Human Reproduction Vol.20, No.9 pp. 2573–2578, 2005 doi:10.1093/humrep/dei061
Advance Access publication May 19, 2005

Quality specifications for seminal parameters based on the state of the art

J.A.Castilla[1,2,7], J.Morancho-Zaragoza[3], J.Aguilar[2,4], R.Prats-Gimenez[3], M.C.Gonzalvo[2,5], E.Fernández-Pardo[3], C.Álvarez[2,6], R.Calafell[3] and L.Martinez[1]

[1]Unidad de Reproducción, Hospital Virgen de las Nieves, E-18014, Granada, [2]Programa de Control de Calidad Externo de Análisis de Semen, Asociación para el Estudio de la Biología de la Reproducción (ASEBIR), [3]Programa de Supervisión Externa de la Calidad, Asociación Española de Farmacéuticos Analistas, AEFA, Madrid, [4]Unidad de Reproducción, Clínica Avicena, Jaén, [5]CEIFER, Granada and [6]Departamento de Ciencias de la Salud, Universidad de Jaén, Jaén, Spain

[7]To whom correspondence should be addressed. E-mail: josea.castilla.sspa@juntadeandalucia.es

BACKGROUND: The aim of this study was to calculate the analytical goal for seminal parameters based on the state of the art, and then to compare these specifications with those previously obtained by our group based on biological variation. METHODS: All data used for analysis were derived from the Spanish programme of external quality control on semen analysis. Over 90 laboratories participated from 1999 to 2003. Using graphs of the state of the art, we also determined the numbers of laboratories that achieved quality specifications. RESULTS: The total allowable error calculated using state of the art graphs is similar to that calculated using biological variation for concentration and total motility. However, it is much higher for morphology and rapidly progressive motility. Over 80% of the laboratories achieved the minimum quality specification based on biological variation for concentration, total and progressive motility. However, only ~30% of the laboratories achieved the minimum quality specification based on biological variation for morphology and rapidly progressive motility. CONCLUSIONS: The study enabled us to identify the state of the art of analytical performance for seminal parameters, and revealed the difficulty inherent in meeting the quality specifications based on biological variation.

Key words: analytical goals/quality control/semen/state of the art

Introduction

Quality control and a quality assurance programme are an essential part of the andrology laboratory (McCulloh, 2004). Quality assurance programmes for the clinical laboratory require an assessment of accuracy, i.e. the closeness of the measured result to the real value, using internal and external quality assessment. The control of the analytical process is concerned with maintaining test results within required limits. The maximum allowable analytical error can be defined by various strategies (Fraser *et al.*, 1999) and these have been ordered hierarchically (Kenny *et al.*, 1999). Ideally, quality specifications should be derived objectively from an analysis of medical needs. Unfortunately, this is very difficult and the necessary calculations have been made for only a few analytes in a limited number of different clinical settings (Petersen *et al.*, 1999). Other strategies that have been recommended for determining quality specifications include professional recommendations (guidelines by national or international expert groups or by expert individuals or institutional groups), those established by regulation or by external quality assessment schemes (EQAS), or those derived from data on the state of the art (Table I).

Quality specifications based on components of biological variation, within and between subjects, have been proposed by various professional groups (Fraser, 2001). Cotlove *et al.* (1970) proposed that a desirable quality specification expressed as the analytical coefficient of variation for assays should be equal to or less than half of the within-subject biological coefficient of variation (<0.5 CVBw). However, for assays that with currently available techniques could not easily meet this analytical goal, Fraser *et al.* (1997) suggested a minimum analytical goal expressed as the analytical coefficient of variation of <0.75 CVBw. For assays for which it is easy to meet desirable standards, the same authors suggested an optimum quality specification expressed as the analytical coefficient of variation of <0.25 CVBw.

The components of a semen analysis are often accepted at face value without regard for errors. However, previous studies have shown that there is a high degree of variation between results (Neuwinger *et al.*, 1990; Jorgensen *et al.*, 1997; Auger *et al.*, 2000). The degree of variation is due in part to inadequate technical training and to an absence of commonly accepted laboratory standards. To address the problem of laboratory standards for semen analyses, the World

Table I. Hierarchy of models to set analytical quality specifications in laboratory medicine (Kenny *et al.*, 1999)

1. Evaluation of the effect of analytical performance on clinical outcomes in specific clinical settings
2. Evaluation of the effect of analytical performance on clinical decisions in general:
a. Data based on components of biological variation
b. Data based on analysis of clinicians' opinions
3. Published professional recommendations
a. From national and international expert bodies
b. From expert local groups or individuals
4. Performance goals set by:
a. Regulatory bodies
b. Organizers EQAS
5. Goals based on the current state of the art
a. As demonstrated by data from EQAS or proficiency testing scheme
b. As found in current publications on methodology

Health Organization (World Health Organization, 1999) and the European Society for Human Reproduction and Embryology (ESHRE) have published manuals in this respect (Kvist and Björndahl, 2002).

We have reported previously a study of analytical goals for semen parameters using the components of biological variation (Álvarez *et al.*, 2003), following the above-mentioned recommendations of Fraser *et al.* (1997) (<0.75 CVBw). The model is based on healthy subjects, sperm donor candidate and strict protocol-controlled conditions (e.g., three to four abstinence days, same period of study, same analytical procedure, same frequency per sample). However, the lack of a standardized methodology used by those seeking to obtain the values of the components of biological variability, together with the fact that it is unclear whether biological variation components derived from healthy subjects can be extrapolated to pathological situations (Ricós *et al.*, 1999), limit its use.

For these reasons, it is necessary to obtain analytical goals from another model. The use of the state of the art has been proposed by many bodies, including the French Society of Clinical Biology (Vassault *et al.*, 1999) and the Spanish Association of Analytic Pharmaceutics (AEFA) (Calafell *et al.*, 2002). Comparison of analytical quality can be accomplished through reference to the performance achieved by the best laboratories participating in EQAS. The aim of this study was to calculate the analytical goal for seminal parameters based on the state of the art, and then to compare these specifications with those obtained previously by Álvarez *et al.* (2003) based on biological variation.

Materials and methods

All data used for analysis were obtained from the Spanish programme of external quality control on semen analysis under the auspices of the Association for the Study of Reproductive Biology (ASEBIR), the Spanish Association of Medical Biopathology (AEBM) and the AEFA, with over 90 laboratories throughout Spain participating in the programme between 1999 and 2003. Proficiency testing programmes were developed for the determination of sperm count, sperm motility, sperm morphology and sperm vitality. The quality control material mailed in 10 (twice per year) trials and used in this study were as follows: two aliquots of formalin sperm suspension for sperm concentration (a total of 20 samples), at least two samples on videotape for motility (24 samples), two unstained (20 samples) and two stained semen smears (20 samples) for morphology and two eosin Y–Nigrosin semen smears (20 samples) for vitality.

The samples used for the Spanish programme of external quality control on semen analysis were obtained from donor candidates, all of whom previously gave their informed consent for their ejaculates to be used in the investigation. Serum studies were performed for HIV, hepatitis B, hepatitis C and syphilis. All were negative.

The data obtained from different laboratories were analysed according to the following protocol. First, outlying results were identified as described by Thienpoint et al. (1987) and excluded from analysis. After outliers had been excluded, all the data were normally distributed as confirmed by the Kolmogorov–Smirnov test. Next, the target value was defined as the mean of the remaining observations and was calculated for each seminal parameter in every control sample. The total error (TE) was obtained by subtracting the target value (TV) from the result submitted by the laboratory ignoring outliers (X) dividing by the former and multiplying by 100:

$$TE = ((X - TV)/TV)*100$$

The number of results sent by each laboratory for each seminal parameter for the whole set of control samples was determined (Table II). The proportion of laboratories that reported results for a given seminal parameter within the error margins for at least 75% of the quality control samples received was calculated. A similar calculation was done for the proportion of laboratories reporting all results (100%) within the error margins. Figures 1 and 2 show, for each seminal parameter, a cumulative percentage of laboratories (*y*-axis) that reported results within an increasing error from the target value (*x*-axis). These results were calculated, on one hand, for laboratories that returned results on at least 75% and, on the other, for those that returned results on 100% of the samples received.

The total allowable error based on the state of the art was calculated according to the three levels of analytical goal (optimum, desirable and minimum) proposed by AEFA (Calafell *et al.*, 2002). Optimum quality specification is that which is obtained by the best 25% of laboratories that reported results for a given seminal parameter on all of the quality control samples received. These are the 25% of laboratories that most closely approach the target value and therefore they are the 25% of all laboratories with lowest TE in their results. The specification is obtained by examining the location in the state of the art graphs (Figures 1 and 2) of the cut-off point

Table II. Number of laboratories that returned data for a given seminal parameter on at least 75% or 100% of the samples received

Seminal parameter	All laboratories	Returned data on at least 75% of the samples received	Returned data on 100% of the samples received
Concentration	104	76	60
Morphology unstained smears	97	72	57
Morphology stained smears	97	76	66
Total motility	89	72	60
Progressive motility	89	73	63
Progressive rapid motility	89	73	65
Vitality	100	77	60

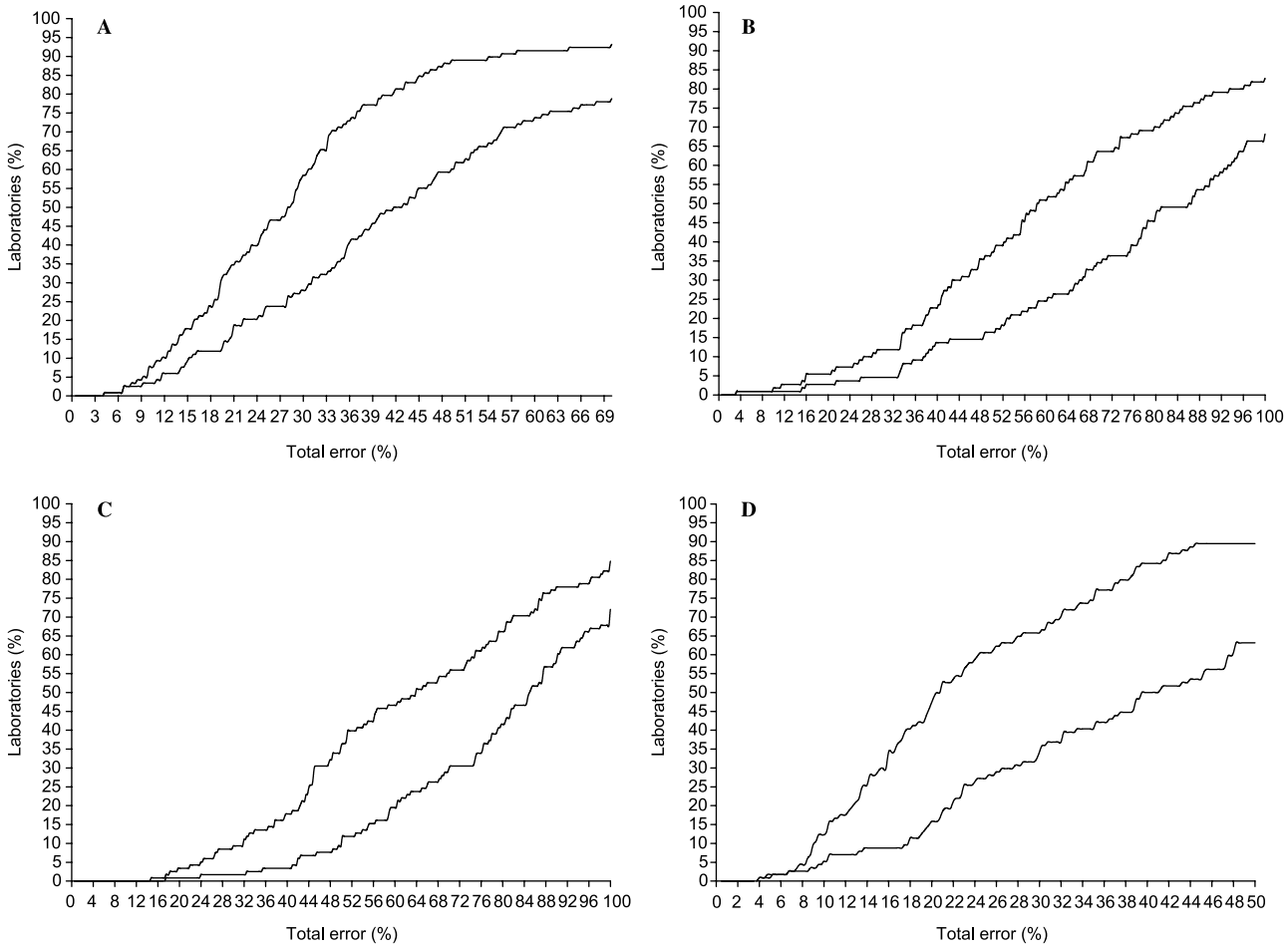

Figure 1. State of the art graphs for concentration (**A**), morphology unstained smears (**B**), morphology stained smears (**C**) and vitality (**D**). Curves calculated using data from laboratories that reported results for a given seminal parameter on at least 75% (upper curve) or on 100% (lower curve) of the quality control samples received.

corresponding to 25% of the laboratories for 100% of the samples processed. The second category, desirable quality specification, is that produced by a majority of laboratories (the best 75%) that reported results for a given seminal parameter on at least 75% of the quality control samples received. This is obtained by examining the location (Figures 1 and 2) of the cut-off point corresponding to 75% of the laboratories for 75% of the samples processed. The third type, the minimum specification, corresponds to that obtained by a majority group of laboratories (the best 90%) that reported results for a given seminal parameter on at least 75% of the quality control samples received. This specification is obtained by examining the location (Figures 1 and 2) of the cut-off point corresponding to 90% of the laboratories for 75% of the samples processed. The laboratories included in the group of the best 25% laboratories for a given seminal parameter are included in the group of the best 75% laboratories, and these are included in the group of the best 90% of laboratories for that seminal parameter.

Table III. The quality specifications expressed as a percentage of total error based on the components of biological variation proposed by Fraser *et al.* (1997)

Performance	TE
Minimum	$\leq 1.65*0.75CV_{Bw} + 0.375\sqrt{CV_{Bw}^2 + CV_{Bb}^2}$
Desirable	$\leq 1.65*0.5CV_{Bw} + 0.250\sqrt{CV_{Bw}^2 + CV_{Bb}^2}$
Optimum	$\leq 1.65*0.25CV_{Bw} + 0.125\sqrt{CV_{Bw}^2 + CV_{Bb}^2}$

TE = percentage of total allowable error; CV_{Bw} = within-subject variability expressed as a coefficient of variation in %; CV_{Bb} = between-subject variability expressed as a coefficient of variation in %.

The quality specifications for total allowable error based on the components of biological variation were calculated from the data reported by us previously (Álvarez *et al.*, 2003) and using the three-level model (optimum, desirable and minimum) proposed by Fraser *et al.* (1997) (Table III) In the daily practice of control material analysis, a laboratory should not commit a TE percentage exceeding that determined in the quality specification. Consider the following example: a sperm suspension has a concentration of 20×10^6/ml, and the analytical goal based on the state of the art, expressed as a percentage of TE, is 28%. This means that the laboratory in question should obtain a result between 14.4×10^6/ml and 25.6×10^6/ml in order to be considered an optimal laboratory.

Using the state of the art graphs, we also determined the number of laboratories that achieved the three levels of quality specifications for total allowable error based on the components of biological variation.

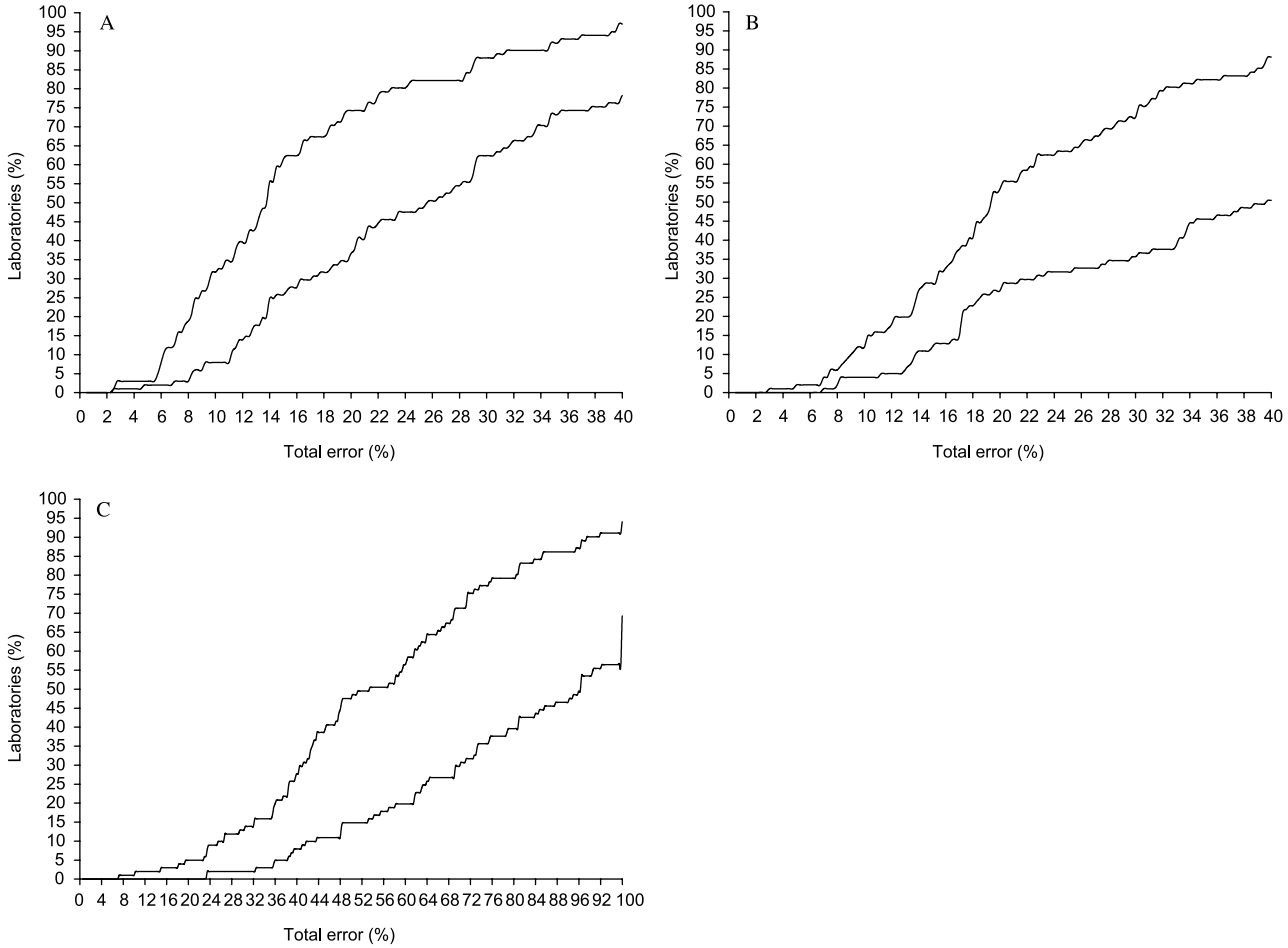

Figure 2. State of the art for total (**A**), progressive (**B**) and progressive rapid (**C**) motility. Curves calculated using data from laboratories that reported results for a given seminal parameter on at least 75% (upper curve) or on 100% (lower curve) of the quality control samples received.

Results

The total allowable error for seminal parameters (Table IV) was calculated from the state of the art graphs obtained (Figures 1 and 2). For morphology (stained and unstained smears), progressive motility and vitality, the minimum criteria were not calculated owing to the high total error obtained.

The total allowable error calculated using the state of the art graphs was similar to that calculated using the biological variation for concentration and total motility. However, it was much higher for morphology and progressive rapid motility.

The percentage of laboratories that achieved the three levels of quality specifications for total allowable error based on the components of biological variation are presented in Table V. Over 80% of the laboratories that reported on at least 75% of the quality control samples received achieved at

Table IV. Quality specifications of seminal parameters expressed as a percentage of TE obtained from biological variation or state of the art graphs

	Optimal		Desirable		Minimal	
	Biological variation	State of the art	Biological variation	State of the art	Biological variation	State of the art
Concentration	18[a]	28	37	37	56	54
Morphology	14	60[b]	28	85[b]	42	
		66[c]		88[c]		
Total mobility	11	14	23	21	35	32
Progressive motility	10	18	21	30	32	
Progressive rapid motility	14	63	29	71	43	93
Vitality	7	23	15	35	23	

[a]Numbers are the percentage of total allowable error in a given seminal parameter.
[b]Calculated using results of EQAS using unstained smears.
[c]Calculated using results of EQAS using stained smears.

Table V. Percentage of laboratories that achieved the levels of quality specifications expressed as a percentage of TE based on biological variations according to the results reported for a given seminal parameter on at least 75% or on all (100%) of the quality control samples received

	Achieved minimal level		Achieved desirable level		Achieved optimal level	
	75%	100%	75%	100%	75%	100%
Concentration	91[a]	70	75	42	24	12
Morphology unstained smears	17	8	10	5	3	1
Morphology stained smears	18	7	8	2	1	0
Total motility	94	74	80	45	35	8
Progressive motility	80	37	55	29	12	5
Progressive rapid motility	36	10	12	2	4	
Vitality	59	27	28	9	3	3

[a]Example: 91% of the laboratories that reported on at least 75% of the quality control samples received achieved at least the minimum level established of quality specifications expressed as a percentage of TE based on biological variations.

least the minimum level established of quality specifications expressed as a percentage of TE based on biological variations for concentration, total and progressive motility. More than 55% of the laboratories that reported on at least 75% of the quality control samples received achieved at least the minimum level established of quality specifications expressed as a percentage of TE based on biological variations for vitality. However, only ~30% of the laboratories that reported on at least 75% of the quality control samples received achieved at least the minimum level established of quality specifications expressed as a percentage of TE based on biological variations for morphology and progressive rapid motility.

Discussion

Quality specifications based on biological variation are similar to state of the art specifications for concentration, and total and progressive motility. However, specifications of analytical quality based on biological variation for morphology, progressive rapid motility and vitality are very different from quality specifications based on the state of the art. These discrepancies are also observed for other biological magnitudes using different models from set analytical quality specifications in laboratory medicine (Lott, 1999; Petersen *et al.*, 2001).

Several limitations of the present study should be mentioned. First, in this study the target value used was the mean value for all laboratories and not the mean value obtained from highly experienced laboratories (reference laboratories) as suggested by other authors (Cooper *et al.*, 1999). This may narrow the conclusions of the study, although different studies have indicated that even experienced groups have a noticeable amount of disagreement for seminal parameters (Neuwinger *et al.*, 1990; Jorgensen *et al.*, 1997; Auger *et al.*, 2000). Secondly, the materials used in the challenges (sperm suspension, video) are different from the semen obtained from patients or that used in the specifications based on biological variation, and participants may adopt special analytical techniques to ensure good performance, as suggested by Kvist and Björndahl (2002). Thirdly, EQAS for semen analysis require standardization, as the same laboratory may obtain different results depending on which programme it is participating in (Cooper *et al.*, 2002). Therefore, it is essential for quality control systems to be set up to ensure the homogeneity of the material analysed in external quality assessment schemes of semen analysis.

The proposed analytical goals based on biological variation for morphology are probably too strict, and would be difficult to meet. When we comply with World Health Organization recommendations and assess the proportion of normal sperm in an ejaculate in 2×200 sperm, with the true percentage of normal form being 14%, the 95% confidence interval is $\pm 24.2\%$ of the real value (10.6–17.4% of normal sperm). If we add to the counting error other error factors such as inter-observer or inter-laboratory variability (Neuwinger *et al.*, 1990; Jorgensen *et al.*, 1997; Auger *et al.*, 2000; Keel, 2004), it seems clear that the minimum (<14%) and desirable (<28%) performance based on biological variation is too strict. Quality specifications based on biological variation should be considered a target, not inflexible criteria of acceptability (Fraser, 1988; Kenny *et al.*, 1999). Quality specifications based on biological variation may not be adequate when we analyse pathologic samples, as the reproducibility of seminal parameters is different in asthenozoospermic and normozoospermic individuals using computer-assisted semen analysis (McKinney and Thompson, 1994).

On the other hand, the optimum performance determined by state of the art criteria for morphology seems too high. If we accepted this level of quality, it would mean that for a male with 14% of normal forms, values of 5–23% could be obtained, if stained smears were used. As the specifications of analytical quality based on the state of the art depend on the performance of laboratories, the former should be periodically checked, although large variations are not to be expected, as various EQAS contributions were used for the study.

There remain to be established definitive values for quality specifications determined by medical necessities. From the quality specifications based on biological variation, we can conclude that the methodology and technology used in the laboratories participating in Spanish EQAS to determine concentration, total and progressive motility is adequate, because a high proportion of laboratories achieved the minimum quality specification for total allowable error based on the components of biological variation. However, quality specifications based on biology are not met by a high number

of laboratories concerning morphology and progressive rapid motility. This may be due to the fact that the methodology or the standardization needs to be improved. Among theoretical solutions could be the use of computer-assisted semen analysers, an improvement in the standardization of criteria and the evaluation of participation results by means of external quality control. The first possibility has not been found to be of great utility in reducing inter-laboratory variability (Keel *et al.*, 2000). However, the second (Franken *et al.*, 2000; Björndahl *et al.*, 2002; Franken *et al.*, 2003) and the third (Cooper *et al.*, 1999) options have been shown to reduce differences between andrology laboratories.

The large difference found between quality specifications based on the state of the art for total and progressive motility versus rapid progressive motility suggests that the assessment of WHO grades a and b motile spermatozoa is a major source of errors, as shown previously (Dunphy *et al.*, 1989).

In summary, the present study enabled us to identify the state of the art of analytical performance for seminal parameters, and to reveal the difficulty inherent in meeting the specifications for quality based on biological variation and in establishing analytical goals based on the state of the art.

References

Álvarez C, Castilla JA, Molina R, Ramírez JP, Vergara F, Yoldi A and Gaforio JJ (2003) Biological variation in seminal parameters in healthy subjects. Hum Reprod 18,2082–2088.

Auger J, Eustache F, Ducot B, Blandin T, Daudin M, Diaz I, El Matribi S, Gony B, Keskes L, Kolbezen M et al. (2000) Intra- and inter-individuality variability in human sperm concentration, motility and vitality assessment during a workshop involving ten laboratories. Hum Reprod 15, 2360–2368.

Björndahl L, Barratt CL, Fraser LR, Kvist U and Mortimer D (2002) ESHRE basic semen analysis courses 1995–1999: immediate beneficial effects of standardized training. Hum Reprod 17,1299–1305.

Calafell R, Domingo J, Fernandez C, Goya R, Guardiola J, Martinez S, Morancho J, Prats R and Salve ML (2002) Especificaciones de calidad analitica. An Clin 27,95–100.

Cooper TG, Atkinson AD and Nieschlag E (1999) Experience with external quality control in spermatology. Hum Reprod 14,765–769.

Cooper TG, Björndahl L, Vreeburg J and Nieschlag E (2002) Semen analysis and external quality control schemes for semen analysis need global standardization. Int J Androl 25,306–311.

Cotlove E, Harris EK and Williams GZ (1970) Biological and analytic components of variation in long term studies of serum constituents in normal subjects: III. Physiological and medical implications. Clin Chem 16, 1028–1032.

Dunphy BC, Kay R, Barratt CL and Cooke ID (1989) Quality control during the conventional analysis of semen, an essential exercise. J Androl 10,378–385.

Franken DR, Smith M, Menkveld R, Kruger TF, Sekadde-Kigondu C, Mbizvo M and Akande EO (2000) The development of a continuous quality control (CQS) program for strict sperm morphology among sub-Sahara African laboratories. Hum Reprod 5,667–671.

Franken DR, Menkveld R, Kruger TF, Sekadde-Kigondu C and Lombard C (2003) Monitoring technologist reading skills in a sperm morphology quality control program. Fertil Steril 79 (Suppl 3),1637–1643.

Fraser CG (1988) Analytic goals are targets, not inflexible criteria of acceptability. Am J Cin Pathol 89,703–705.

Fraser CG (2001) Biological Variation: From Principles to Practice. 1st edn. AACC Press, Washington, DC, USA.

Fraser CG, Petersen PH, Libeer JC and Ricós C (1997) Proposals for setting generally applicable quality goals solely based on biology. Ann Clin Biochem 34,8–12.

Fraser CG, Kallner A, Kenny D and Petersen PH (1999) Introduction: strategies to set global quality specifications in laboratory medicine. Scand J Clin Lab Invest 59,477–478.

Jorgensen N, Auger J, Giwercman A, Irvine DS, Jensen TK, Jouannet P, Keiding N, Le Bon C, MacDonald E, Pekuri AM et al. (1997) Semen analysis performed by different laboratory teams: an intervariation study. Int J Androl 20,201–208.

Keel BA (2004) How reliable are results from the semen analysis? Fertil Steril 82,41–44.

Keel BA, Quinn P, Schmidt CF Jr, Serafy NT Jr, Serafy NT and Schalue TK (2000) Results of the American Association of Bioanalysts national proficiency testing programme in andrology. Hum Reprod 15,680–686.

Kenny D, Fraser CG, Petersen PH and Kallner A (1999) Consensus agreement. Scand J Clin Lab Invest 59,585.

Kvist U and Björndahl L (eds) (2002) ESHRE Monographs: Manual on Basic Semen Analysis. Oxford University Press, Oxford, UK.

Lott JA (1999) Setting process control limits for enzyme tests in serum. Scan J Clin Lab Invest 59,527–530.

McCulloh DH (2004) Quality control: maintaining stability in the laboratory. In Gardner DK, Weissman A, Howles CM, and Shoham Z (eds) Texbook of Assisted Reproductive Techniques. Taylor & Francis, London, UK, pp. 25–39.

McKinney KA and Thompson W (1994) Reproducibility of sperm motility measurements in asthenozoospermic and normozoospermic individuals using the Hamilton–Thorn motility analyzer. Arch Androl 33,151–155.

Neuwinger J, Behre HM and Nieschlag E (1990) External quality control in the andrology laboratory: an experimental multicenter trial. Fertil Steril 54,308–314.

Petersen PH, Fraser CG and Kenny D (1999) Strategies to set global analytical quality specifications in laboratory medicine. Scand J Clin Lab Invest 59,1–164.

Petersen PH, Brandslund I, Jorgensen L, Stahl M, Defineolivarius N, Borch K and Ohnsen J (2001) Evaluation of systematic and random factors in measurements of fasting plasma glucose as the basis for analytical quality specifications in the diagnosis of diabetes. 3. Impact of the new WHO and ADA recommendations on diagnosis of diabetes mellitus. Scand J Clin Lab Invest 61,191–204.

Ricós C, Álvarez V, Cava F, Garcia-Lario JV, Hernández A, Jiménez CV, Minchinela J, Perich C and Simon M (1999) Current database on biologic variation: pros, cons and progress. Scand J Clin Lab Invest 59,491–500.

Theinpoint LM, Steyaert HL and De Leenherr AP (1987) A modified statistical approach for the detection of outlying values in external quality control: comparison with other techniques. Clin Chem Acta 168,337–346.

Vassault A, Grafmeyer D, de Graeve J, Cohen R, Beaudonnet A and Bienvenu J (1999) Analyses de biologie médicale: spécification et normes d'acceptabilité à l'usage de la validation de techniques. Ann Biol Clin 57,685–695.

World Health Organization (1999) WHO Laboratory Manual for the Examination of Human Semen and Semen-Cervical Mucus Interaction. 4th edn. Cambridge University Press, Cambridge, UK.

Submitted on December 17, 2004; resubmitted on April 4, 2005; accepted on April 7, 2005

QUALITY, INNOVATION, AND SERVICE—IT'S ALL AT THE CENTER OF EVERYTHING WE DO.

From developing assisted reproductive technologies that maximize performance, like the first ART media and cultures, to expertise that streamlines productivity, FUJIFILM Irvine Scientific brings together decades of industry expertise with a powerhouse of innovation, turning opportunities into realities. Together, we're working to support healthy futures—from retrieval to realization.

ALL IN *FOR LIFE.*
irvinesci.com/*ALLIN*

FUJIFILM
Value from Innovation

IrvineScientific