**Original Article**

**Myths of Sperm Morphology- Inter-observer Variance in Human Sperm Morphology Assessment**

Uma Maheshwari A*, Dhanashri D**, Ranjani S***, Radha P****, Pandiyan N*****  
*Clinical Embryologist, SB Fertility Centre, Thanjavur, **Clinical Embryologist, ***Senior Clinical Embryologist, ****Senior Consultant, *****Chief Consultant & HOD, Dept of Andrology & Reproductive Medicine, Chettinad Super Speciality Hospital, Chettinad Academy Of Research & Education, Chennai, India.

Dr. Uma Maheswari did her Bachelor Of Dental Surgery from SRM Dental College and did her PGDiploma in Clinical Embryology at Chettinad Hospital & Research Institute. She is currently working as Clinical Embryologist in SB Fertility Centre, Thanjavur.

**Abstract**

Male infertility has gained focus in the last few decades and in the pursuit of treating the male, more in depth investigations have emerged into the semen analysis and sperm function. Though some authors described that sperm morphology is also a good indicator for male fertility, numerous publications have highlighted its analytical weakness with wide inter and intra observer variations.

**Aim:** The following study was aimed to identify the inter-observer variations of human sperm morphology assessment according to WHO guidelines 2010.

**Materials and Methods:** This was a prospective blinded study conducted in the Department of Reproductive Medicine, Chettinad Super Speciality Hospital, Kelambakkam, and Tamilnadu, India. Semen samples from men who reported for routine semen analysis as a part of their infertility investigations were recruited in this study. A total of 1100 spermatozoa were assessed according to WHO 2010 criteria by five senior embryologists simultaneously. The feedback forms were analyzed and compiled for statistical significance.

**Result:** From this study we observed that there was an interobserver variance between the five embryologists in assessing the head and tail defect with P value=0.000 which is statistically significant.

**Conclusion:** This study also shows much subjective variations for morphology assessment of spermatozoa. More details about abnormal forms can lead to more anxiety for both the patients and the clinicians. This is our concern. It is time to revisit the necessity of the laborious process of training for morphology assessment or should semen analysis be simplified to binary assessments with only concentration and motility as its components.

**Key Words:** Sperm morphology, Semen analysis, Variability.

**Introduction**

Since the discovery of spermatozoa by Leeuwenhoek in 1677, routine semen analysis and studies of spermatozoa in detail have become integral parts of infertility investigations. Male infertility has gained focus since the last few decades and in the pursuit of treating the male, more profound insights into the semen analysis and sperm function tests have emerged as essential tools in the armamentarium of male infertility management. Wagner first explored the relationship between sperm morphology and development of spermatozoa in 1837. 1 MacLeod and Gold were the first to report the morphological assessment of human spermatozoa. As the morphological modifications that take place during spermiogenesis are not homogeneous in human and are more physiological, the fundamental questionaries as to “what is a normal spermatozoon?” 2,3 WHO has been reducing the reference values for morphology by stricter criteria in fourth & fifth editions (1999 & 2010).

This brings to light that only 4% spermatozoa with normal morphology (5th percentile) are needed for a possible biological pregnancy.3 While it is accepted that there exists a marked difference in interpretation which is related to different observers and their levels of training, it is also evident that preparation of the slide and staining techniques can affect the morphology.4 Besides, it is clear that lack of quality control & changes in international standards in spermatozoa morphological interpretation also influence the reports. Though some authors described that sperm morphology is also a good indicator for male fertility, numerous publications have highlighted its analytical weakness with wide inter and intra observer variations.5-8

With this background, we strived to study the possible weakness of this step in semen analysis.
Aim
The following study was aimed to identify the inter-observer variations of human sperm morphology assessment according to WHO guidelines 2010.

Materials and Methods
This was a prospective blinded study conducted in the Department of Reproductive Medicine, Chettinad Super Speciality Hospital, Kelambakkam, and Taminadu, India. Semen samples from men who reported for routine semen analysis as a part of infertility evaluation were recruited for the study after obtaining a written informed consent. Semen samples with Azoospermia were excluded from the study.

Preparing a Stained Glass Slide With The Semen Sample
After preliminary semen analysis, an aliquot (10 µl) of undiluted semen was placed on a glass slide and was spread by pulling forward with a slide angled (45°, slide-feathering method) as described1. The air dried smear was stained by differential quick staining slide-feathering method) as described.

Strict Kruger’s Criteria
Head
Oval configuration with a smooth contour Length is 5.6 µm, Width is 3.5-5 µm And the width/length ratio is 1/2-3/5.

Acrosome
Comprising 40-70% of the distal part of the head

Mid piece
No cytoplasmic droplets of more than half of the sperm head are accepted

Tail
Short, multiple, broken, smooth hairpin bends, sharply Angulated bends, of irregular width, coiled, or any combination of these.

Photographs of the numbered spermatozoa were projected onto a screen for morphology assessment. All the five embryologists simultaneously and individually assessed morphology and recorded on separate forms. The feedback forms were analyzed and compiled for statistical significance.

Results

<table>
<thead>
<tr>
<th>TOTAL NUMBER OF OBSERVERS</th>
<th>NORMAL SPERM</th>
<th>HEAD DEFECT SPERM</th>
<th>MIDPIECE DEFECT SPERM</th>
<th>TAIL DEFECT SPERM</th>
<th>MULTIPLE DEFECT SPERM</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n(%)</td>
<td>n(%)</td>
<td>n(%)</td>
<td>n(%)</td>
<td>n(%)</td>
<td>n(%)</td>
</tr>
<tr>
<td>1</td>
<td>142(100)</td>
<td>102(71.7)</td>
<td>19(13.4)</td>
<td>139(97.9)</td>
<td>530(34.9)</td>
<td>103</td>
</tr>
<tr>
<td>2</td>
<td>57(40.0)</td>
<td>347(68.3)</td>
<td>19(3.3)</td>
<td>134(46.4)</td>
<td>398(68.6)</td>
<td>57</td>
</tr>
<tr>
<td>3</td>
<td>29(1.7)</td>
<td>188(63.8)</td>
<td>1(0.6)</td>
<td>96(32.8)</td>
<td>179(62.4)</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>5(0.3)</td>
<td>149(24.7)</td>
<td>0(0.0)</td>
<td>69(12.4)</td>
<td>75(12.6)</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>5(0.4)</td>
<td>149(24.7)</td>
<td>0(0.0)</td>
<td>55(9.5)</td>
<td>49(8.6)</td>
<td>5</td>
</tr>
<tr>
<td>TOTAL</td>
<td>233(100.0)</td>
<td>1009(43.0)</td>
<td>21(0.0)</td>
<td>377(16.4)</td>
<td>1945(82.6)</td>
<td>233</td>
</tr>
</tbody>
</table>

Table 1. Statistical Analysis of Observers Morphological Assessment.

Statistically significant if P<0.05 P value=0.000; Significant.

The values in Table 1. shows
- All five embryologists scored for spermatozoa at the same time and gave the assessment of 5 normal forms, 343 head defect forms and 55 multiple defect forms.
- Four among the five observers gave the same assessment of 5 normal forms, 249 head defect forms and 69 multiple defect forms.
- Three among the five observers gave the same assessment of 29 normal forms, 188 head defect forms, 5 mid-piece defect forms, 1 tail defect form and 69 multiple defect forms.
- Two among the five observers gave the same assessment of 52 normal forms, 127 head defect forms, 19 mid-piece defect forms, 1 tail defect form and 139 multiple defect forms.
- Each observer differed from all the other observers in the assessment of of 142 normal forms, 102 head defect forms, 79 mid-piece defect forms, 19 tail defect forms and 218 multiple defect forms.

Table 2 - Coefficient of variation
The values in Table 2 show:
- Tail Defect Sperm C.V% = 775.67 being the highest.
- Midpiece defect sperm C.V% = 341.13.
- Normal sperm C.V% = 229.93.
- Multiple defect sperm C.V% = 124.19.
- Head defect sperm C.V% = 49% being the least from other groups.

The present data confirms the wide variability in the morphological assessment of:
- Tail defect sperm (C.V% - 775.67%) followed by
- Midpiece defect sperm (C.V% - 341.13%)
- Normal sperm (C.V% - 229.93%)
- Multiple defect sperm (C.V% - 124.19%)
- Head defect sperm (C.V% - 49.00%).

The bar diagram (Fig. 2) represents the concurrence of assessment in a group of all five observers, four observers, three observers, two observers and variability in assessment.

Morphological Assessment

Same morphological assessment done by all five observers for five spermatozoa (Fig. 3 to Fig. 7) out of 1100 sperms as normal forms.

Morphological assessment done by four among the five observers for five sperms (fig. 8-fig. 12) out of 1100 sperms as normal forms.
Morphological assessment by three among the five observers for 29 sperms (fig.13-fig.41) out of 1100 sperms as normal forms.
Discussion

Based on the criteria recommended for morphological assessment in semen analysis, there is a significant variability among the observers in the morphological assessment of each spermatozoon. It implies that inter-individual laboratory variations do exist. From our observation with five senior embryologists who had rich experience in semen analysis there was much variation in identifying even the percentage of normal spermatozoa. Compared to other defects high variance was identified in assessing the tail defects though there was correlation in identifying the head defects.

WHO has revised its criteria for morphological assessment thrice since 1987. Initially it was based on Tygerbergs criteria in 1987 and Kruger’s criteria in 1992. Later in 1999 it was changed into strict Kruger’s criteria which was based on consensus arrived by the articles included from 14 countries in four continents and four study population.

Apart from WHO classification there are other classifications that are followed to assess morphology of the spermatozoa. One among that is Davids classification in which the Multiple Anomalies Index (MAI) was calculated (ratio between the total number of abnormalities and the number of abnormal spermatozoa) and similar result has been noticed in a study conducted in France by Eustache and Auger. It is highly confusing to have so many classifications and criteria to assess the morphology of spermatozoa. It is not a surprise that this aspect of semen analysis is so nebulous and is the subject of debate and confusion. Other authors such as Matson et al., Ombelet et al., Keel et al have also reported strong inter observer variability in their studies.

Reports on morphology of human spermatozoa can also be intimidating to the clinician and couples. Unless it is a case of total Teratozoospermia (e.g.: Globozoospermia), is it necessary to delve into more details about normal and abnormal forms? A recent study conducted by Kovac et al showed that even men with complete Teratozoospermia had higher rates of spontaneous pregnancy without assisted reproduction. Hence categorizing the type of treatment for morphological abnormalities of spermatozoa either with IVF/ICSI is also questionable.

Conclusion

Due to lack of standardization of sperm morphology assessment for the past 30 years, Gatimel et al in his recent review paper questioned the use of sperm morphology assessment in routine semen analysis until significance to clinical outcome of a given treatment is quantified. Our study also shows much subjective variations for morphology assessment of spermatozoa. More details about abnormal forms can lead to more anxiety to both the patients and the clinicians. This is our concern. It is time to revisit the necessity of the laborious process of training for morphology assessment and the unnecessary time taken during routine semen analysis. Should sperm morphology still be included as a routine in semen analysis or should semen analysis be simplified to binary observations with only concentration and motility? We hope we will get the answers soon….

References


