The stallion breeding soundness evaluation: revisited
Charles C. Love
Section of Theriogenology, Department of Large Animal Clinical Sciences, College of Veterinary Medicine, Texas A&M University, College Station, TX

Abstract
The stallion breeding soundness evaluation (BSE) was originally formalized in 1983 in pamphlet form by the Society for Theriogenology and called the Manual for Clinical Fertility Evaluation of the Stallion. Since that time it has provided an invaluable guideline for practitioners to critically evaluate the reproductive capability of a stallion by providing an outline on how to perform and interpret those findings. The horse breeding industry has changed and new information has become available concerning aspects that affect the performance and interpretation of the BSE. This proceeding discusses several aspects of the BSE including the evaluation of sperm quality, interpretation of BSE findings and prediction of fertility.

Keywords: Stallion, sperm, motility, morphology, breeding soundness, fertility

Introduction
The guidelines for performing a stallion breeding soundness evaluation (BSE) were introduced in pamphlet form in 1983 by the Society for Theriogenology. In performing a thorough evaluation of the stallion the practitioner is able “to eliminate from consideration…or at least alert the owner of potential problems”. It is further stated “the examination will assist identifying the cause(s) of reduced fertility and the findings used to develop guidelines for management of the stallion to enable it to achieve its maximum fertility”.

The intent of the stallion BSE is different from those performed on food animals in which the goal is to identify those males that fall below a specific threshold for sperm quality and other factors and eliminate them from the population. In contrast, the goal of the stallion BSE is to identify the cause and pursue, in many cases, a resolution of the problem. This approach is often questioned with the concern that the cause of the reduced fertility/subfertility may be hereditary in nature and therefore the trait should not be propagated by breeding the stallion. While the level of reduced/subfertility in the stallion population due to hereditary causes is unknown, non-hereditary causes such as illness and advanced age are well established and likely account for the majority of inherent stallion subfertility. In addition, other causes of “stallion subfertility” often overlooked include mare and management limitations.

The Manual suggested that the results of the BSE would classify a stallion as Satisfactory, Questionable, or Unsatisfactory based on the ability of a stallion to “render pregnant” 75% of 40 mares by natural cover, or 120 mares by artificial insemination. This suggests a 50%/cycle pregnancy rate if each mare is exposed to the stallion for at least two estrous cycles. These book sizes of 40 and 120 mares were not randomly chosen but rather resulted from what at the time, were the number of shares sold at the time stallions were syndicated (40 for Thoroughbreds; 120 for Standardbreds). Therefore, these recommendations were intended to accommodate industry expectations, primarily Thoroughbred and Standardbred, and provide conservative endpoints to evaluate and render an opinion the reproductive capability of a stallion. They were not, however, based on a biological threshold for fertility, nor are the sperm quality guidelines in the manual directly related to a specific endpoint (i.e. 75% pregnancy rate in a book of 40 mares).

The BSE Manual has provided the veterinary clinician an invaluable framework to examine, diagnose, and treat reproductive problems in the stallion. Since 1983 the horse breeding industry has been dynamic, with the introduction and common application of cooled-shipped and frozen semen. In addition, the Thoroughbred industry has increased the number of mares (i.e. 100-200 mares) to which the more popular stallions are bred. These changes have created challenges that were not addressed in the BSE manual such as how to determine the number of mares a stallion can be bred to for those larger book sizes. The manual recommendations for book size, particularly for Thoroughbred stallions (40 mares),
was probably a very modest challenge to the fertility of even stallions of lesser fertility as the authors of
the Manual were probably aware. Another area of interest includes the evaluation of sperm quality and
how it relates to fertility. At the time the manual was published, evaluation of sperm quality primarily
involved fresh semen, since then however, the expanded use of cooled-shipped and frozen semen has
raised questions about the evaluation of these sperm types. The number of sperm quality assays and the
ability of these assays to describe sperm quality objectively have advanced considerably as well as the
ability of these assays to describe the relationship of sperm quality to fertility.

Based on the expanded knowledge, it would benefit the discipline to update concepts and
guidelines set forth in the BSE manual. This proceeding is not intended to be all-inclusive but rather to
provide critique and suggestion regarding several aspects of the stallion BSE.

The case against progressive motility

It has been suggested that progressive sperm motility is “essential” for fertility in the stallion as
well as the “most critical aspect of motility”, however, “progressive” motility has not been shown to be
more important than total motility or other measure of sperm quality.

Sperm motility is historically the most common assay performed to determine sperm quality,
particularly when samples are evaluated as part of routine semen collection and processing, commonly in
concert with breeding activities or the BSE. Sperm activity can be described as either total or
progressive, but for diagnostic purposes the concept of progressive motility was initially introduced and
has been adopted. Historically, sperm motility has been determined using the “subjective method” by
visual microscopic examination. While it is unclear why the concept of “progressive motility” was
adopted, it appears to convey an additional subjective quality measure that indicates sperm that are
“straight” and “fast enough”, which further suggests a “goodness” that “non-progressive” sperm do not
possess. This has translated into a more “precise/sensitive/predictive” index of fertility. While there have
been many studies that use progressive sperm motility as their endpoint to measure sperm quality, there
have been none that have validated this measure to determine if it “means” anything more than simply
total motility or other assays of sperm quality. Sperm motility was initially adopted as a measure of
sperm quality because microscopy was the only method available to evaluate sperm, however, there have
been additional assays developed that can add to the clinician’s ability to render a diagnosis. Sperm
motility has been adopted “because we can” rather than because many assays have been compared and it
was the single assay that rose to the top as the “best”.

The definition of total motility is straightforward and is simply the percentage of sperm that
display any type of motility regardless of the “quality” of that motility. In contrast, progressive motility,
while conceptually well-defined, is practically without definition. Conceptually, progressive means
“moving relatively straight across the microscopic field” or “actively moving forward”, however,
practically, these vague definitions have resulted in an assay that lacks inter-operator repeatability
resulting in confusion regarding the “quality” of a semen sample. While it may be considered an
academic discussion the inability to describe the threshold that separates a progressive from a non-
progressive sperm is a clinically critical question, because that separation is what has been used to infer
the difference between a fertile and infertile sperm. Fertile sperm are straight and fast (progressive),
while infertile sperm are not. More recently computerized systems have been introduced to provide more
“objective” results. One limitation of this assay is repeatability, or the ability of the assay (progressive
motility) to give the same measure when different individuals measure the same sample. A previous
study determined that the coefficient of variation was high for subjective assessment of progressive sperm
motility (20%) compared to total sperm motility (10%) and determined that five ejaculates were required
for an error rate of 10% for progressive motility while only one sample was needed for total motility. In
another study, the between-breeding season coefficient of correlation was higher for total than
progressive sperm motility for both subjective (r = 0.63 vs. 0.48) and computer assisted motility analysis
(r = 0.63 vs. 0.34). In addition, both total and progressive motility were able to discriminate high and low
fertility stallions while mean spermatozoa linearity (progressive movement) was not.
The lack of repeatability of progressive motility is easily demonstrated if a sperm sample is projected on an overhead screen to a group of clinicians or students and they blindly record total and progressive motility values. There will often be considerable agreement between total motility scores but far less for progressive motility.

Why is a discussion about progressive motility anything more than an academic exercise? It has been widely adopted in both lay and veterinary jargon as the sole determinant of a stallion’s reproductive capability. Instead of thoroughly evaluating a semen sample, using assays that are more repeatable and diagnostic, both the clinicians and researchers continue to refer to this poorly defined assay to describe sperm quality. In addition, a stallion’s present or future fertility “reputation” is often based on the results of this assay.

The advent of computerized motility analyzers would seem to have solved the problem of assay repeatability, since “progressive” can be objectively defined by machine settings. Unfortunately, having a repeatable assay is of little use if the assay is not interpretable or does not relate to a relevant biological endpoint such as fertility.

A threshold for “adequate” progressive motility has been randomly set for frozen-thawed stallion semen at 30% and sometimes (European Union) at 35%. This threshold too is arbitrary and is often the sole criterion determining whether a stallion’s ejaculate “passes” or “fails” the post-thaw motility assay. Compounding this problem is that other more relevant sperm quality measures (morphology, longevity of motility, viability) are ignored. And perhaps the single most important measure, the total number of sperm inseminated, is not controlled at all.

Why is this relevant to the clinician? A common clinical question is “what do these sperm assays mean?” Many clinicians are familiar with a stallion that had low sperm motility in his ejaculate that is “fertile” or the stallion that has excellent sperm quality yet whose fertility is “low”. Often the conclusion is that sperm quality assays means little because of the lack of correlation. This conclusion is often drawn with little knowledge of the breeding circumstances (mare and management) and the assumption is then made that these “other” factors have no influence on the reported fertility outcome.

As diagnosticians “low” progressive motility is used to diagnose a stallion with a history of inadequate fertility. In effect we are saying this stallion has pathologically abnormal sperm quality sufficient to explain the fertility reduction described in the history. The lack of “progressive” sperm motility, however, is commonly not pathologic. Non-progressive sperm motility can be a transient occurrence caused by temperature or subtle osmotic changes. In addition, there are stallions whose sperm just circle, but who are highly fertile, perhaps due to the abaxial position of the midpiece attachment on the sperm head. A closer look at the sperm quality of these stallions will often find that they have a high percentage of morphologically normal sperm, high viability and DNA quality, and that their longevity of total sperm motility does not change after cooled storage, characteristics that are consistent with a fertile stallion. Longevity of sperm motility in the cooled state is a useful assay to determine whether a sperm sample with “low progressive motility” is pathologic. If the total motility is maintained for 24 hours, it is less likely that the “low” progressive motility is a fertility limiting factor.

Suggestion:
Sperm motility is recognized as an important factor that plays a role in fertility, but it should not be used alone as a determinant of a stallion’s reproductive potential simply because of its historic application. The consequences of incorrectly diagnosing a subfertile stallion simply because of low progressive motility particularly in the face of high total motility, are profound and when incorrect result in a loss of credibility to the profession and the procedures that are so critical to our discipline. We currently have many other assays available to supplement the evaluation of sperm motility as a measure of sperm quality to provide a more thorough evaluation of sperm quality and render a clearer diagnosis.

Morphologic evaluation of sperm quality
Enumerating the shape of a stallion sperm population is an important part of the BSE. The percent morphologically normal is used to calculate the total number sperm that are progressively motile
and morphologically normal. This number, in the second ejaculate (at least 1.0 billion in the month of December), is considered an important value when determining the final classification status (Satisfactory, Questionable, Unsatisfactory). A relevant criticism of this composite number is that both progressively motile and morphologically normal sperm are, to a certain degree, measuring the same endpoint (i.e. a good sperm) and therefore the total number of sperm that a stallion has in the second ejaculate is being evaluated twice and that stallion may unfairly be “downgraded” in classification status. It should be recognized that there are sperm abnormalities such as proximal droplets, distal droplets, as well as abnormal heads and midpieces that may be “progressively motile” and if not accounted for in the composite number may artificially inflate a stallion’s “normal” sperm number. While this is an imperfect number, an alternative would be to replace progressively motile sperm with total motile sperm.

Evaluation of sperm morphology in the stallion is performed differently than other species that originally adopted the primary, secondary, and tertiary nomenclature to describe the origin of the abnormality (testis, epididymis, iatrogenic). In contrast, the stallion BSE identifies the specific sperm shape (abnormal head, detached head, coiled tail, etc). While the Manual stated that “the numbers of morphologically abnormal or non-motile forms is relatively unimportant concerning stallion fertility and are important only in calculating the number of normal sperm” the type of abnormality may be important diagnostically, particularly when the percent of normal sperm is low. As an example, a stallion may have a low percent normal sperm due to transient defects such as distal droplets (a better prognosis) or due to defects associated with sperm production in the testes such as abnormal heads and midpieces (potentially a worse prognosis). Diagnostically, identification of specific abnormalities not only suggest the origin of the defect, but may narrow the diagnosis and aid the clinician in developing a prognosis and treatment.

The evaluation of sperm morphology may be the single most important sperm assay available to the practitioner because sperm shape in most cases is not altered by iatrogenic causes such as improper handling, and the results reflect what the stallion is ejaculating, without the bias inherent in sperm motility evaluations. It also lends itself to remote evaluation since sperm samples can be fixed on-farm and transported to specialists for evaluation and interpretation.

**Fertility and prediction**

One challenge of the BSE is that the practitioner/owner would like a prediction of fertility. The best measure of fertility, as is often stated, is breeding females and determining pregnancy. This is however, simplistic because the level of “satisfactory” fertility is difficult to define in the horse industry due to the varying owner expectations. Some owners/syndicates may view acceptable fertility based on the number of foals produced to satisfy an economic endpoint, while others may be view acceptable as the production of an individual foal to continue a particular line from an old subfertile stallion or mare, to satisfy an individual goal.

The goal of most examinations is the prediction of an outcome. In the case of the stallion BSE, fertility prediction based on a set of measurable variables would be desirable. It is commonly stated that the BSE, regardless of species, cannot predict fertility, but rather is better at identifying the low/sub-fertile individual. This conclusion is often based on the reliability of the clinical outcome.

For instance, stallions with poor sperm quality are more likely to also have low fertility because regardless of the effects of mare or management, sperm quality will be the factor that limits fertility. Therefore, there is essentially only one outcome for a male with poor sperm quality and that is a reduction in fertility making the clinical diagnosis “correct” most of the time. In contrast, the stallion with “good” sperm quality has two options. One, fertility is representative of the level of sperm quality (i.e. good, excellent); or two, fertility is not representative and thus sperm quality is perceived as not “predictive” of fertility. The latter example is what leads to the conclusion that the BSE cannot predict fertility, which is unfortunate because in this case measuring the role of the female and management is critical, but often ignored, leading to the conclusion that the sperm test is no good. It is therefore likely that if the fertility of mare quality and management could be measured and accounted for, that good sperm quality would be predictive of fertility. A recent study highlighted the prominent role that mare (increasing mare age) and
a management factor (the use of cooled-shipped semen) had in reducing fertility in a group of stallions in which sperm quality was uniformly high and did not affect fertility.8

Previous authors have described the relationship between fertility groups (high and low) and sperm motility6 or morphology9 as well as the relationship between morphologic features and fertility.7 A recent study10 compared the relationship between sperm motility, morphology and three fertility groups (high, average, and below average; Table). This study described the relationship between three fertility measures (seasonal pregnancy rate, percent pregnant per cycle and pregnant per first cycle) and sperm quality measures. The only sperm quality measure able to discriminate groups of high and low seasonal pregnancy rate was percent total motility. However, when percent pregnant per cycle and percent pregnant per first cycle were used, numerous sperm quality measures discriminated different fertility groups. Why then are there stallions with excellent sperm quality that do not have good fertility? These stallions truly exist, such as those whose acrosomes do not react in the face of excellent sperm quality (motility, morphology).11

An additional factor that limits our ability to predict fertility is the fertility measure itself. Seasonal pregnancy rate is a common measure used to describe fertility compared to pregnant per cycle or pregnant per first cycle. Yet when these three measures were used to describe fertility in a group of clinically fertile stallions that were divided into high, average and low fertility, all three of the seasonal pregnancy rate (SPR) groups had a stallion with a 100% seasonal pregnancy rate (Table 1).10 Clinically, this suggests why some stallions are “fertile” based on SPR while their sperm quality suggests they are less fertile based on a more sensitive measure such as percent pregnant per cycle or pregnant per first cycle. The limitation of seasonal pregnancy rate as a measure of fertility is that it is an accumulation of breedings and not a measure of breeding efficiency allowing less fertile stallions to “catch up” by breeding mares more cycles.

Suggestion

Poor sperm quality, by itself, can be the limiting factor in fertility, but good sperm quality does not assure that other factors (mare, management) are also optimal and not in themselves a limiting factor. Instead of this finding being included as part of the diagnosis (i.e. this is actually a fertile stallion, look elsewhere for the cause of subfertility), the conclusion is that this is truly a subfertile stallion and this sperm assay is inadequate. If you change the assumption from the “the assay is inadequate” to “the assay is always inherently correct” it means the diagnostician must then pursue another reason for the “stallion’s” subfertility.

There have been few studies that have compared fertility levels with sperm quality features. A common outcome is that the relationship (usually correlative) may be low, but statistically significant. The conclusion drawn from these results is that the relationship of sperm quality to fertility is of little use clinically (i.e. not predictive). This is unfortunate because included in these “low” correlations are mare and management factors that are not accounted for in the analysis. If these factors were accounted for, correlations would likely be “higher”.

The dilemma of classification

Similar to the bull, a classification system (Satisfactory, Questionable, Unsatisfactory, Classification Deferred) is provided in the BSE manual. The two classification systems (bull and stallion) however, have different intents. The intent of the bull system is to identify those individuals that, based on sperm criteria as well as potential physical and behavioral characteristics, do not attain a certain threshold, and eliminate them from the breeding population. The stallion BSE, as mentioned earlier, has the intent of providing a framework to perform a thorough examination allowing the practitioner to identify causative fertility limiting factors. The classification categories are intended to provide a basis for the practitioner to give perspective (small testes, average sperm quality) and interpretation about a stallion’s reproductive capability. The classification system is not a Pass/Fail system which suggests that stallions that Fail should or cannot be bred. The concept of Fail may have long-standing effects on a stallion’s reputation that reduce the stallion’s economic viability, particularly if the fertility limiting
condition is transient. The classification system is not intended to provide a platform for the practitioner to pass judgment as to whether a stallion should or should not be bred. This approach is particularly important regarding conditions that may be potentially genetically based (cryptorchidism, low sperm numbers or quality) versus those that are definitively genetically based (hyperkalemic periodic paralysis [HYPP], hereditary equine regional dermal asthenia [HERDA]). As information becomes available about conditions like cryptorchidism it may be apparent that they are multifactorial (genetic and environmental) rather than exclusively genetic in origin.12 Even in the case of those conditions that have been established as genetic, the ultimate resolution should be the purview of the breed associations. It should be the responsibility of the practitioner to provide accurate information regarding the condition so the owner and breed association can render a responsible opinion.

The Society for Theriogenology Manual for Clinical Fertility Evaluation of the Stallion has been an extremely useful tool for the veterinary practitioner since it was introduced in 1983. There has been considerable information available since the introduction of the manual that can would be useful in updating and further promoting this valuable publication.

References
Table. The mean (+ SD) and range () for stallion sperm motility and morphology variables for the percent mares pregnant / cycle. (reprinted from Love, Theriogenology 2011)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Seasonal Pregnancy rate (%)</strong></td>
<td>97 ± 4 (90-100)</td>
<td>86 ± 10 (50-100)</td>
<td>61 ± 29 (12-100)</td>
</tr>
<tr>
<td><strong>Pregnant / cycle (%)</strong></td>
<td>91 ± 10 (75-100)</td>
<td>56 ± 6 (45-74)</td>
<td>32 ± 13 (8-45)</td>
</tr>
<tr>
<td><strong>Pregnant / first cycle (%)</strong></td>
<td>91 ± 10 (75-100)</td>
<td>58 ± 16 (0-88)</td>
<td>34 ± 14 (0-50)</td>
</tr>
<tr>
<td><strong>Sperm motility values</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total motility (%)</strong></td>
<td>83 ± 5 (76-91)</td>
<td>76 ± 12 (44-92)</td>
<td>48 ± 21 (18-66)</td>
</tr>
<tr>
<td><strong>Progressive motility (%)</strong></td>
<td>77 ± 6 (64-85)</td>
<td>71 ± 12 (43-88)</td>
<td>44 ± 20 (16-63)</td>
</tr>
<tr>
<td><strong>Path velocity (µ/s)</strong></td>
<td>196 ± 26 (163-230)</td>
<td>190 ± 19 (157-241)</td>
<td>162 ± 41 (104-194)</td>
</tr>
<tr>
<td><strong>Progressive velocity (µ/s)</strong></td>
<td>174 ± 20 (146-207)</td>
<td>171 ± 17 (143-208)</td>
<td>139 ± 34 (88-166)</td>
</tr>
<tr>
<td><strong>Morphology</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Normal</strong></td>
<td>67 ± 8 (50-76)</td>
<td>48 ± 15 (17-83)</td>
<td>41 ± 27 (7-85)</td>
</tr>
<tr>
<td><strong>Abnormal heads</strong></td>
<td>9 ± 6 (1-22)</td>
<td>12 ± 9 (0-45)</td>
<td>17 ± 13 (1-43)</td>
</tr>
<tr>
<td><strong>Detached heads</strong></td>
<td>2 ± 2 (1-8)</td>
<td>2 ± 2 (1-10)</td>
<td>6 ± 11 (0-44)</td>
</tr>
<tr>
<td><strong>Proximal droplets</strong></td>
<td>8 ± 3 (5-14)</td>
<td>20 ± 14 (2-51)</td>
<td>25 ± 18 (4-60)</td>
</tr>
<tr>
<td><strong>Distal droplets</strong></td>
<td>6 ± 4 (2-16)</td>
<td>8 ± 7 (0-28)</td>
<td>2 ± 2 (1-7)</td>
</tr>
<tr>
<td><strong>Bent midpieces</strong></td>
<td>1 ± 1 (0-4)</td>
<td>0.2 ± 0.4 (0-1)</td>
<td>1 ± 5 (0-17)</td>
</tr>
<tr>
<td><strong>General midpiece abnormality</strong></td>
<td>6 ± 4 (1-14)</td>
<td>7 ± 6 (1-29)</td>
<td>10 ± 6 (2-26)</td>
</tr>
<tr>
<td><strong>Hairpin tail</strong></td>
<td>4 ± 3 (0-9)</td>
<td>4 ± 4 (0-16)</td>
<td>5 ± 6 (0-20)</td>
</tr>
<tr>
<td><strong>Coiled tail</strong></td>
<td>1 ± 1 (0-3)</td>
<td>2 ± 2 (0-7)</td>
<td>5 ± 5 (0-14)</td>
</tr>
<tr>
<td><strong>Premature germ cell</strong></td>
<td>1 ± 2 (0-6)</td>
<td>1 ± 2 (0-8)</td>
<td>2 ± 2 (0-7)</td>
</tr>
</tbody>
</table>

Percent pregnant / cycle groups – Group 1- ≥ 76% and ≤ 100%; Group 2- ≥ 46% and < 76%; Group 3- ≥ 0% and < 46%.

1-P<0.0001
2-P<0.001
3-P<0.0002
4-P<0.04
5-P<0.04
6-P<0.008