

# Semenology training programs: 8 years' experience

Daniel R. Franken, Ph.D.,<sup>a</sup> Natalie Aneck-Hahn, D.Tech.,<sup>b</sup> Carl Lombaard, Ph.D.,<sup>c</sup> and Thinus F. Kruger, M.D.<sup>a</sup>

<sup>a</sup> Department of Obstetrics and Gynaecology, Tygerberg Hospital, University of Stellenbosch, Tygerberg; <sup>b</sup> Department of Urology, Steve Biko Academic Hospital, University of Pretoria, Pretoria; and <sup>c</sup> Department of Biostatistics, South African Medical Research Council, Tygerberg, Cape Town, South Africa

**Objective:** To evaluate the immediate effect of hands-on training on the quality of technical skills of laboratory technicians.

**Design:** Retrospective analytical study.

**Setting:** Academic institutions and private infertility clinics.

**Participants:** One hundred and ten laboratory technicians.

**Intervention(s):** None.

**Main Outcome Measure(s):** First, 110 participants from 16 African countries attended 5-day semenology workshops at Tygerberg Hospital. During these workshops the methodology as described by the World Health Organization manual for the analysis of human semen, namely, sperm concentration, motility, vitality, and morphology, formed part of the curriculum. Second, two experienced morphology readers from the Tygerberg group presented 23 1-day sperm morphology workshops in nine other countries.

**Result(s):** The semenology workshops indicated a significant improvement in the evaluation of sperm concentration only, whereas pretraining and posttraining results for motility did not differ. Vitality reports did not differ from those of the experienced worker in the first place. Calculation by means of a linear regression model showed a significant decrease in the mean posttraining scores from the pretraining scores for the sperm morphology courses.

**Conclusion(s):** Training improved technical skills for the evaluation of morphology and sperm concentration. The evaluation of graded sperm motility seems to be more difficult to teach over a short period. (Fertil Steril® 2010; ■ : ■–■. ©2010 by American Society for Reproductive Medicine.)

**Key Words:** Semenology, sperm, morphology, training

The analysis of human semen remains, especially in developing countries, the cornerstone of male fertility investigations, and therefore laboratory technician quality assurance should be handled with great responsibility. Although many laboratories claim to use the World Health Organization (WHO) manual for the analysis of human semen as a guideline, a recent survey (1) concluded that only 5% of United Kingdom laboratories adhered to the current WHO rules for the evaluation of sperm morphology. This was also the case for staining, classification, and sampling techniques (1). These reports are possibly partly responsible for the concerns expressed that analysis of human semen has become a neglected test and should be regarded as a technique of the past (2, 3).

Except for a few reports, interlaboratory comparisons between andrology laboratories are still lacking (4, 5). Reliable comparisons can be done only in cases where all laboratories use not only comparable or similar techniques but also similar standards for the evaluation of human semen, especially for sperm morphology. Despite the fact that the guidelines set for the analysis of human semen are described in detail in the WHO manual (6), our experience indicated that technicians follow self-made rules or evaluate sperm by hand-down information obtained from senior colleagues.

Hands-on training sessions seem to be mandatory to ensure that the trainees understand and follow the correct guidelines set for eval-

uating normal spermatozoa. We have demonstrated that, to maintain the reading skills needed to evaluate morphology, the role of hands-on training, as well as refresher courses, is vital (7).

The aim of the present report is to evaluate the immediate impact of hands-on training on the morphology reading skills of individuals participating in multiple international morphology workshops; and to record the technical skills of technicians in evaluating semen parameters, that is, sperm morphology, concentration, progressive motility, motility, and vitality, during annual semenology workshops presented at Tygerberg Hospital. The Department of Obstetrics and Gynaecology (University of Stellenbosch) in conjunction with the WHO's Human Reproductive Programme presented the semenology workshops.

## MATERIALS AND METHODS

### Semen and Slide Preparation

One of the important factors during sperm morphology evaluation is the preparation and staining techniques of the slides. During the current study only precleaned frosted glass slides were used. Semen smears were made according to the WHO's standards for slide preparation (6). The volume of semen that was used to prepare the slides depended on the sperm concentration in the ejaculate. In cases with a sperm concentration of  $>20 \times 10^6$  cells per milliliter we used 5  $\mu$ L semen, and if the concentration was  $<20 \times 10^6$  cells per milliliter we used 10 to 20  $\mu$ L semen. The feathering method was used to prepare the smears (6). The ideal smear will provide 8 to 12 sperm cells per  $\times 1,000$  magnification. If more cells are present the overlapping of cells usually causes poor morphology scores (4). All the workshop slides were stained by means of the Papanicolaou staining procedure (8).

### Guidelines for Identifying Normal Sperm

Strict sperm morphology is based primarily on the identification of normal sperm. For a spermatozoon to be considered normal the sperm head, neck,

Received December 7, 2009; revised April 14, 2010; accepted April 17, 2010.

D.R.F. has nothing to disclose. N.A.-H. has nothing to disclose. C.L. has nothing to disclose. T.F.K. has nothing to disclose.

Reprint requests: Daniel R. Franken, Ph.D., Department of Obstetrics and Gynaecology, Tygerberg Hospital, University of Stellenbosch, Tygerberg, South Africa (FAX: 27-21-933-3270; E-mail: drf@sun.ac.za).

midpiece, and tail must adhere to published criteria (9). Additionally, there should be a well-defined acrosomal region comprising 40% to 70% of the head area. The midpiece should be slender,  $<1 \mu\text{m}$  in width, about one and a half times the length of the head, and attached axially to the head. The tail should be straight, uniform, thinner than the midpiece, uncoiled, and approximately  $45 \mu\text{m}$  long (6). Papanicolaou staining was used in the study because it is our standard staining method (9). Apart these descriptions of normality, emphasis was placed during the training period on sperm-head morphology. The descriptions require that all "borderline" forms be considered abnormal. The sperm head should be  $4$  to  $5 \mu\text{m}$  in length and  $2.5$  to  $3 \mu\text{m}$  in width. The total length-to-width ratio should be  $1.5:1.75$ .

Sperm that are slightly tapered in the posterior region or sperm with a slightly narrower head also are considered normal, as are those with a slightly narrower head combined with a slightly tapered ( $\text{Width} \times \text{Length} = 4 \times 2.5 \mu\text{m}$ , slightly tapered post acrosomal area). These minor variations of spermatozoa represent the forms that are considered normal. Borderline forms or slightly amorphous heads are those forms with configurations that fall outside the range of normal variation (9). Strict criteria morphology therefore allows, within the classification for normality, spermatozoa with head configurations that fall outside the range of normal cells (oval shaped), but in which the variation is not pronounced enough to be classified as abnormal.

### Guidelines for Identifying Abnormal Sperm

*Head defects* include large, small, tapered, pyriform, round, and amorphous heads, vacuolated heads ( $>20\%$  of the head area occupied by unstained vacuolar areas), heads with small acrosomal area ( $<40\%$  of head area), and double heads, or any combination of these (6).

*Neck and midpiece defects* include "bent" neck (neck and tail form a  $>90$ -degree angle to the long axis of the head), asymmetric insertion of the midpiece into the head, thick or irregular midpiece, abnormally thin midpiece (i.e., no mitochondrial sheath), or any combination of these.

*Tail defects* include short, multiple, hairpin, broken tails, bent tails ( $>90$  degrees), tails of irregular width, coded tails, or any combination of these.

### Sperm Morphology Workshops

Two experienced sperm morphology readers (T.F.K. and D.R.F.) presented 23 hands-on workshops during the study period 2000 to 2009. Two hundred ninety-seven individuals enrolled for these 1-day sperm morphology workshops. The results of 35 workshops presented in Argentina ( $n = 3$ ), Brazil, Egypt, India ( $n = 3$ ), Singapore, The Netherlands, sub-Saharan African countries ( $n = 16$ ), Saudi Arabia ( $n = 3$ ), Switzerland ( $n = 3$ ), and South Africa ( $n = 3$ ) were evaluated in this report. The morphology results for all workshops were pooled. For each workshop a set of four semen smears was prepared and stained with Papanicolaou stain to serve as training material. Slides 1 and 2 were used to record the morphology reading skills of the participants during a pretraining session, and slides 3 and 4 evaluated the immediate impact of hands-on training on the reading skills during a post-training session.

The format of the workshops usually consisted of four sessions; session 1 involved a pretraining evaluation period, during which the delegates were to record the percentage of normal sperm present on two prestained slides. The second session included detailed lectures on morphologic characteristics of normal and abnormal sperm as described by Tygerberg strict criteria and the WHO 1999 manual (6). The third session was a consensus training session, during which high-quality photographs of numbered sperm cells were projected onto a large screen for evaluation by the group, and the fourth was regarded as a posttraining or test period and participants were requested to evaluate 100 sperm cells on two Papanicolaou-prestained slides.

**Sperm concentration** Following the guidelines set out by the WHO, preliminary estimation of sperm concentration was made with use of  $10 \mu\text{L}$  semen with a coverslip  $22 \times 22 \text{ mm}$ . The appropriate dilution was calculated, and with the aid of an improved Neubauer hemocytometer chamber the sperm concentration was determined.

**Motility** Sperm motility was categorized into four subgroups, namely a, b, c, and d according to whether it presented as:

- Rapid progressive:  $\geq 25 \mu\text{m/s}$  ( $\pm$  five sperm head lengths per half a tail length)
- Slow/sluggish progressive:  $\sim 24 \mu\text{m/s}$
- Nonprogressive:  $5 \mu\text{m/s}$  (less than one sperm head length)
- Immotile: —

**Progressive motility** The extent of progressive sperm motility relates to pregnancy rates (10, 11). Within each field, all rapidly progressive spermatozoa (WHO 99 class a) and slowly progressive spermatozoa (WHO class b) were counted first—and care was taken to count only cells that are present in the field at the same time. When all progressive spermatozoa had been counted, nonprogressive spermatozoa (WHO 99 class c) and immotile spermatozoa (WHO 99 class d) were counted in the same field.

**Vitality** Sperm vitality is used to test cellular integrity, assessing the ability of the sperm plasma membrane to exclude extracellular substances such as dyes, or measuring the leakage of intracellular markers, particularly enzyme. Vitality assessment also provides a check on the accuracy of motility assessments, as the percentage of live spermatozoa should slightly exceed the total percentage of motile spermatozoa. Vitality is tested with use of eosin-nigrosin (one-step staining technique) (12). This method uses nigrosin to increase the contrast between the background and the sperm head, which makes them easier to discern. It also permits slides to be stored for reevaluation and quality control purposes (13). The percentage of viable cells normally exceeds the percentage of motile cells. Spermatozoa with red or dark pink heads are considered dead whereas those with white or faint pink heads are considered alive (6).

### Statistical Analysis

The Mann-Whitney test for independent samples was used to compare the sperm concentration, motility, and vitality readings for the recorded pretraining and posttraining sessions (MedCalc for Windows, Statistics for Biomedical Research version 9.6, 2008; MedCalc Software, Mariakerke, Belgium). The morphology reading of the slides differed (from low to high values), and therefore the difference in the reading and the true morphology value was standardized with use of a SD from binomial distribution. The number of sperm read by the expert to determine the true value was 200. A mixed-effect linear regression model was used with the participant as a random effect and the timing of the morphology tests as the fixed effect.

## RESULTS

### Tygerberg Strict Criteria Sperm Morphology Workshops

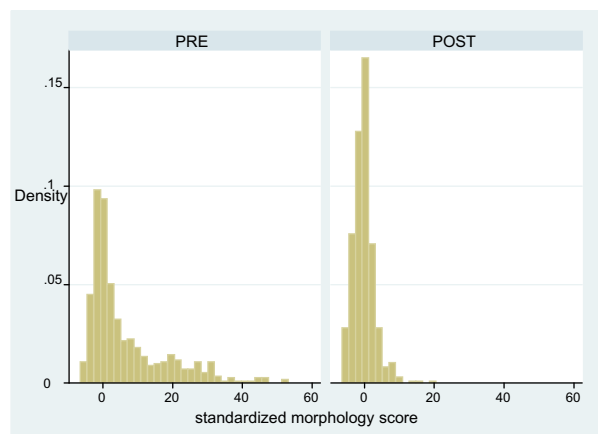
Two hundred nineteen participants were trained in 18 training events. For each training event a different set of morphology slides was prepared for each participant. Two readings were made before training and three readings after training. The pretraining scores are skewed to the right whereas the posttraining readings are more symmetrically distributed around the expected value of zero. The expert (true value) will obtain a zero score if the reading by the participant is the same as the expert reading.

### Comparison of the Pretraining and Posttraining Standardized Morphology Score

The mixed-effect linear regression model showed a significant decrease in the mean scores from the pretraining scores,  $P < .001$ . The adjusted mean scores were calculated from the mixed-effects model. The posttraining mean is equal to  $-0.19$  with 95% confidence interval of  $-0.98$  to  $0.61$ . Because this interval spans 0 it shows that the mean morphology reading after training by the participants was not significantly different from 0 ( $P = > 0.05$ ). After training the participants read the morphology slides close to the true value. In contrast the pretraining readings were substantially biased: the Z-score of 6.5 indicates this (Fig. 1).

**FIGURE 1**

Histogram of standardized scores before and after training.



Franken. Semenology training. Fertil Steril 2010.

The results of the morphology readings are presented in Table 1. Table 1 illustrates the misconceptions that exist among participants about the normal sperm cell as defined by the Tygerberg criteria. When the percent difference between slides 1 and 3 is compared, the power of information as far as normal is concerned is evident. The mean percent difference recorded for a slide containing >14% normal cells evaluated during the pretraining session decreased dramatically when a similar slide was presented in the posttraining session, that is,  $104.0\% \pm 189.1\%$  versus  $13.0\% \pm 8.01\%$  ( $P=.0001$ ). The mean percent difference recorded for a slide containing <14% normal cells evaluated during the pretraining session remained almost equal when similar slides were presented in the posttraining session, that is,  $13.0\% \pm 15.8\%$  (slide 2, pretraining),  $13.0\% \pm 8.1\%$  (slide 3, posttraining),  $8.0\% \pm 5.2\%$  (slide 4, posttraining).

## Semen Analysis Workshops

The Mann-Whitney test for independent samples recorded during the general semen analysis course for pretraining versus posttraining sessions showed the following statistical results: the data recorded for sperm concentration showed a highly significant improvement in the pretraining and posttraining values ( $P=.001$ ). The reported data for the motility evaluation, however, revealed no change in the results recorded during the pretraining and posttraining session ( $P=.25$  and  $.06$ ). Similar findings were recorded for the vitality results ( $P=.5$ ). Likewise the results recorded for vitality measurements did not improve after training ( $P=.5$ ). The morphology scores, however, again showed a dramatic improvement after the training sessions ( $P=.0001$ ).

## DISCUSSION

Laboratory technologists often are confronted with the question “How good are we at evaluating the semen sample?” The analysis of human semen in developing countries remains the cornerstone of male fertility investigations.

Since the introduction of proficiency testing, numerous reports have indicated that participation in organized proficiency testing programs has resulted in a decrease in interlaboratory SDs and coefficients of variation (14). Continual proficiency testing has revealed a dramatic improvement in the quality of clinical testing and has served to ensure better agreement of results among laboratories (15, 16). Sperm concentration evaluation showed a significant improvement from pretest to posttraining examination. There was also a decrease in interparticipant variability, as seen in other courses (7, 15).

Studies by Auger et al. (16) and Jorgensen et al. (17) found variations in motility assessments. Sperm motility assessments clearly are influenced by time, temperature, and chamber depth (17–19). The time it takes for a course participant to evaluate sperm motility can influence the final motility assessment and result in bias when compared with the expert’s evaluation, especially if the participant is a novice. However, this should not have a marked influence on the overall WHO motility grading (16). The amount of variation probably could be related to the level

**TABLE 1**

Descriptive statistics of 290 participants’ morphology scores during 18 pretraining and posttraining sessions.

	Normal cells recorded during pretraining (%)		Normal cells recorded during posttraining (%)	
	Slide 1	Slide 2	Slide 3	Slide 4
Reference slide value	15	5	14	9
Lowest value	0.0	0.0	0.0	−8.0
Highest value	98.0	75.0	60.0	33.0
Mean	29.7	12.6	13.2	4.4
95% confidence interval	26.6–33.2	10.5–14.7	12–14	3.4–5.4
Median	17.0	6.5	13.0	2.0
SD	26.6	16.2	8.0	5.6
Percent difference from mean	$104.0 \pm 189.1^{a,b}$	$13.0 \pm 15.8^{c,d}$	$13.0 \pm 8.1$	$4.2 \pm 5.7$

<sup>a</sup> Compared with slide 3,  $P=.0001$ , Mann-Whitney test for independent samples.

<sup>b</sup> Compared with slide 4,  $P=.001$ , Mann-Whitney test for independent samples.

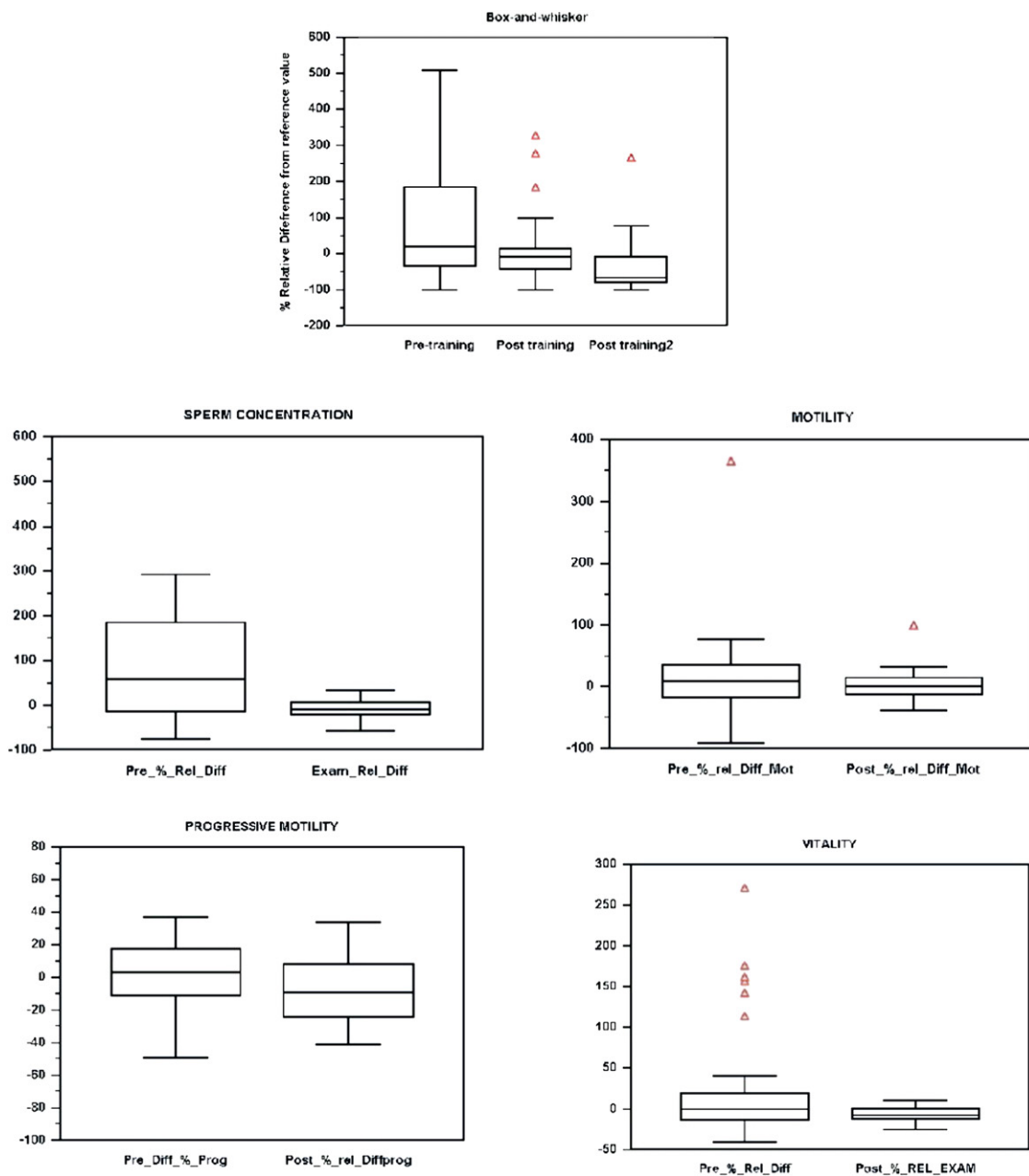
<sup>c</sup> Compared with slide 3,  $P$ =not significant, Mann-Whitney test for independent samples.

<sup>d</sup> Compared with slide 4,  $P$ =not significant, Mann-Whitney test for independent samples.

Franken. Semenology training. Fertil Steril 2010.

FIGURE 2

Box-whisker plots recorded for pretraining and posttraining results of sperm concentration, motility, and vitality. Rel\_Diff = % relative difference; Diff\_Mot = % relative difference for motility.



Franken. Semenology training. Fertil Steril 2010.

of training of the observer and also the natural ability of the observer for this subjective task (16, 20). The results also indicate that perhaps the motility assessment training received during the present course is insufficient. It must be kept in mind that the 4- to 5-day course does constitute a comprehensive training program for andrologist technologists, but it is essentially an initial baseline training for theory and practical education (12). One way of possibly overcoming the time bias is to use videotaped or DVD-taped images. Although this

deprives the participant of the experience of looking through a microscope or assessing the mixing and sampling of the semen, all participants are able to see exactly the same image (16). Variability then can be attributed directly to technologist analysis rather than biologic variations or sampling errors (21).

Similarly to the findings of Auger et al. (16) and Björndahl et al. (13), the evaluation of sperm vitality did not improve in these courses. The proportion of results within 10% of the expert values was high in the pretest but improved with training as did the group variability (Fig.

2). This can be attributed to the principle of the test (immobilized sperm, with or without staining) and its quantitative nature (18).

Identifying abnormal cells compared with the evaluation of normal cells seems to be less troublesome. Although the evaluation of sperm morphology also is done subjectively, the concept of a normal spermatozoon is understood rapidly by the technologists in training. This observation highlights the importance of hands-on training in sperm morphology.

We firmly believe that global quality control measurements in andrology laboratories will eventually become compulsory. A high-quality semen analysis still represents the cornerstone of the investigation of the infertile couple. To maintain low intratechnician and intertechnician variation and high-quality proficiency testing among laboratory technicians, continual teaching programs should be available to all (22).

## REFERENCES

- Riddell D, Pacey A, Whittington K. Lack of compliance by UK andrology laboratories with World Health Organization recommendations for sperm morphology assessment. *Hum Reprod* 2005;20:3441–5.
- McDonough P. Editorial comment: Has traditional sperm analysis lost its clinical relevance? *Fertil Steril* 1997;67:585–7.
- Chong AR, Waiters CA, Weinreb SA. The neglected laboratory test. The semen analysis. *J Androl* 1983;4: 280–2.
- Ombelet W, Menkveld R, Kruger TF, Steeno O. Sperm morphology assessment: historical review in relation to fertility. *Hum Reprod Update* 1995;1: 543–7.
- Coetzee K, Kruger TF, Lombaard CJ. Predictive value of normal sperm morphology: a structured literature review. *Hum Reprod Update* 1998;4: 73–84.
- World Health Organization. Laboratory manual for the examination of human semen and semen–cervical mucus interaction. New York: Cambridge University Press, 1999.
- Franken DR, Aneck-Hahn N. Sperm morphology training programmes: the African experience. *Hum Reprod Monograph* July 2008, p60–63.
- Menkveld R, Oettle EE, Kruger TF, Swanson RJ, Acosta AA, Oehninger SC. Atlas of human sperm morphology. Baltimore: Williams & Wilkins, 1991.
- Menkveld R, Stander FSH, Kotze TJDW, Kruger TF, van Zyl JH. The evaluation of morphologic characteristics of human spermatozoa according to stricter criteria. *Hum Reprod* 1990;5:586–92.
- Larsen L, Scheike T, Jensen TK, Bonde JP, Ernst E, Hjøllund NH, et al. Computer-assisted semen analysis parameters as predictors for fertility of men from the general population. The Danish First Pregnancy Planner Study Team. *Hum Reprod* 2000;15:1562–7.
- Zinaman MJ, Uhler ML, Vertuno E, Fisher SG, Clegg ED. Evaluation of 9549 computer-assisted semen analysis (CASA) with IDENT stain to determine sperm concentration. *J Androl* 1996;17:288–92.
- Björndahl L, Soderlund I, Kvist U. Evaluation of the one-step eosin-nigrosin staining technique for human sperm vitality assessment. *Hum Reprod* 2003;18: 813–6.
- Björndahl L, Barratt CLR, Fraser LR, Kvist U, Mortimer D. ESHRE basic semen analysis courses 1995–1999: immediate beneficial effects of standardized training. *Hum Reprod* 2002;7:1299–305.
- Franken DR, Barendson R, Kruger TF. A continuous quality control (CQC) program for strict sperm morphology. *Fertil Steril* 2000;74:721–4.
- Keel BA, Quinn P, Schmidt F, Serafy NT Jr, Serafy NT Sr, Schalue TK. Results of the American Association of Bioanalysts national proficiency testing programme in andrology. *Hum Reprod* 2000;5:680–6.
- Auger J, Eustache F, Ducot B, Blandin T, Daudin M, Diaz I, et al. Intra- and inter-individual variability in human sperm concentration, motility and vitality assessment during a workshop involving ten laboratories. *Hum Reprod* 2000;15:2360–8.
- Jorgensen N, Andersen AG, Eustache F, Irvine DS, Suominen J, Petersen JH, et al. Regional differences in semen quality in Europe. *Hum Reprod* 2001;16: 1012–9.
- Toft G, Rignell-Hydbom A, Tyrkiel E, Shvets M, Giwercman A. Quality control workshops in standardization of sperm concentration and motility assessment in multicentre studies. *Int J Androl* 2005;28:144–9.
- Le Lannou D, Griveau JF, Le Pichon JP, Quero JC. Effects of chamber depth on the 9248 motion pattern of human spermatozoa in semen or in capacitating medium. *Hum Reprod* 1992;7:1417–21.
- Kraemer M, Fillion C, Martin-Pont B, Auger J. Factors influencing human sperm kinematic measurements by the Celltrak computer-assisted sperm analysis system. *Hum Reprod* 1998;13:611–9.
- Dunphy BC, Kay R, Barratt CLR, Cooke ID. Quality control during the conventional analysis of semen, an essential exercise. *J Androl* 1989;10:378–85.
- Brazil C, Swan SH, Tollner CR, Treece C, Drobnis EZ, Wang C, et al. Study for Future Families Research Group. Quality control of laboratory methods for semen evaluation in a multicentre research study. *J Androl* 2004;25:645–56.