*, 664, 662-665.
- Shibahara, H., Obara, H., Ayustawati, Hirano, Y., Suzuki, T., Ohno, A., Takamizawa, S. & Suzuki, M. (2004). Prediction of pregnancy by intrauterine insemination using CASA estimates and strict criteria in patients with male factor infertility. *Int J Androl*, 272, 63-68.
- Slama, R., Eustache, F., Ducot, B., Jensen, T. K., Jørgensen, N., Horte, A., Irvine, S., Suominen, J., Andersen, A. & Auger, J. (2002). Time to pregnancy and semen

parameters: a cross-sectional study among fertile couples from four European cities. *Human Reproduction*, 172, 503-515.

Sripada, S., Townend, J., Campbell, D., Murdoch, L., Mathers, E. & Bhattacharya, S. (2010). Relationship between semen parameters and spontaneous pregnancy. *Fertil Steril*, 942, 624-630.

Tomlinson, M., Harbottle S., Woddward B., Lindsay K. (Eds.) (2012). Association of Biomedical Andrologists Laboratory andrology guidelines for good practice.

Tomlinson, M. J., Kessopoulou, E. & Barratt, C. L. (1999). The diagnostic and prognostic value of traditional semen parameters. *J Androl*, 205, 588-593.

Tomlinson, M. J., Pooley, K., Simpson, T., Newton, T., Hopkisson, J., Jayaprakasan, K., Jayaprakasan, R., Naeem, A. & Pridmore, T. (2010). Validation of a novel computer-assisted sperm analysis (CASA) system using multitarget-tracking algorithms. *Fertil Steril*, 936, 1911-1920.

Vanweert, J., Repping, S., Vanvoorhis, B., Vanderveen, F., Bossuyt, P. & Mol, B. (2004). Performance of the postwash total motile sperm count as a predictor of pregnancy at the time of intrauterine insemination: A meta-analysis. *Fertility and Sterility*, 823, 612-620.

W.H.O (1987). WHO laboratory manual for the examination of human semen and sperm-cervical mucus interaction, Cambridge, Cambridge Univ. Pr.

W.H.O (2010). Chapter 2 Standard Procedures. In: Cooper, T. G. (Ed.) WHO laboratory manual for the examination and processing of human semen.(pp. 34). Cambridge: World Health Organization.

W.H.O. (2010a). Apendix 1 Reference values and semen nomenclature. In: Cooper, T. G. (Ed.) WHO laboratory manual for the examination and processing of human semen.(pp. 223-225). Cambridge: World Health Organization.

W.H.O. (2010b). Chapter 2 Standard Procedures. In: Cooper, T. G. (Ed.) WHO laboratory manual for the examination and processing of human semen.(pp. 33). Cambridge: World Health Organization.

W.H.O. (2010c). Chapter 2 Standard Procedures. WHO laboratory manual for the examination and processing of human semen. World Health Organization.

W.H.O. (2010d). Chapter 2 Standard Procedures. In: Cooper, T. G. (Ed.) WHO laboratory manual for the examination and processing of human semen.(pp. 35,39-40,42). Cambridge: World Health Organization.

W.H.O. (2010e). Chapter 3 Optional Procedures. In: Cooper, T. G. (Ed.) WHO laboratory manual for the examination and processing of human semen.(pp. 136,137). Cambridge: World Health Organization.

Zinaman, M. J., Brown, C. C., Selevan, S. G. & Clegg, E. D. (2000). Semen quality and human fertility: a prospective study with healthy couples. *J Androl*, 211, 145-153.

Zinaman, M. J., Uhler, M. L., Vertuno, E., Fisher, S. G. & Clegg, E. D. (1996). Evaluation of computer-assisted semen analysis (CASA) with IDENT stain to determine sperm concentration. *J Androl*, 173, 288-292.