BRITISH JOURNAL OF BIOMEDICAL SCIENCE, 2018
VOL. 75, NO. 2, 53–60
https://doi.org/10.1080/09674845.2017.1377591

# Sperm toxicity testing: UK best practice guideline from the Association of Biomedical Andrologists

S Long[a], B Woodward[b] and M Tomlinson[c]

[a]Andrology Department, Heart of England NHS Foundation Trust, Heartlands Hospital, Birmingham, UK; [b]X&Y Fertility, 144a New Walk, Leicester, UK; [c]Fertility Unit, Nottingham University Hospital, Nottingham, UK

**ABSTRACT**
In order to ensure the quality and integrity of diagnostic semen analysis results, materials used should be tested to ensure that they do not interfere with sperm function. As a toxicity test, complex sperm function testing may be considered controversial, since the fertilizing capacity of single sperm can never be assured. In preference, sperm motility offers a unique means of assessing the toxicity of reagents and materials before they are used in routine practice. Motility is the semen parameter most likely to be influenced by the external environment. Indeed, it is the main reason that laboratories insist on supplying their own approved specimen containers and ensuring that patients, as far as possible, adhere to strict conditions for sample collection and transport prior to testing. This differs to other indirect tests of toxicity such as the mouse embryo assay, whereby the rate of mouse pre-implantation embryo development to the blastocyst stage is compared. This guideline is aimed at health care scientists who deal with andrology in both general pathology and specialised fertility laboratories, and provides a model approach to sperm toxicity testing. For assisted reproduction clinics, the same methodology can be used to test any consumables that are used for sperm processing, and as an indirect guide for any consumables that come into direct contact with oocytes and pre-implantation embryos.

**ARTICLE HISTORY**
Received 28 July 2017
Accepted 18 August 2017

**KEYWORDS**
Toxicity testing; andrology; semen analysis; ISO15189; accreditation; quality control

## Introduction, background and scope of guideline

In order to ensure the quality and integrity of diagnostic semen analysis results, materials used should be tested to ensure that they do not interfere with sperm function. This is in accordance with the latest World Health Organisation (WHO) guidelines [1] that state that processes must be optimised in order that the patient can provide a suitable specimen for testing and that external influences on the test result are kept to a minimum. For assisted reproduction clinics, the same methodology can be used to test any consumables that are used for sperm processing, and as an indirect guide for any consumables that come into direct contact with oocytes and pre-implantation embryos.

As a toxicity test, complex sperm function testing may be considered controversial, since the fertilizing capacity of single sperm can never be assured. Rather, sperm motility offers a unique means of assessing the toxicity of reagents and materials before they are used in routine practice. Motility is the semen parameter most likely to be influenced by the external environment. Indeed, it is the main reason that laboratories insist on supplying their own approved specimen containers and ensuring that patients, as far as possible, adhere to strict conditions for sample collection and transport prior to testing [2]. This differs to other indirect tests of toxicity such as the mouse embryo assay (MEA), whereby the rate of mouse pre-implantation embryo development to the blastocyst stage is compared, and therefore may be indicative of human embryo toxicity.

In Europe, another selection criterion for consumables may be whether they are CE-marked. CE-marking shows that the manufacturer has checked that a product meets European Union (EU) safety, health or environmental requirements, and is an indicator of a product's compliance with EU legislation [3]. This, however, does not necessarily indicate a higher level of quality and CE-marking does not negate the necessity to toxicity test.

The aim of this guideline is to assist health care scientists who deal with andrology in both general pathology and specialised fertility laboratories, by providing a model approach to sperm toxicity testing (STT).

## Testing seminal fluid or a standard sperm suspension media

Since diagnostic testing conducted in an andrology laboratory usually involves assessment of semen samples where sperm are present, some might consider that sperm should be tested within seminal fluid as part of the STT programme. At first, the idea of using unprocessed

**CONTACT** S Long enquiries@aba.uk.net

© 2018 British Journal of Biomedical Science

semen samples might appear to have some clear advantages, especially in terms of convenience and time management. However, there are a number of reasons why use of unprocessed semen for STT should be avoided. Firstly, seminal fluid is a complex suspension of salts, sugars, proteins, lipids, nucleic acids, reactive and non-reactive intermediates [4]. The exact composition of these constituents not only varies between men, but also differs in samples from the same man. Therefore, the behaviour of sperm in response to many of these biochemicals upon exposure to different laboratory materials will vary from sample to sample. For this reason, the sperm should be removed from seminal plasma prior to testing and resuspended in a suitably consistent media, which will support sperm motility for several hours. The process of seminal plasma removal can either be via washing through density gradient centrifugation (DGC) or by a technique called 'swim-up', whereby a semen sample is overlaid with media and motile sperm swim up to the top layer. Both methods allow an enriched population of sperm to be harvested with a high degree of motility. From this starting sperm population, any significant change to motility may be more easily detected than from a largely heterogeneous sample containing debris, non-sperm cells and immotile sperm.

Most commercially available sperm wash/re-suspension buffers are considered suitable. It is recommended that the same type of media is consistently utilised thereafter, since there may be subtle differences in both composition and performance in terms of sustaining motility over a prolonged period. For example, Figure 1 shows how motility varied for sperm incubated in three different makes of commercially available media over a 24-h period. Significant differences in motility were detectable at both early (3-h) and late (24-h) stages. This does not necessarily preclude the changing from one media to another, but laboratories should be aware that this can lead to significantly altered performance. Commercial suppliers of sperm wash media and gradients provide instructions for the use of these products. These should be followed, unless in-house methodologies have been suitably validated. All procedures should be verified in-house to demonstrate acceptance prior to use.

## Preparation method: the difference between density gradient centrifugation and swim-up

DGC is a method that separates sperm cells based on their density and therefore results in the distribution of sperm within the gradient layers that match their density. Mature sperm usually have more compact chromatin and are denser than immature sperm.

Commercial DGC media consist of culture media containing a suspension of microscopic beads, usually made from silane-coated silica particles. The percentage concentration of the microscopic beads varies. Most DGC media contain a pre-prepared 'high' and 'low' level suspension, where the particular v/v% (concentration) may also be different. High-level suspensions usually contain 80% microscopic beads, whilst low level suspensions usually contain 40% microscopic beads. Some DGC packs are supplied as a stock solution of 100% microscopic beads that requires diluting with a buffered medium prior to use. A gradient is prepared by layering a lower density suspension above a higher density suspension (see Technique 1). Following the DGC preparation (see Technique 2 and Figure 2), in a solution gradient where the lower layer has a specific gravity of 1.10 g/ml (80%), only the most mature and normal sperm should penetrate this layer [5]. The 'swim-up' method (see Technique 3 and Figure 3) is based on the ability of sperm to migrate through cervical mucus *in vivo*. This separates the population of sperm that may be more likely to have fertilizing potential from the population that may not. The swim-up method uses the same principle, and is much simpler to perform than the DGC procedure.

## How should motility be assessed?

All motility assessment should be undertaken via microscopic examination. Manual motility assessment is the recommended method by the current WHO guidelines [1]. However, despite its technical complexity and subjectivity, the 4-tier grading system (a, b, c and d) based on estimates of progressive swimming speed, allows the operator to attempt to distinguish between rapidly progressive motility and sluggish progressive motility. This is the method recommended by the Association of

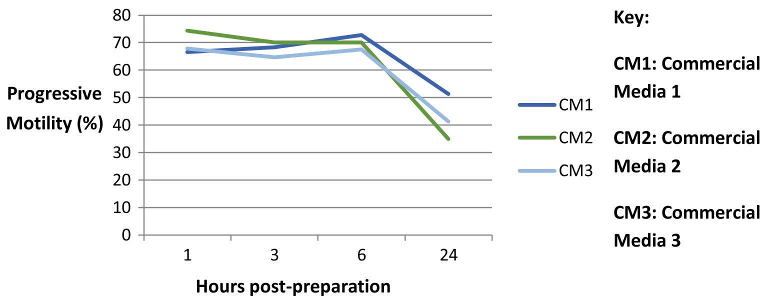


**Figure 1.** The effect of different media on sperm motility following preparation via density gradient centrifugation.
Note: This figure demonstrates how motility varied for a pooled pellet of sperm, which was prepared from semen samples from 15 donors via density gradient centrifugation (DGC). The pellet of sperm was equally divided and then washed through one of three different commercial media. Motility was compared at 1, 3, 6 and 24 h post-preparation. There were clear differences in motility depending of the brand of the media used, especially at 24 h.

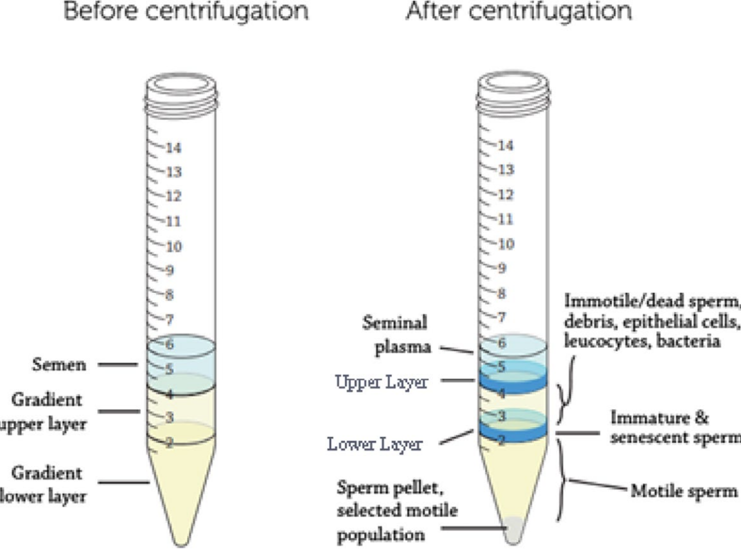


**Figure 2.** Density gradient tubes before and after centrifugation.
Note: Motile sperm should be present in the pellet of the sample after centrifugation. This pellet should be washed via centrifugation through medium prior to use (adapted from 6).

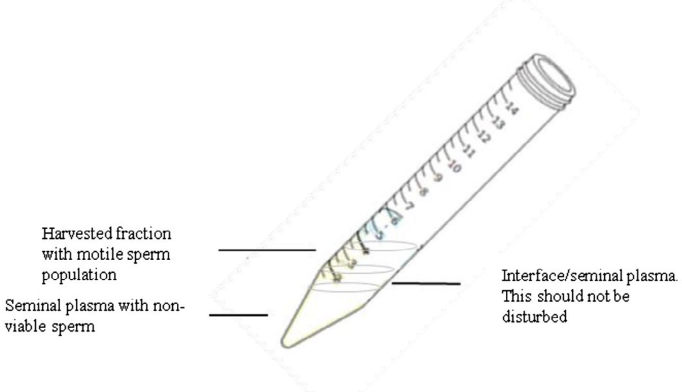


**Figure 3.** Swim up preparation after initial preparation.
Note: The uppermost layer should be harvested following direct swim-up, as this should contain the motile sperm [6].

Biomedical Andrologists (ABA) [7,8]. In clinical terms, the three-grading system recommended by the current WHO guidelines [1] (where a + b are combined as 'progressively motile sperm') can be misleading since ability to conceive is considered to rely on both the number of sperm moving forward and also the speed of progression [9–11]. Therefore, laboratories using a manual motility assessment to determine STT should be aware of this potential limitation of their testing procedure, and document it accordingly in line with uncertainty of measurement protocols [12].

Computer-aided semen assessment (CASA) can also be used for motility assessment. However, it is recommended that CASA system also provides robust measures of sperm concentration, since the accuracy of the motility measurement is dependent on the ability to initially correctly identify all the sperm and distinguish them from non-sperm objects. Use of indirect CASA methods based on light scatter and provided by algorithm only, should be viewed as inherently risky and is not recommended.

## Which products require testing?

The list of products requiring testing depends on the scope and extent of the service and if necessary should cover all diagnostic and processing materials which sperm may be exposed to. Although it is difficult to exclude any item which may come into contact with sperm, the duration of testing can be limited to the duration of the likely maximum exposure limit (MEL). For example, there is little point in testing the toxicity of a pipette tip for several hours if the MEL is less than 1 min. Testing duration therefore needs to be pragmatic and tailored to the MEL for each material/consumable. A schedule is suggested in Table 1.

## Routine examples and applications to consider

Sperm added to a glass slide should always be covered with a coverslip since they are used together. The STT may be undertaken by adding an aliquot, e.g. standard 10 µL to a slide, adding a coverslip and then testing against a

suitable control. Repeat measurement could take place at the designated time period(s) provided the slide is kept in a humidified environment, since a non-humidified environment might cause the slide to dry out, which might adversely affect the volume of suspension and motility. Testing of specimen collection containers (or sheaths) has to be aligned with the scope of service and testing need only take place if procedures suggest that samples are likely to remain in these containers for extended periods of time. The MEL should not exceed 4-h, since this represents the maximum recommended delivery time that may be used for a post vasectomy sample [13].

Plasticware used for cryopreservation represents something of an anomaly, since there is an argument for a MEL of many years, even decades, within the cryopreserved environment. True representation of the toxic effects of cryopreservation plasticware is therefore difficult to represent. Motile sperm in suspension placed in a cryo-vial or straw may be cryopreserved within minutes and then removed rapidly after warming. Therefore, although the testing procedure should account for at least this exposure period, extending the time to a 24-h period may be of value in the first instance in order to satisfy the service that the long-term risks to cryopreserved material are small.

Petri dishes may be used in some andrology services for the dissection and analysis of surgically retrieved sperm and extended (even overnight) incubation may take place. In such instances, laboratories can consider analysing sperm which have been incubated in a petri dish with a deliberately 'scratched' surface, particularly if damage to the plastic surface is associated with the dissection of materials using sharp objects such as sterile needles or sterilised glass slides. Moreover, cryopreservation or *in vitro* fertilization (IVF) services might incubate sperm for extended periods and therefore toxicity testing should reflect this time period.

**Table 1.** Expected routine laboratory exposure of sperm to materials used in short-term (<2 h) or long-term (24 h) toxicity experiments.

| Short-term exposure (≤1 h) | | Long-term exposure (≤24 h) | |
|---|---|---|---|
| Item | Routine exposure (min) | Item | Routine exposure (h) |
| Microfuge tubes | 15–20 | Specimen collection container | 1–4* |
| Sterile/non-sterile pipettes | 1–5 | Specimen collection sheath | 1–4* |
| Pipette tips | 1–5 | Conical centrifuge tubes | 1–4* |
| Cryo-straw filling nozzles | 5 | 21-gauge needles | 0.5 |
| CASA slides | 5–10 | Cryo-straws | 1–24 h |
| Large-volume fixed-depth slides | 60 | | |
| Plain glass slides/ coverslip | 5–10 | Cryovials | 1–24 |
| IUI catheters | 30 | Petri dishes | 24 |

*Depends on scope of service.

## What is a 'suitable control'?

A control is generally a consumable that has been verified as non-toxic by MEA or HSSA (Human Sperm Survival Assay). Testing the control will often involve using a consumable that has previously been tested for toxicity in order to assess the motility, i.e. a glass slide/coverslip. The STT Index (described later) will need to be calculated for the control and be within acceptable limits (unless another method is employed).

## Frequency of testing and numbers to be tested

Although testing of all the products listed in Table 1 might appear to be a lengthy task, STTs do not have to be carried out simultaneously for all consumables. Rather, as new batches/products are to be introduced for use in the laboratory, often at different times throughout the year, then testing can be staggered accordingly.

The STT procedures can also be used as a dual method to simultaneously assess reagents, i.e. the sperm preparation medium alongside the centrifuge tube used. Slides and coverslips can be tested together and if there is a batch failure, investigations can be implemented to discover which item may be the cause. This can be incorporated into the quality management system (QMS). Retained stock should be suitably represented within the testing procedure, e.g. by performing a risk assessment if large quantities of the same batch are ordered ensuring sufficient numbers are tested. Large numbers ordered at the same time may also be subject to differences in batch numbers, and this will therefore require additional STT tests for each separate batch.

## Patient consent and ethics

The provision of samples for STT should not need any ethical approval. The Royal College of Pathologists (RCPath) guideline [14] states that laboratories can use specimens which are considered outside the remit of the Research Ethics Committees (RECs), to enable a pragmatic approach to use of patient samples for low risk activities. In diagnostic laboratories, excess 'waste' sample, which would otherwise be discarded, can be used for the STT procedures under this guidance [14]. However, it is recommended that diagnostic laboratories obtain consent from patients for the use of residual sample following examination procedures.

For excess samples that have been processed for treatment, relevant information should be provided to, and written consent should be obtained from, the person providing the sample before it is used for quality purposes. In some countries, this is a legal requirement. Such sperm should only be used in line with what the sperm provider has understood and explicitly allowed, and which is documented.

## Staff responsibilities, quality control and quality assurance

It is the responsibility of laboratory management to ensure that there are adequate resources for the undertaking of Quality Assurance (QA) procedures. All personnel undertaking STT should be sufficiently trained and assessed as competent to conduct the procedure, including all necessary competency assessments to demonstrate relevant proficiency. When defining the STT procedure, the laboratory should have robust Internal Quality Control (IQC). This should demonstrate an appropriate level of agreement between all operators. The laboratory should also ensure that they follow all appropriate QA procedures to minimize anything that could impact the results, e.g. temperature monitoring and equipment calibration.

## Uncertainty of measurement of motility

Current WHO guidance on motility assessment [1] encourages the adoption of a simpler approach by categorization of sperm into only 3 grades of motility (progressive, non-progressive or immotile). The intention was to remove a degree of subjectivity which appears to be inherent in manual assessment of motility [15]. However, by placing the emphasis on only those sperm which have progressive motility, there is no means of distinguishing 'how well' the sperm progress, despite this impacting on for the likelihood of conception.

Experienced operators may focus on a moving object or observe a microscopic field for several minutes whilst highly motile sperm will have entered and left. This has the potential for only the immotile/non-progressive fraction to be enumerated with any accuracy and an over-estimation of motile sperm, which then can be compounded in samples with higher density and sperm with high velocity. This may not be highly significant provided the level of error is consistent and repeated measures are performed to demonstrate reproducibility of measurement both across all operators and by a single operator. CASA may reduce this bias and enable estimation of velocity.

Over 25 years ago, two studies of motility assessments of frozen-thawed sperm by different staff, demonstrated a 9–14% inter-operator variation [16] and coefficients of variation of 21% across 10 laboratories [17]. Since frozen-thawed sperm motility is often reduced compared to fresh motility, the variation in fresh samples with highly motile sperm is likely to be compounded. Moreover, providing completely 'blinded' (i.e. without operator bias from prior knowledge of the first result) repeat manual motility measurement are technically difficult to demonstrate on fresh specimens. For this reason, it is recommended that centres consider using inter-operator variation and inter-laboratory comparison (ILC). Whatever variation is demonstrated within the laboratory, it should be accounted for when deciding upon an appropriate threshold for determination of toxicity.

Since sperm motility is highly temperature sensitive, STT should be carried out at a consistent temperature, e.g. by use of a heated microscope stage, which has been validated to ensure the motility assessment takes place at 37 °C (taking into account uncertainties associated with temperature readings)(Technique 4).

## Testing procedure, recording of data from testing and retention periods

A basic testing procedure must be in place with all motility values being undertaken in accordance with the current WHO guidelines [1]. The baseline data should be assessed from the control ($t = 0$) and the sperm suspension applied as directed in Table 2. Motility should be assessed at appropriate intervals, based on the MEL (Table 2) alongside the control. The exact procedure used will depend on the method employed to assess toxicity. All assessments should be recorded using controlled documents that are fully auditable and traceable. All reagents and consumables used in the procedure should also be recorded along with actions taken in the event of adverse test results (or at a minimum the reference to the electronic records).

## Batch rejection criteria – how toxic is toxic?

The main objective of STT is to ensure that consumables from a new batch perform to the same level as the previous batch without any increased level of toxicity to sperm. It is also important to establish whether the physical consistency of the consumables affect the outcome of a test. For example, a laboratory's standard centrifuge tube may demonstrate adequate batch-to-batch consistency, yet conceivably the motility of the sperm within the tube may be significantly lower than when in another consumable, e.g. the specimen collection container. There are three ways to consider the outcome of STT data include: gathering baseline data from all consumables tested; using an index as a measure of the toxicity; or statistical analysis. The laboratory should decide which method is best for them to apply for routine practice.

## Using baseline data

It may be prudent to create 'baseline' data which provides an indication of whether different products used within the laboratory provide the same conditions for sperm motility or determine whether they may affect it in a negative way. Having preliminary toxicity data from a number of samples allows the laboratory to gain a sense of perspective and expectation of the likelihood of any toxicity across a range of materials. Furthermore, as well as testing a range of materials, it is recommended that laboratories consider using a number of different semen samples, since 'one off' testing may be misleading.

Perhaps the best analogy is the observation of variability amongst samples that are cryopreserved. Cryopreservation commonly shows high inter-patient variation and the majority simply do not adequately

**Table 2.** Application of sperm suspension to product.

| Item under test | Potential method of adding suspension | Considerations |
|---|---|---|
| Specimen collection containers<br>Centrifuge tubes, catheters, straws, vials, etc. | Place a minimum of 250 µL of well-mixed washed sperm suspension into the test product | Some specimen collection containers will require more than 100 µL added as the suspension should have made contact with as much of the container internally as possible |
| Pipettes, pipette tips | For pipettes and pipette tips draw up a minimum volume (this depends on the pipette type and use) | Some pipettes may not have a volume of 100 µL. These should be set at maximum volume<br>Pasteur pipettes may require more than 100 µL |
| Glass slides, coverslips | For plain glass slides, add 10 µL and cover with a 22 × 22 mm coverslip | Leave in a humid environment |
| Fixed depth chambers | Fill the chamber with appropriate volume for chamber, wipe off the excess at the fill port | Leave in a humid environment, e.g. petri dish containing moist tissue |
| Collection condoms (sheath) | See 'specimen collection containers' | The testing method will be determined by the method of removal of semen. Ensure this is reflected within the procedure |

withstand the addition/removal of cryo-protectant and cooling. Moreover, samples from the same individuals where cryopreservation has been undertaken show a reduced quality or death of some sperm [17]. Therefore, it is only to be expected that variation will also be observed in the way samples respond to long-term incubation and contamination with a variety of laboratory materials. Baseline data from at least 10 samples can assist in gaining proper perspective across all materials in terms of the average motility over an appropriate time-course.

This need not be an onerous task and can be separated into short- and long-term exposure experiments with data gathered over several weeks. If no previous batch is available for comparison, thresholds for rejection can be based on motility in the test container at any given time point using Standard Errors (×2 or ×3) from the baseline data for the control (see Table 3). Each laboratory's definition of toxicity should be generated from their own reference data but taking into account the relative sensitivity of the analysis provided.

## Using a sperm toxicity testing index (STTI)

Sperm survival assays have previously been used to test the potential reprotoxicity of products utilised for assisted conception over a 24-h period [18]. This group described a threshold for toxicity (termed a 'SpST index') of < 0.85 at 24 and 96 h of exposure. A similar threshold could be adopted for acute toxicity [19]. However, selection of a suitable threshold should take into account the sensitivity and reliability of the measure (motility) in question. For example, there is no justification in setting a rejection threshold at 85% of control if the percentage of progressive sperm has a measurement sensitivity of +/– 20%. As outlined in the previous section on uncertainty of measurement, sensitivity of the laboratory's motility measurement should first be calculated from either repeated measures (performed on separate aliquots) or inter-operator variation. For this reason, it may be prudent to make allowance and adopt a less stringent approach which may relate better to the baseline approach described above. See Table 3 for example of how to calculate the STTI.

## Statistical assessment

Current WHO guidance [1] describes several methods to use in QA and QC procedures, stating the specific use of a paired *t*-test for assessing whether consumables are toxic to sperm. The use of a particular test depends on the data distribution and whether it is normally or non-normally distributed. Other statistical methods such as analysis of variance (ANOVA) may also be used to demonstrate overall differences in data distribution, but the eventual test of choice should depend on the data, and if in doubt a statistician should be consulted.

## Investigation of an adverse test outcome

Whether the baseline or STT approach is used, rejection of a batch of materials should be confirmed by re-testing. Providing that acceptance criteria have been carefully considered, confirmatory testing using a second semen sample should provide sufficient internal justification for withdrawal of the batch.

However, laboratories considering raising the test results as a performance issue with the manufacturer should only do so having performed a more detailed investigation. The same applies should a laboratory decide that materials are clearly acutely toxic and feel obliged to inform the wider andrology community. This could include regulators such as the Human Fertilisation and Embryology Authority (HFEA) or the Medicines and Health care Products Regulatory Authority (MHRA), but should only be considered after further investigation and appreciation of any wider implication related to

**Table 3.** Example for calculating sperm toxicity and the sperm toxicity testing index using standard error.

1. For a set of 10 samples, each prepared in the same standard sperm wash buffer gives: 75,80,95,69,88,78,90,72,80. Mean of 81.2, SD 8.25
   Rejection criteria at mean-3xSD = mean (81.2%) - 2xSD (24.8) = 56.4%
2. Using STT Index (as opposed to baseline data)

   Another method of comparing results is to use a comparison of the test motility values to the control motility values, although this is usually within the context of longer term survival assays. This can be undertaken by a simple calculation once motility has been ascertained of the ratio between the two results:

$$\text{STT Index} = \frac{\text{Progressive Motility (\%) of TEST sample}}{\text{Progressive Motility (\%) of CONTROL sample}}$$

Where TEST is from the new batch and CONTROL is material from the previous batch.

reputational risk. More detailed investigation in this context should state clearly the level of toxicity likely to be observed and include power calculations which identify the numbers of samples required to demonstrate this beyond a reasonable doubt. In many cases, as few as 10–20 samples may be sufficient and motility data are compared between test and control groups. Analysis need not be complex but, where there is doubt, should be performed in consultation with a statistician.

## Special considerations

*Samples to use:* It is recommended that only non-viscous samples without evidence of agglutination or aggregation are used for STTs as such characteristics will introduce further uncertainty into the process. Other considerations may influence the laboratory decision for using a sample, but these are at the discretion of the laboratory. *Overall risk:* Despite the importance of demonstrating that materials which are in contact with sperm have undue influence on their function, the overall risk associated with material toxicity remains low. This guideline should be used in conjunction with the laboratory procedures on STT to ensure that risk associated with material toxicity can be adequately detected if it at all exists.

## Disclaimer

The ABA is not recommending any suppliers or reagents. This guideline should be used in conjunction with manufacturer's recommendations. The ABA does not support and will not take responsibility for the use of this guideline in any correspondence with a manufacturer about the quality of their product. Substandard results, as determined when following the guideline, may not be directly be related to product quality and the ABA strongly advises to proceed cautiously.

## Summary

STT is vital in assuring the reliability of the results being sent out to users of the service. Correct application of the principles within this guideline will allow the laboratory to take some steps in assuring their results. It is vital that each service assesses the risk of undertaking testing and give justifications, where appropriate as to why the protocols are set as they are.

## Acknowledgements

We are grateful to all of the andrologists, particularly the members of the ABA Executive Committee, who provided useful comments to help improve this guideline.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## ORCID

*S Long* http://orcid.org/0000-0002-2172-8637
*B Woodward* http://orcid.org/0000-0001-8669-4776

## References

[1] World Health Organization. WHO laboratory manual for the examination and processing of human semen. 5th ed. Geneva: WHO Press; 2010.
[2] Woodward BJ, Tomlinson MJ. How strict should specimen acceptance or rejection criteria be for a diagnostic semen analysis? Hum Fert. 2015;18(2):87–91.
[3] CE Marking. Guidance: CE marking [Internet]. Mölndal [updated 2012; cited 2017 Aug 14]. Available from: https://www.gov.uk/guidance/ce-marking
[4] Owen DH, Katz DF. A review of the physical and chemical properties of human semen and the formulation of a semen stimulant. J Andrology. 2005;26(4):459–469.
[5] Bjorndahl L, Mortimer D, Barratt CLR, et al. A practical guide to basic laboratory andrology. Cambridge: Cambridge University Press; 2010. ISBN 1139482491, 9781139482493.
[6] Nidacon International. Products [Internet]. Mölndal. [2017; cited 2017 Aug 14]. Available from: http://nidacon.com/ifu/density_gradient.pdf
[7] World Health Organization. WHO Laboratory Manual for the Examination of Human Semen and Semen-Cervical Mucus Interaction. 4th ed. Geneva: WHO Press; 1999.
[8] Tomlinson MJ, Harbottle SJ, Woodward BJ, et al. Association of biomedical andrologists. Association of biomedical andrologists – laboratory andrology guidelines for good practice version 3 – 2012. Hum Fert. 2012;15(4):156–173.
[9] Barratt CLR, Tomlinson MJ, Cooke ID, et al. Prognostic significance of computerised motility analysis for *in vivo* fertility. Fertil Steril. 1993;60:520–525.
[10] Larsen L, Scheike T, Jensen TK, et al. The danish first pregnancy planner study team. Computer-assisted semen analysis parameters as predictors for fertility of men from the general population. Hum Reprod. 2000;15:1562–1567.
[11] Garrett C, Liu DY, Clarke GN, et al. Automated semen analysis: 'zona pellucida preferred' sperm morphometry and straight-line velocity are related to pregnancy rate in subfertile couples. Hum Reprod. 2003;18:1643–1649.
[12] Sanders D, Fensome-Rimmer S, Woodward B, et al. Uncertainty of measurement in andrology: UK best practice guideline from the Association of Biomedical Andrologists. Br J Biomed Sci. 2017;74:157–162.
[13] Hancock P, Woodward BJ, Muneer A, et al. Laboratory guidelines for postvasectomy semen analysis: Association of Biomedical Andrologists, the British Andrology Society and the British Association of Urological Surgeons. J Clin Pathol. 2016;69(7):655–660.
[14] Marks D. Guidance on the use of clinical samples for a range of purposes that are not within the remit of research ethic committees [Internet]. [2012; cited 2017 Jul 28]. Available from: https://www.rcpath.org/resourceLibrary/guidance-use-of-clinical-samples-nov-12.html
[15] Cooper TG, Yeung CH. Computer-aided evaluation of assessment of "grade a" spermatozoa by experienced technicians. Fertil Steril. 2006;85:220–224.
[16] Cooper TG, Neuwinger J, Bahrs S, et al. Internal quality control of semen analysis. Fertil Steril. 1992;58(1):172–178.
[17] Neuwinger J, Behre HM, Nieschlag E. External quality control in the andrology laboratory: an experimental multicenter trial. Fertil Steril. 1990;54(2):308–314.
[18] Nijs M, Franssen K, Cox A, et al. Reprotoxicity of intrauterine insemination and *in vitro* fertilization-embryo transfer

disposables and products: a 4-year survey. Fertil Steril. 2009;92(2):527–535.
[19] Critchlow JD, Matson PL, Newman MC, et al. Quality control in an *in vitro* fertilization laboratory: use of human sperm survival studies. Hum Reprod. 1989;4(5):545–549.

## Technique 1. Preparation of tubes for density gradient centrifugation

(1) Take two 15 ml conical centrifuge tubes. Optimally these should be pre-tested and accepted as non-toxic to sperm.

Allocate one tube for the DGC media and one tube for the wash buffer.

(2) Allow all reagents (DGC media and wash buffer) to reach room temperature.

(3) To prepare the DGC media tube, use a pipette to add 1.0 ml of the lower layer (80% v/v). Using a new pipette, slowly dispense 1.0 ml of the upper layer (40% v/v) on top of the lower layer. Take care not to expel too quickly or cause mixing of solutions. Two distinct layers should be seen.

(4) To prepare the wash buffer tube, use a sterile pipette to add 5 ml wash buffer 5 ml.

Note, suggested volumes for the DGC layers and wash buffer have been provided, but this can be adjusted as required.

Note, this instruction is based on a 40/80% (v/v) pre-prepared DGC kit and a suitable buffer solution for washing the sperm. This is provided as an example of how to set up a DGC only. Andrologists are advised to seek advice from those who regularly perform sperm preparation, as variations exist depending on different manufacturer instructions. Any variation from manufacturer instructions should be suitably validated or verified.

## Technique 2. Processing the sperm via density gradient centrifugation

(1) Carefully overlay 1.0 ml liquefied semen onto the density gradient upper layer.

(2) Centrifuge this tube at 300 *g* for 20 min, ensuring the centrifuge is balanced.

(3) Gently remove the tube from the centrifuge. A sperm pellet should be visible at the base of the tube. Using a sterile pipette, remove most of the supernatant, aspirating from the top of the layers, and discard this supernatant. Only the visible pellet should remain.

(4) Using a fresh sterile pipette, transfer the pellet into the wash buffer tube. Resuspend this pellet by gently agitating the media in the tube.

(5) Centrifuge the wash buffer tube at 300 *g* for 5 min. A sperm pellet should then be visible at the base of the tube. Using a sterile pipette, remove most of the supernatant, aspirating from the top of the layers, and discard this supernatant. Only the visible pellet should remain in approximately 0.25 ml.

(6) Resuspend pellet in wash buffer by gently agitating the tube.

(7) Dilute this concentrated sperm solution with additional wash buffer as appropriate, e.g. if there are several consumables to test then dilute to a larger volume. However, it is recommended that a concentration of at least 10 million sperm/ml is available for each STT.

Figure 2 accompanies this instruction.

## Technique 3. Processing the sperm via swim up

(1) Label two 15 ml conical centrifuge tubes. Use a sterile pipette to transfer 1.0 ml of semen to Tube #1.

(2) Carefully layer 1.5 ml of wash buffer over the semen using a new pipette

(3) Without mixing the layers, place the tube into an incubator set at 37 °C for 60 min. The surface area can be increased if the tubes are incubated at a 45° angle although this should not be adjusted during the incubation period or when harvesting.

(4) After 60 min, use a sterile pipette to carefully remove 0.5–1 ml of the upper layer. This should contain motile sperm.

(5) Transfer the harvested media into the second wash buffer tube.

(6) Centrifuge at 300 *g* for 10 min without using the break. A sperm pellet should then be visible at the base of the tube. Using a sterile pipette, remove most of the supernatant, aspirating from the top of the layers, and discard this supernatant. Only the visible pellet should remain in approximately 0.25 ml.

(7) Resuspend pellet in wash buffer by gently agitating the tube.

(8) Dilute this concentrated sperm solution with additional wash buffer as appropriate.

Note, conical tubes are advised for this procedure, which has been adapted from the WHO 2010 guidelines (1).
Figure 3 accompanies this instruction.

## Technique 4. The procedure for performing a sperm toxicity test

(1) Using the enriched population of prepared motile sperm, perform a baseline motility test at time $t = 0$.

(2) Depending on satisfactory results from baseline data, the preparation or centrifuge tube containing the sperm may be used as the control tube.

(3) Add the sperm suspension to each item to be tested (see Table 2).

(4) Perform motility assessment at appropriate intervals in line with likely routine exposure on the test sample and the control (Table 1 is given as a guide to testing intervals and is intended to suitably assess the risk of toxicity in line with normal procedures).

(5) At the required interval, remove an aliquot to assess motility (or use control).

(6) All points of '0' are to be regarded as the control baseline measurement. The other points are to be undertaken on the testing consumables suspension and the control.

Note, all motility assessments should grade minimum of 200 sperm in duplicate and at 37 °C, using consumables that have previously been toxicity tested and passed as acceptable. Consider stating acceptable sample preparation criteria for acceptance, depending on what you expect to test, i.e. very low numbers of sperm in the suspension may require you to re-prepare this if you need multiple aliquots. For STT, motility assessment should be made as easy as possible, i.e. numbers suitable enough to undertake 2 × 200 counts.