RBMO

ARTICLE

A technical note on the assessment of human sperm vitality using eosin–nigrosin staining

BIOGRAPHY

David Mortimer is co-owner of Oozoa Biomedical, a Vancouver-based reproductive biomedicine consulting company that designs ART laboratories, upgrades laboratory systems, and performs Total Quality Management audits. He is author of over 150 publications including Practical Laboratory Andrology, Quality and Risk Management in the IVF Laboratory, and A Practical Guide to Basic Laboratory Andrology.

David Mortimer*

KEY MESSAGE

Counting 200 spermatozoa from one eosin–nigrosin stained slide provides a result with sufficient accuracy for the clinical purpose of ascertaining whether the immotile spermatozoa in a human ejaculate showing low sperm motility are alive or dead.

ABSTRACT

Research question: How many spermatozoa and slides need to be counted to make a reliable assessment of sperm vitality? Currently, various authorities recommend assessing human sperm vitality on counts of at least 200 cells, but on one or two slides.

Design: This was an observational study on duplicate eosin–nigrosin stained sperm vitality slides made from 58 ejaculates. Assessments were made using counts of 2 × 100 and 1 × 200 cells per slide, all performed by the same trained expert observer.

Results: Although assessments tend to show fewer and smaller outlier values when based on counts of 200 spermatozoa than 100, and on 2 × 200 than 1 × 200, counting 200 spermatozoa from one slide provides a result with sufficient accuracy for the clinical purpose of ascertaining whether the immotile spermatozoa in an ejaculate showing low sperm motility are alive or dead.

Conclusions: While the increased accuracy of results derived from counts of 2 × 200 cells might be important in research studies where sperm vitality is the specific end-point of interest, counting at least 200 spermatozoa from one smear is sufficiently accurate for the clinical purpose of establishing whether the immotile spermatozoa seen in ejaculates with low sperm motility are alive or not. Consequently, the extra workload of performing replicate counts and the associated calculations does not increase the clinical value of the result, and hence is unnecessary in routine semen analysis. Careful laboratory technique as well as proper staff training and competence are essential. The conclusions might not be applicable to staining methods other than the recommended one-step eosin–nigrosin technique.

Oozoa Biomedical, Inc., Box 93012 Caulfeild Village, West Vancouver BC V7W 3G4, Canada

*Corresponding author. E-mail address: david@oozoa.com (D Mortimer). https://doi.org/10.1016/j.rbmo.2020.03.002 1472-6483/
Declaration: The authors report no financial or commercial conflicts of interest.

KEYWORDS
Eosin–nigrosin
Semen analysis
Sperm vitality

INTRODUCTION

The assessment of human sperm vitality is an integral part of a basic semen analysis (*Björndahl et al., 2004*, *2010*; *Mortimer, 1994*), particularly in ejaculates with low motility, usually defined as <40% motile (*Barratt et al., 2011*; *Björndahl et al., 2010*). The purpose of the assessment is to ascertain whether the immotile spermatozoa in an ejaculate showing a low proportion of motile cells are alive or dead. The former situation can arise owing to flagellar structural abnormalities, the extreme case being Kartagener ('immotile cilia') syndrome (for review, see *Mortimer, 2018*), while the latter might be caused by cytotoxic anti-sperm antibodies or some chemical toxicity (*Björndahl et al., 2010*; *Mortimer, 1994*; *Mortimer et al., 2013*).

The most widely recommended method is the combined eosin–nigrosin ('E–N') stain, where eosin is a supravital stain and the purple nigrosin serves as a counterstain so that the unstained live spermatozoa can be seen under bright field light microscopy (*Björndahl et al., 2003*, *2004*, *2010*; *Mortimer, 1994*).

However, there is controversy as to how many slides, and how many spermatozoa per slide, need to be counted to obtain an accurate result. The current 5th edition of the World Health Organization's semen analysis manual ('WHO5'; *World Health Organization, 2010*) says to assess 200 spermatozoa in each of two slides (section 2.6.1.2, #8), whereas the textbook for the ESHRE Special Interest Group in Andrology's Basic Semen Analysis courses (*Björndahl et al., 2010*), and technical recommendations for semen analysis (*Björndahl et al., 2016*), say to assess at least 200 spermatozoa, but do not require the use of replicates.

Given the specific purpose of assessing sperm vitality within a routine semen analysis, a highly precise result is not necessary, begging the question as to whether the extra effort of counting at least 200 spermatozoa in duplicate slides is of value. The purpose of this technical note is to investigate the measurement error of human sperm vitality assessments on the basis of counting 100 or 200 spermatozoa in one or two slides.

MATERIALS AND METHODS

Replicate E–N smears were prepared from 60 ejaculates provided for analysis at a specialized diagnostic andrology laboratory. This number of samples was chosen because according to the central limit theorem, the 95% range is 2.0 × SD when $n = 60$. However, the motility assessments on two ejaculates were found to have been made outside the standard time window and these samples were excluded from the analysis. The E–N method used was exactly as described in *Björndahl et al. (2010)*, and performed in complete accordance with the *Björndahl et al. (2016)* 'How to count sperm properly' guidelines.

The sperm motility in these 58 ejaculates was 62 ± 18% (mean ± SD), ranging between 23% and 86%, covering a wide clinical range of samples. Because the end-point of the study was simply the accuracy of reading a series of sperm vitality slides, neither details of the clinical provenance of the samples nor their general semen characteristics were pertinent to the study's performance or findings.

Liquefied semen was mixed thoroughly without the use of a vortex mixer and a 30 µl aliquot mixed with an equal volume of a commercial one-step E–N stain (Sperm VitalStain, Nidacon International AB, Göteborg, Sweden) either on a porcelain spotting plate or in a 1.5 ml Eppendorf tube. After 30 s, 40 µl of the semen + stain mixture was transferred onto one of a pair of clean, labelled 3″ × 1″ glass microscope slides as a 'streak' down the middle of the slide. The second slide was then placed face down on top of the first slide and the semen + stain mixture allowed to spread between the two slides, after which the slides were smoothly pulled apart horizontally to make a pair of smears. Care was taken to make the smears thin enough so that the counterstain was not too dark for clear visualization of the spermatozoa. After allowing the smears to air dry, they were mounted as soon as practicable using 50 × 22 mm #1 coverslips and a permanent mountant and allowed to dry at least overnight. These slides can be stored indefinitely at ambient temperature.

For each slide, two counts of at least 100, and one count of at least 200 spermatozoa were performed at a magnification of 1000 × under oil immersion using a 100 × bright field (i.e. not phase contrast) objective and correctly adjusted Köhler illumination. In each count the entire last field was always assessed in its entirety to remove potential counting bias, so the 100-cell counts ranged from 100 to 120 spermatozoa and those of 200 cells ranged from 200 to 232 spermatozoa. All counts were performed by the same trained observer (see below), and the proportion of live spermatozoa was expressed as an integer percentage value.

Care was taken not to confuse any nigrosin stain underlying the anterior portion of the sperm head for pink eosin staining of the cell. White (unstained) spermatozoa were counted as 'live' and those showing pink or red colouration as 'dead'. Only those that showed a faint pink colouration in the neck region alone ('leaky necks') were considered as not dead, as per the validated method (*Björndahl et al., 2003*, *2004*, *2010*); the method described in WHO5 is incorrect in this regard (*Barratt et al., 2011*). This staining artefact arises from a small amount of eosin entering the neck region of some spermatozoa as a result of shrinkage during drying of the smears that causes a breach of the plasma membrane behind the posterior ring structure that separates the sperm head and neck regions (*Mortimer, 2018*).

The observer was trained using reference slides and the iterative method first described by *Mortimer (1994)*, and since employed in the ESHRE Special Interest Group in Andrology's Basic Semen Analysis courses (*Björndahl et al., 2010*). Competence was defined as achieving a mean difference of less than ±0.5% live spermatozoa, as well as a 95% range of discrepancies of less than or equal to ±5%, compared with the reference values over a series of at least 20 slides. The laboratory participated in the ESHRE External Quality Control Programme.

Data were analysed and *Bland and Altman (1986)* plots created using MedCalc v19.1 software (MedCalc, Ostend, Belgium; www.medcalc.org) and graphs combined into a single figure using CorelDraw 2019 (Corel Corporation, Ottawa, Canada; www.coreldraw.com). Variances of the differences between assessment series were compared using the F-test (MedCalc). In simple Bland and Altman

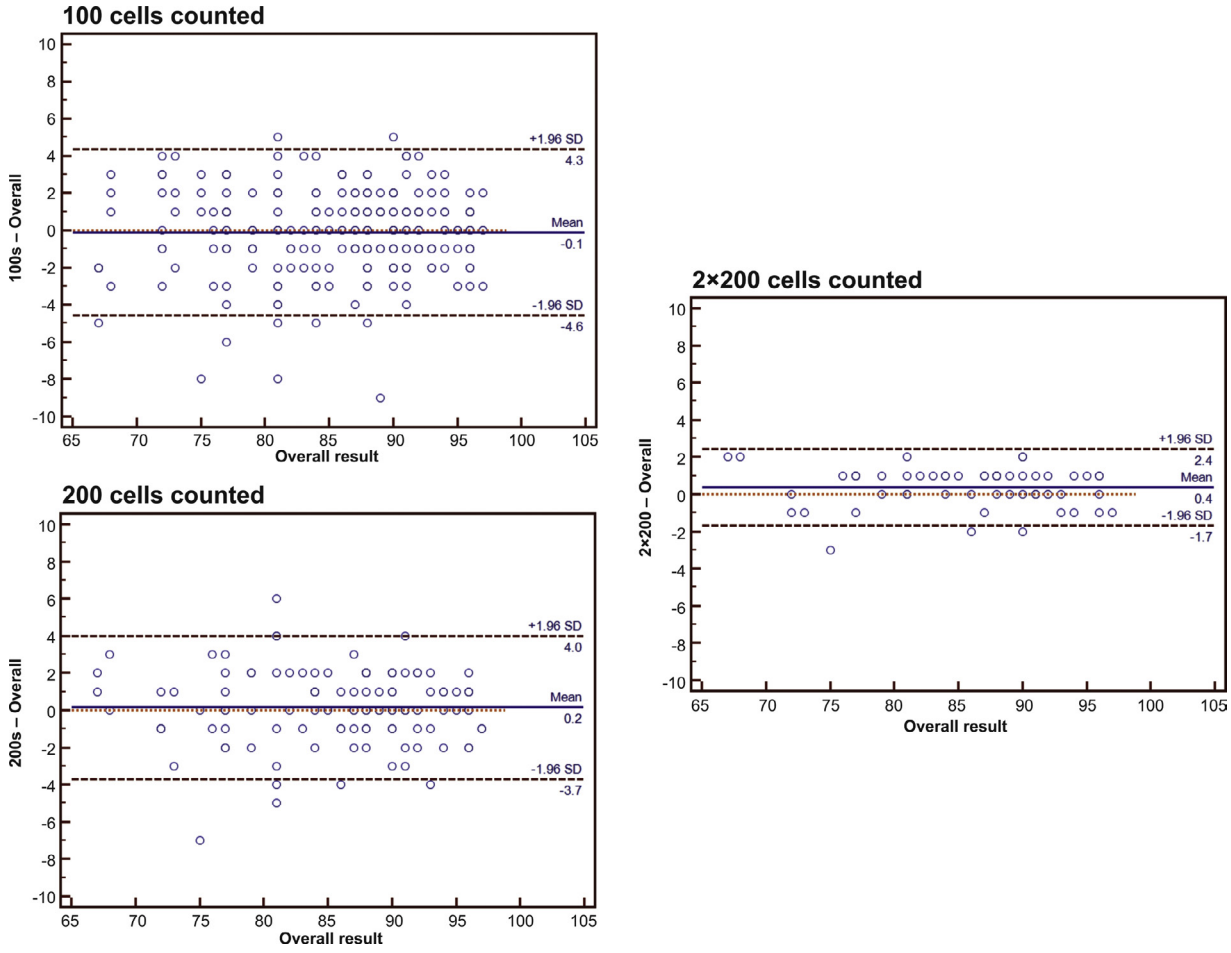

FIGURE 1 Bland and Altman plots for the assessment of human sperm vitality in 58 pairs of eosin–nigrosin stained smears based on counts of 100, 200 and 2 × 200 spermatozoa by a trained observer.

plots the difference between two methods of measurement is plotted against the average of the methods when both methods are expected to have equivalent likelihood of providing a correct result. In the present study the overall result was used instead of the average as it was expected to be the most accurate value, being based on the greatest number of assessments.

RESULTS

Among the 58 ejaculates used in the study the sperm vitality ranged between 67% and 97% (86 ± 7%, mean ± SD); these overall results were calculated using the actual count data from all six readings performed on each specimen (i.e. ≥800 spermatozoa assessed).

FIGURE 1 comprises Bland and Altman plots for the three separate assessments, using the overall result as the reference value because it represents the most accurate assessment. In all three series the 95% range of differences between the individual and overall results was <5%, but the ranges decreased from the 100 cells counted series to the 200 cells counted to the 2 × 200 cells counted series. The differences data (calculated as individual result – overall result, so a negative value denotes the individual count was lower than the overall value) are summarized in TABLE 1. The mean differences were all essentially zero, with SD of ≤2.5%, and hence calculated 95% ranges of no more than ±5%. The variances of the 100 and 200 cells counted series were not different ($F = 1.3278$, $P = 0.088$), but the variance of the 2 × 200 cells counted was lower than either single count series ($F = 3.5201$, $P < 0.001$ and $F = 4.6738$, $P < 0.001$ for the 100 and 200 cells counted series, respectively). There were occasional outliers, but all <10% (±10% being the target for measurement error in semen analysis; see *Björndahl et al., 2010*). The higher outlier values were seen among assessments based on 100 spermatozoa (–8 × 2, –9 × 1) than on 200 spermatozoa (maximum = –7), indicating better precision of the latter.

When duplicate counts were performed on 200 spermatozoa per slide the results were closer to the overall value compared with the individual counts of 200 spermatozoa, but that is entirely to be expected because they were derived using 50% of the cells counted to determine the overall value (≥800 cells).

DISCUSSION

The purpose of assessing sperm vitality within a routine semen analysis is to elucidate whether the immotile spermatozoa seen in ejaculates with low sperm motility are alive or dead. As expected, assessments tend to be more reliable, in terms of fewer and smaller outlier values, when based on counts of 200 spermatozoa than 100, and on 2 × 200 than 1 × 200. However, the effort involved is obviously doubled when

TABLE 1 DIFFERENCE VALUES (Δ%) BETWEEN THE INDIVIDUAL COUNT RESULTS AND THE OVERALL VALUE FOR THE PERCENTAGE OF LIVE SPERMATOZOA SEEN IN 58 SEMEN SAMPLES

Slides	Readings	No. of results	Mean ± SD	Min and max	Outliers beyond ±5%
			Δ%	Δ%	
A	1st 100	58	−0.2 ± 2.5	−6 to +5	−6
	2nd 100	58	−0.5 ± 2.0	−8 to +5	−8
	200	58	0.2 ± 2.0	−7 to +6	−7, +6
B	1st 100	58	0.2 ± 2.5	−9 to +4	−9, −8
	2nd 100	58	0.2 ± 2.0	−4 to +4	
	200	58	0.1 ± 2.0	−5 to +4	
A and B	All 100s	232	−0.1 ± 2.3	−9 to +5	−9, −8 × 2, −6
A and B	All 200s	116	0.2 ± 2.0	−7 to +6	−7, +6
A and B	2 × 200	58	0.4 ± 1.1	−3 to +2	

The 'Readings' column lists the results for the first and second counts of 100 spermatozoa, and the subsequent count of 200 spermatozoa, for each of the 58 A slides and the 58 B slides; the 'A and B' rows show the differences for all 232 counts of 100 spermatozoa, all 116 counts of 200 spermatozoa, and the 2 × 200 counts for the 58 pairs of slides.

counts are on at least 200 spermatozoa than 100 spermatozoa, but more than quadrupled when counting 2 × 200 cells and performing the replicate comparison calculations.

The Bland and Altman plots in the present study used the overall result rather than the average of the two values being compared. While this approach has been criticized as often showing a relationship between the measurements whether there is a true association between difference and result magnitude or not (*Bland and Altman, 1995*), this is not a concern here.

FIGURE 1 illustrates that counting 200 cells reduced both the measurement error of the assessments and the prevalence of cases with outliers of >5%, yielding results that are certainly fit-for-purpose in clinical semen analysis terms. Counting 2 × 200 cells significantly reduced assessment variability, but while the increased accuracy of results derived from duplicate counts might be important in research studies where sperm vitality is the specific end-point of interest, the extra effort would not provide any greater clinical value, and hence is unnecessary in routine semen analysis.

The use of a one-step E–N stain permits reliable assessment of sperm vitality and is superior to the use of eosin alone, as was recommended by the World Health Organization at that time (*Björndahl et al., 2003*, *2004*). Without the nigrosin counterstain, reliable evaluation of unstained live spermatozoa is rendered much more difficult unless negative phase contrast optics are used, and these are very difficult to obtain from modern microscope manufacturers, and rarely seen in andrology laboratories except where they are integral to certain CASA instruments. The commercial E–N stain used in the study (Sperm VitalStain, Nidacon International AB) is directly equivalent to the in-house prepared one-step E–N stain (*Björndahl et al., 2003*, *2010*).

This study has shown that the routine method of counting at least 200 spermatozoa from one smear (*Björndahl et al., 2010*, *2016*) gives sufficiently accurate results for the intended clinical purpose of establishing whether the immotile spermatozoa seen in ejaculates with low sperm motility are generally alive or not. This method should therefore continue to be employed in the ESHRE Special Interest Group in Andrology's Basic Semen Analysis courses. However, it must be stressed that besides careful laboratory technique (*Björndahl et al., 2003*, *2004*, *2010*), effective training and establishment of competence for all staff performing sperm vitality assessments are essential, and in accordance with clauses 5.1.5b and 5.1.6 of ISO 15189:2012 *Medical laboratories – Requirements for quality and competence* (*International Organization for Standardization, 2012*).

REFERENCES

Barratt, C.L.R., Björndahl, L., Menkveld, R., Mortimer, D. **The ESHRE Special Interest Group for Andrology Basic Semen Analysis Course: A continued focus on accuracy, quality, efficiency and clinical relevance.** Hum. Reprod. 2011; 26: 3207–3212

Björndahl, L., Mortimer, D., Barratt, C.L.R., Castilla, J.A., Menkveld, R., Kvist, U., Alvarez, J.G., Haugen, T.B. 2010 **A Practical Guide to Basic Laboratory Andrology.** Cambridge University PressCambridge, UK

Björndahl, L., Söderlund, I., Kvist, U. **Evaluation of the one-step eosin-nigrosin staining technique for human sperm vitality assessment.** Hum. Reprod. 2003; 18: 813–816

Björndahl, L., Söderlund, I., Johansson, S., Mohammadieh, M., Pourian, M.R., Kvist, U. **Why the WHO recommendations for eosin-nigrosin staining techniques for human sperm vitality assessment must change.** J. Androl. 2004; 25: 671–678

Björndahl, L., Barratt, C.L.R., Mortimer, D., Jouannet, P. **'How to count sperm properly': Checklist for acceptability of studies based on human semen analysis.** Hum. Reprod. 2016; 31: 227–232

Bland, J.M., Altman, D.G. **Statistical methods for assessing agreement between two methods of clinical measurement.** Lancet 1986; 327: 307–310

Bland, J.M., Altman, D.G. **Comparing methods of measurement: why plotting difference against standard method is misleading.** Lancet 1995; 346: 1085–1087

International Organization for Standardization. 2012 **ISO 15189:2012 Medical laboratories – Requirements for quality and competence.** International Organization for StandardizationGeneva, Switzerland

Mortimer, D. 1994 **Practical Laboratory Andrology.** Oxford University Press New York, USA

Mortimer, D. **The functional anatomy of the human spermatozoon: relating ultrastructure and function.** Mol. Hum. Reprod. 2018; 24: 567–592

Mortimer, D., Barratt, C.L.R., Björndahl, L., de Jager, C., Jequier, A.M., Muller, C.M. **What should it take to describe a substance or product as 'sperm-safe'.** Hum. Reprod. 2013: i1–i45

World Health Organization. 2010 **WHO Laboratory Manual for the Examination and Processing of Human Semen, fifth ed.** World Health Organization Geneva, Switzerland

Received 29 January 2020; received in revised form 26 February 2020; accepted 3 March 2020.