Uncertainty of Measurement in semen analysis – information for users

EDITION No	9
DATE OF ISSUE	25.05.2022
REVIEW INTERVAL	Two yearly
AUTHOR	A Allan
OWNER	R Gregoire
LOCATION OF COPIES	Q-Pulse

Document review history		
Reviewed date	**Reviewed by**	**Version**
28.05.2020	A Allan	1. Period of review altered following UKAS assessment 2. P/VAS limits of detection corrected from 250 sperm per ml to 500. 3. Addition to each section of 'quantification of uncertainty' detailing conditions to which re-assessment of the parameter would occur. 4. Addition to each section of 'quantification of uncertainty' of dates in which re-assessment last occurred
03/12/2020	A Allan	Conditions to which re-assessment of the parameter has been met. So, 1. Recalculation of between-observer variability for motility and concentration 2. Recalculation of within-observer variability for motility and concentration 3. Recalculation of lower limits of uncertainty, including introduction of large volume fixed depth slides
21.07.2021	S Brooks	6.1 updated to detail the review of monthly means at AMR to assess if review of UoM data is required.
24/05/2022	A Allan	Minor grammatical corrections.

CONTENTS

1. WHAT IS UNCERTAINTY?

In ordinary use, the word 'uncertainty' does not inspire confidence. However, when used in a technical sense as in 'measurement of uncertainty' or 'uncertainty of a test result' it carries a specific meaning and tells us something about the quality of the result, and helps describe the doubt that exists about the result of any measurement. In everyday speak, uncertainty might be expressed as ‘give or take’, e.g. a stick might be two metres long, give or take a centimetre.

Clinicians and scientists are generally comfortable with the concept of uncertainty in relation to a blood test to determine for example a hormone level, but of course, a semen analysis comprises a combination of different test results. As such it is important to consider the measurement of uncertainty in relation to semen analysis testing and the mechanisms that are in place to attempt to minimize uncertainty measurement when assessing semen samples.

2. EXPRESSING UNCERTAINTY MEASUREMENT

Since there is always a margin of doubt about any measurement, we need ask ‘How big is the margin?’ and ‘How bad is the doubt?’. This is conventionally achieved by considering the width of the margin and the confidence level (which states how sure we are that the ‘true value’ is within that margin).

It is important to realise that uncertainty is not the same as error. Error is the difference between the measured value and the ‘true value’ of the thing being measured, whereas uncertainty is the quantification of the doubt about the measurement result. Any error whose value we do not know is a source of uncertainty.

It is also important to consider that a semen analysis result is often described as for example, ‘normal’, ‘abnormal’ or ‘subnormal; with these categorisations being based on one or more results of tests performed within the analysis. Whilst it is possible to quantify the measurement uncertainty of some of the individual quantitative elements of a semen analysis, it is not possible to derive the uncertainty measurement of the semen analysis result as a whole due to several unavoidable and uncontrollable confounding factors described below.

3. WHY IS IT IMPORTANT TO CONSIDER UNCERTAINTY?

Uncertainty is a (usually quantitative) indication of the quality of the result. It gives an answer to the question, how well does the result represent the value of the quantity being measured? It allows users of the result to assess its reliability, for example for the purposes of comparison of results from different sources or with reference values or ranges.

In the case of a semen analysis, a result is often compared to a reference range. In this case, knowledge of the uncertainty shows whether the result is well within the reference range or only just makes it. Sometimes a result is so close to the limits of the reference range that the risk associated with the possibility that the measured parameter may not fall within the limit, once the uncertainty has been allowed for, must be considered. In other words, results which fall just outside the normal range may in reality be within the normal range.

We often encounter patients who get differing results (and therefore sometimes conflicting advice) when they have semen analyses performed in more than one laboratory. Whilst this is, more often than not, due to the inherent within-individual variability of semen samples a further complication

may be arising. To illustrate this point imagine (although it is of course impossible) a patient were able to have the same sample analysed simultaneously in two Andrology labs. Would we expect the laboratories to get identical results? Only within limits, we may answer, but when the results are close to the specification limit it may be that one laboratory indicates 'normality' whereas another indicates an 'abnormality'. From time to time accreditation bodies have to investigate complaints concerning such differences. This can involve much time and effort for all parties, which in many cases could have been avoided if the uncertainty of the result had been known by the service user.

4. WHERE DO UNCERTAINTIES IN SEMEN ANALYSIS COME FROM?

Many things can undermine a measurement of a semen analysis parameter and importantly these flaws in the measurement may be visible or invisible. Although patients and Andrologists do their best, the nature of semen analyses dictate that they are rarely performed under absolutely perfect conditions and as such, errors and uncertainties can arise from the areas detailed in the table below. However, in some areas it is possible to attempt to control for and minimise these errors and uncertainties and the ways in which we attempt to do this is also included below.

Source of error or uncertainty	Control/minimisation methods
The laboratory equipment used to perform measurements. Measuring instruments (pipettes, counting chambers etc.) can suffer from errors including bias, changes due to ageing, wear, or other kinds of drift, poor readability, noise (for electrical instruments) and many other problems. Please note that the uncertainty of measurement generated by the pipettes used during semen analysis pales into insignificance compared with the other sources of uncertainty described in sections below	• Formal installation and validation • Regular maintenance • Calibration
The patient It is well recognised that the 'quality' of semen samples produced can vary hugely for a variety of reasons, not least of which is normal biological variation. As such it is foolhardy to base a diagnosis on only one semen analysis.	• Performance of repeat semen analyses to help derive a 'representative' diagnosis
The semen sample itself Human semen is a heterogeneous fluid which undergoes a process of liquefaction	• The Andrology Laboratory examines the sample within 60 minutes of it being produced wherever possible

shortly after ejaculation. The constituents of seminal plasma are not capable of sustaining sperm motility and viability over prolonged periods	• Semen samples are well mixed before aliquots are removed for assessment purposes • Awareness that sampling a non-liquefied sample may lead to an erroneous result
Semen sample collection The way in which a semen sample is collected can hugely affect its quality. • Duration of abstinence • Collection method • Collection vessel • Incomplete collection • Exposure to adverse temperature • Ejaculation to analysis interval	• Patients are advised to abstain from ejaculation for a minimum of two and a maximum of seven days. • Patients are advised to collect their samples by masturbation • Patients are advised to only use the container provided by the Andrology Laboratory • Patients are advised to inform the Andrology Laboratory if any of the sample was spilled. • Patients are advised to protect the sample form extremes of temperature • Patients who produce their sample off-site are advised to deliver the sample to the Andrology laboratory within 60 minutes of it being produced
Imported uncertainties Calibration of for example, pipettes or heated-stages will have an inherent uncertainty which is then built into the uncertainty the measurement being made. NB. The uncertainty due to not calibrating equipment would obviously be much worse!	• It is not possible to control for this *per se* although it is essential that equipment is regularly calibrated.
Operator skill and judgment Some measurements (eg assessment of sperm motility by eye) depend upon the skill and judgment of the person looking down the microscope. For example a sperm is deemed to be progressing rapidly (ie grade A) if it is moving >25µm/sec which equates approximately to 5 x the length of a sperm. Such an assessment is highly	• Training • IQC • EQA • Use of computer-aided semen analysis

subjective! Similarly, the human eye is unavoidably drawn to moving objects and as such is inclined to overestimate sperm motility	
Sampling issues The measurements that are made relating to a particular semen sample must be properly representative of the semen sample itself. This is particularly relevant as human semen is a heterogeneous fluid which undergoes a process of liquefaction shortly after ejaculation.	• Semen samples are well mixed before aliquots are removed for assessment purposes • Awareness that sampling a non-liquefied sample may lead to an erroneous result
The environment Temperature, air pressure, humidity and many other conditions can affect the measuring instrument or indeed, the sample being measured	• Patients are advised to protect the sample from extremes of temperature • Patients are advised to only use the container provided by the Andrology Laboratory • All motility assessments are performed at 37°C

So, we can see from the table above that there are many very real issues which may cause uncertainty of measurement in relation to a semen analysis and although we can do our best to control for these, many of the control methods listed above rely heavily on patient compliance.

5. PERIOD OF REVIEW OF UNCERTAINTY WITHIN ANDROLOGY

This document is reviewed every two years as per QMS review schedule. However, this review date may be sooner if new information pertaining to uncertainty and its relevance to Andrology becomes available – for instance,

I. New guidance in the literature
II. Whenever a new process is introduced
III. Whenever a significant change to an existing process or procedure is introduced
IV. A new product is implemented
V. A significant change in staffing personnel

If there is no change to the above parameters then uncertainty of measurement limits described in this SOP will not be reviewed. The lack of review will be justified with each document review.

6. QUANTIFICATION OF UNCERTAINTY

This section seeks, where possible, to quantify the uncertainty of measurement for individual parameters within a semen sample. This allows individual parameters within the analysis to be appropriately interpreted.

At the end of each section, we have attempted to relate the data to clinical practice by suggesting some **points for consideration**, which may help your interpretation.

6.1 Within-patient variation

To demonstrate just how much the quality of semen samples can vary within the same individual, the table below shows data relating to 20 donors who provided a total of 754 ejaculates (minimum of 10 each). The parameter means for all 20 donors are shown together with the mean CV, with the latter representing the between-sample variation for each donor.

	Volume	Concentration	Progression (a+b)	Total Count
Mean (Range)	3.3 (1.1-6.5)	74.8 (43.5-136)	56.0 (44.5-68.5)	227.5 (85.3-557.1)
Mean CV between sample (Range)	26.1 (15.3-42.3)	33.2 (13.6-49.6)	15.2 (10.1-21.5)	43.9 (20.7-68.4)

Points for consideration - clearly, men will produce samples of very variable quality. Diagnosis of sub-normality should not be based on a single semen sample.

NB. These data were calculated in 2014. Due to the absence of a frequent change in the local population, these data will only be reassessed following the assessment of monthly mean data at the Annual Management review. If deemed a shift in population data these experiments will be repeated. This gap will not exceed 10 years.

6.2 Volume measurement

Semen volume is measured by weight. Uncertainty related to the measurement of semen sample volume is very small as demonstrated below where the weight of the same semen sample + pot was determined 10 times. *(June 2017)*

Mean weight (gms)	13.871
Range	13.870-13.872
CV	0.006

Clearly, the same sample measured repeatedly gives almost exactly the same result.

Points for consideration – the reported volume of a semen sample is extremely reliable.

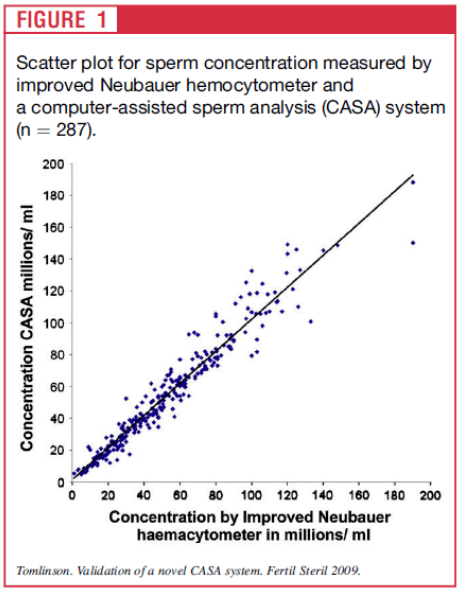
NB. Re-assessment of this measure of uncertainty will only be required if the sample container type or the manufacturing process of the sample container changes. Otherwise, re-assessment is not required. These measurements have not been reassessed in this version of the document as the product, manufacturer and supplier have not changed since June

2017. If this does change however, the assessments will be repeated prior to the new product implementation into the laboratory.

6.3 Concentration

6.3.1 Manual vs computer-aided concentration measurements

Figure 1 below shows the correlation between CASA and 'manual' sperm concentration measurements for 287 semen samples.



These data clearly show that the Sperminator CASA machine can measure sperm concentration at least as well as the conventional haemocytometer method (Tomlinson 2014)

Points for consideration – whenever possible concentration measurements are performed using a CASA machine.

6.3.2 Measurement of sperm concentration using CASA

Although the Hewitt Fertility Centres use CASA machines wherever possible, measurement of uncertainty still exists. The table shows the results of 10 CASA concentration measurements performed on the same aliquot of 10 different samples.

Variability observed when using CASA to measure sperm concentration on the same slide 10 times for 10 different samples (June 2017)

n=10 x 10	
CV	2.85-26.84

The CV shown above is quite high. However, when we take repeated aliquots of the same semen sample (see table below) the CV is much larger due to the sampling variability associated with sampling a heterogeneous fluid such as human semen.

Variability observed when using CASA to measure sperm concentration on 10 aliquots of the same sample (June 2017)

n=10	
Mean concentration (millions/ml)	51.5
Range	48-57
CV	6.02

Points for consideration – the uncertainty associated with measuring sperm concentration, even using CASA technology is quite high.

NB. These measurements have not been reassessed in this document version as the CASA software and hardware have not changed since June 2017. If this does change, the assessments will be repeated prior to implementation into the laboratory.

6.3.3 Manual measurement of sperm concentration

Unfortunately, CASA systems require between 8 and 80 million sperm per ml to function optimally. Therefore, the andrology service remains reliant on manual measurement of sperm concentration for semen samples outside these ranges. As such, it is prudent to examine the uncertainty associated with the manual measurements of sperm concentration.

6.3.3.1 Within-observer variability

The table below shows the results of 10 manual concentration measurements performed on the same sample by the same operator.

Variability observed with the same observer performing a manual concentration measurement on the same sample 10 times (June 2020)

n=10	
Mean concentration (millions/ml)	117.2
Range	111-124.5
CV	4.24

The CV is reasonable and similar to that those derived using the CASA system.

NB. These measurements have been reassessed recently as criteria in Section 5 have been met ie staff changes.

6.3.3.2 Between-observer variability

The table below shows the results of 5 operators performing manual concentration measurements on 10 different samples. Please note that the variability seen here may be a combination of true 'between-observer' variability together with sampling error.

*Variability observed with 5 observers performing a manual concentration measurement on the same sample at the same time (*Data completed Dec 2020*)*

		Operator									
		1	2	3	4	5	Mean	Max	Min	SD	CV
Sample Number	1	40	53.5	53	40	42	45.7	53.5	40	6.942622	15.19173
	2	47	56	63	47	63	55.2	63	47	8.01249	14.51538
	3	46	49.4	33	45	52	45.08	52	33	7.302876	16.19981
	4	34	39.2	30	38.5	34	35.14	39.2	30	3.768023	10.72289
	5	28	33	25	31.5	30.1	29.52	33	25	3.126819	10.59221
	6	17	19.4	15	14.6	16.5	16.5	19.4	14.6	1.905256	11.54701
	7	67	80	83	75.6	58.3	72.78	83	58.3	10.09465	13.87009
	8	20	25	19	21.4	14.8	20.04	25	14.8	3.70783	18.50214
	9	31.7	30	30.1	36	29	31.36	36	30	2.768212	8.827208
	10	100.5	92.3	125.5	82	93	98.66	125.5	82	16.38454	16.60707

Once again, the CVs are large.

NB. These measurements have been reassessed recently as criteria in Section 5 have been met ie staff changes.

Points for consideration – the uncertainty associated with manually measuring sperm concentration can be large, particularly at lower sperm concentrations.

6.4 Motility

There are principally four areas in which uncertainty of measurement can be introduced when measuring sperm motility these being

i. the time interval between ejaculation and analysis
ii. the effect of temperature
iii. the effect of the operator or CASA system and
iv. the difference between operators

6.4.1 Time interval between ejaculation and analysis

Sperm motility in some semen samples will start to decline after approximately 50 minutes. As such, the Andrology Laboratory endeavors to perform all motility analyses within 50 minutes of ejaculation.

Points for consideration –Patients who produce samples off-site should be strongly advised to deliver the samples to the Andrology Lab within 50 minutes of ejaculation.

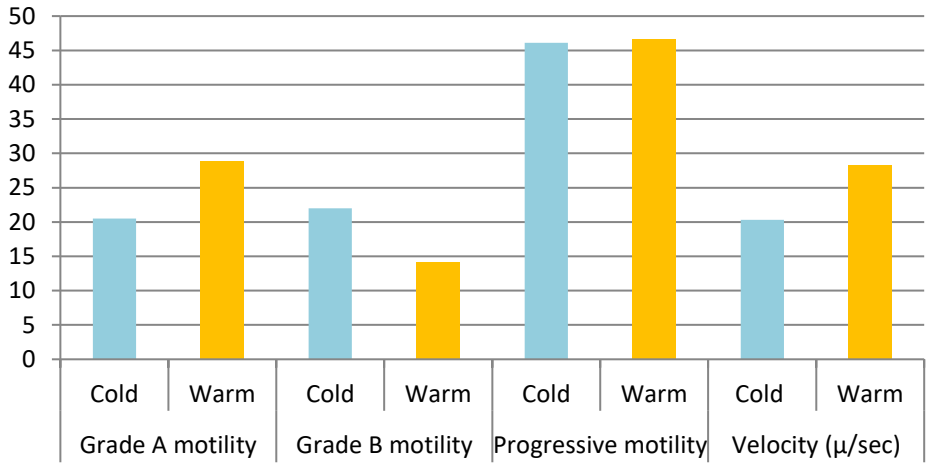
NB. The time interval between ejaculation and analysis will not be reassessed until evidence in the literature, (WHO 2010) or professional guidance suggests to time interval should be reconsidered. If these do change, the assessments will be completed prior to implementation into the laboratory.

Any instance where sperm motility is reduced and the motility assessment was performed over 50 minutes after ejaculation (e.g. where the sample was produced off-site) will be highlighted on the report.

6.4.2 Effect of temperature

Figure 3 below, demonstrates that more sperm swim faster at 37°C than at room temperature.

Figure 3. The effect of temperature on sperm motility
(Cold 19-23°C; warm 36-37°C, n=25)



Points for consideration - All motility analyses performed within the Hewitt Fertility Centre Andrology Laboratories are performed at 37°C.

Failure by patients to follow instructions regarding sample production and transport will be noted on the report.

NB. This has not been reassessed in this document version as evidence in the literature (WHO 2010) outlines no change in guidelines to temperature. If this does change, the assessments will be completed prior to implementation into the laboratory.

6.4.3 Manual vs. computer-aided motility measurements

Error and uncertainty associated with motility measurement is often difficult to demonstrate as the industry lacks a 'gold standard' methodology. Indeed manual motility measurements are almost impossible to validate unless they are compared directly to an objective measurement such as those derived by computer-aided semen analysis (CASA). The table below shows the results for manual vs computerised motility measurements for 100 semen samples.

Manual versus computer derived motility measurements for the same samples (median %, n=100)

Motility grade (median)	Manual	CASA	Correlation coefficient (r)	Significance
Rapid (a)	25	21	0.863	<0.0001
Medium (b)	16	11.5	0.327	<0.001
Slow (c)	8	15	0.168	NS
Static (d)	50	50	0.845	<0.0001
Progressive (a+b)	42	32.5	0.845	<0.0001

These data clearly show that the human eye tends to over-estimate the speed at which sperm are progressing by up to 20% compared to the CASA system.

Points for consideration – whenever possible motility measurements are performed using a CASA machine to remove subjectivity. Please note, however, that CASA machines cannot be used for samples with low sperm concentrations.

NB. This has not been reassessed in this document version as the CASA software and hardware have not changed since June 2017. If this does change, the assessments will be completed prior to implementation into the laboratory.

6.4.4 Measurement of sperm motility using CASA

Although we use CASA machines wherever possible measurement uncertainty still exists. The table below shows the results of 10 CASA motility measurements performed on the same aliquot of 10 different semen samples..

Variability observed when using CASA to measure sperm motility on the same slide 10 times for 10 different samples (June 2017)

Motility grade (n=10 x 10)	**CV**
Rapid (a)	5.8 – 58.7
Medium (b)	8.2 – 29.8
Slow (c)	18.7 – 29.6
Static (d)	6.0 – 39.1

The CVs shown in the table above are high. However, when we take repeated aliquots of the same semen sample (see table below) the CV is higher still due to the sampling variability associated with sampling a heterogeneous fluid such as human semen.

Variability observed when using CASA to measure sperm motility on 10 aliquots of the same sample (June 2017)

Motility grade (n=10)	**Mean**	**Range**	**CV**
Rapid (a)	24	20-29	12.88
Medium (b)	20.9	15-26	16.34
Slow (c)	8	5-10	25
Static (d)	47.1	41-55	10.37

Points for consideration – the uncertainty associated with measuring sperm motility, even using CASA technology is high.

NB. This has not been reassessed in this document version as the CASA software and hardware have not changed since June 2017. If this does change, the assessments will be completed prior to implementation into the laboratory.

6.4.5 Manual measurement of sperm motility

Unfortunately, CASA systems require between 8 and 80 million sperm per ml to function optimally. Therefore, the andrology service remains reliant on manual measurement of sperm concentration for semen samples outside these ranges. As such, it is prudent to examine the uncertainty associated with the manual measurements of sperm motility.

6.4.5.1 Within-observer variability

The table below shows the results of 10 manual motility measurements performed on the same sample by the same operator. Please note i. that some variability may occur due to natural 'deterioration' of sperm motility over the time taken to perform the 10 measurements and ii. The variability seen here will be a combination of true 'within-observer' variability together with sampling error.

*Variability observed with the same observer performing a manual motility assessment on the same sample10 times (*June 2020*)*

Motility grade (n=10)	**Mean**	**Range**	**CV**
Progressive	55.8	48-62	8.31
Non-progressive	6.8	4-9	28.41
Non-motile	37.4	33-44	10.26

Once again very high CVs are observed as with the CASA system. Please note that a significant degree of unwitting bias may be inherent in this experiment as the operator may subconsciously 'adjust' their measurement as they already know previous results!

NB. This has been recently reassessed as criteria in Section 5 have been met ie staff changes.

6.4.5.2 Between-observer variability

The table below shows the results of 5 observers performing manual motility measurements at the same time on 10 samples. Please note that the variability seen here may be a combination of true 'between-observer' variability together with sampling error.

Variability observed with 5 observers performing a manual motility assessment on the same sample at the same time (Data completed Dec 2020)

Motility grade (n=50)	**Mean CV**	**Range CV - Low**	**Range CV - High**
Progressive	14.931467	8.279673061	20.16977663
Non-progressive	44.131173	17.43041722	83.16309141
Immotile	10.705606	3.662480267	19.28079354

Once again, the CVs are very large.

Points for consideration – the uncertainty associated with manually measuring sperm motility is high.

NB. This has been recently reassessed as criteria in Section 5 have been met ie staff changes.

6.4.6 Effect of numbers counted when using CASA

The table below shows the difference in CV obtained when estimating sperm concentration and motility by asking the CASA machine to count either 200 or 400 sperm for 5 different samples measured 10 times each.

Motility grade (n= 5x10)	Mean CV for counting 200 sperm	Mean CV for counting 400 sperm
Concentration	11.5	7.75
Rapid (a)	22.55	16.47
Medium (b)	21.89	19.16
Slow (c)	27.98	27.49
Static (d)	10.68	9.16

The table above shows the CVs obtained when measuring concentration and motility repeatedly (n=50) on the same sample according to whether 200 or 400 sperm were counted. A slight reduction in CV is observed when 400 sperm are counted compared to 200 (as might be expected). However, the CV observed when 400 sperm are counted remains high and as such there is little benefit in counting 400 rather than 200 sperm as its effect on measurement uncertainty is minimal.

NB. Re-assessment of this measure of uncertainty will only be required if the assessment software were to change, a new CASA system is installed or the referring demographic population were found to have *significantly altered. Otherwise, re-assessment is not required.

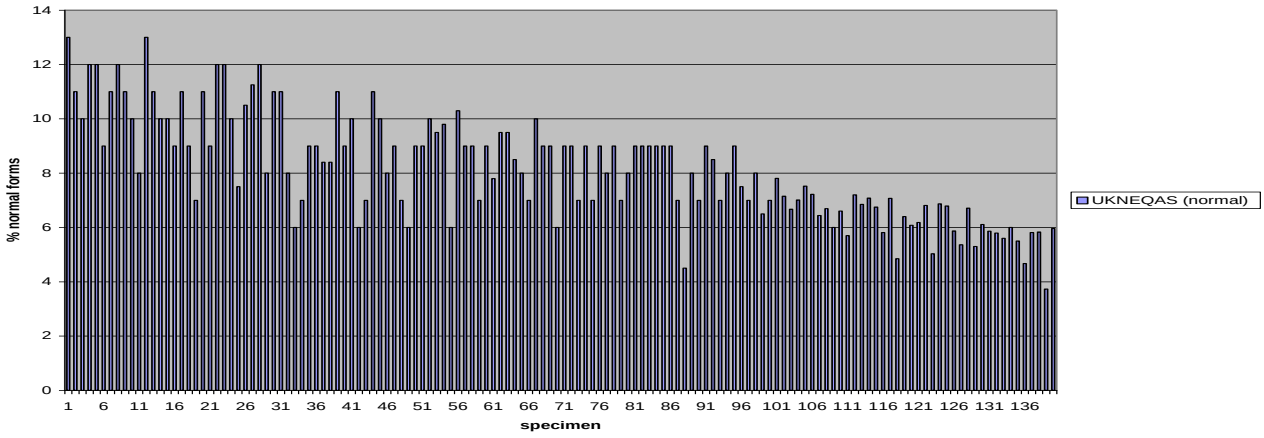
****It is the responsibility of the Person Responsible or Scientific Director to deem when a 'significant' change has occurred.***

6.5 Morphology

The assessment of sperm morphology is fraught with difficulty for many reasons and significant measurement uncertainty exists. Some examples of these difficulties are given below.

The figure below would suggest that a laboratories' perception of 'a normal sperm' is slowly changing to meet the needs of the new reference ranges, despite using the same sperm shape and size criteria to work to. Figure 1 below shows the target values for % normal forms from EQA samples over the past 8 years. There is a clear relationship showing generally stricter scoring with time in response to a gradual adoption of a lower reference range.

EQA target values for % normal forms from 2005-2011 in the UK NEQAS scheme

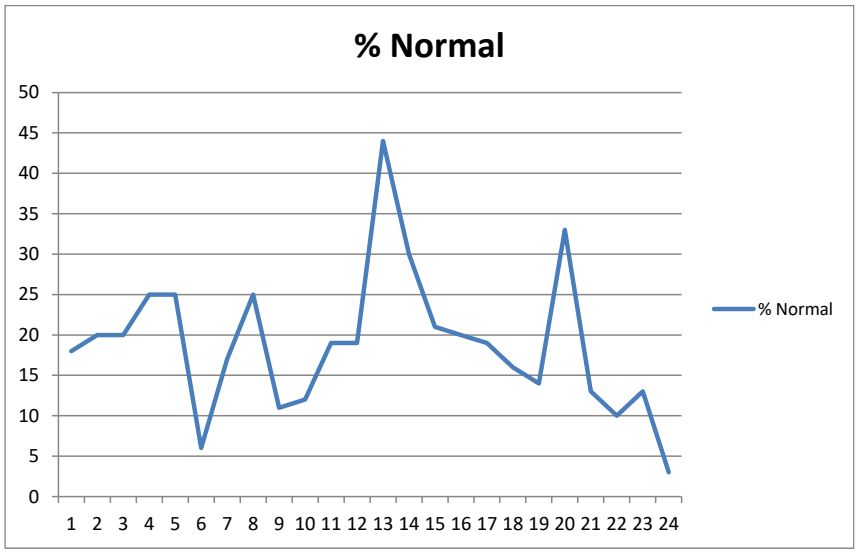


Secondly, Andrologists are routinely trained to follow a philosophy which labels any sperm which does not meet pre-defined size and shape definitions as being abnormal. By definition the group of abnormal forms then includes a significant number of 'unknowns' which could include: borderline forms; artefacts

created by slide preparation; or indeed those which become adhered to artefacts such as debris or non-sperm cells. The consequence of adopting this strategy is that the uncertainty surrounding those 'unknowns' (and therefore for the entire measurement) cannot be assessed.

Thirdly, and to compound the difficulties yet further, not only can differences in fixation and staining make a difference to the overall result but individual interpretation of exactly the same sperm images show a remarkable lack of consistency across a range of operators.

A recent small study using a series of clear images sent out by the laboratory at Nottingham University Hospital to a number of centres showed that even in experienced hands, agreement on whether a sperm is normal or abnormal varied considerably. The figure below shows the % normal forms as reported by 24 individuals (fully trained to perform semen analysis) based in six different laboratories. The mean from 160 sperm images was 18.9% normal with a range from 3% to 44%!



Identification of normal forms from the same set of micrographs assessed by 24 staff in six different centres (Tomlinson 2014)

Clearly there is a very large and unquantifiable uncertainty associated with sperm morphology assessment as currently performed, and it seems that estimating sperm morphology in terms of percentage 'normal forms' is difficult (if not impossible) with subjectivity remaining a significant problem.

However, there are certain situations where the performance of sperm morphology is associated with an extremely low (if not zero) level of uncertainty and these are where the morphological defect applies to every sperm and such conditions are easily recognisable. This might include conditions such as globozoospermia (where the head size is increased and no acrosome is present), pin-head sperm (where the sperm heads are missing) or gross tail defects. Such conditions are often 'sterilising'.

Points for consideration – the high degree of uncertainty around assessing % normal forms raises questions about its clinical value. However, sperm morphology assessment is of considerable value in identifying gross morphological abnormalities.

NB. Re-assessment of this measure of uncertainty will only be required if new evidence in the literature or new guidelines outline a change to sperm morphology assessment.

7. UNCERTAINTY BUDGET

Uncertainty Budget

Type A: which are carried out by calculations from a series of repeated observations using statistical methods	Type B: evaluations where measurements are derived from other sources e.g. This could be information from past experience of the measurements, from calibration certificates,
Calculate SD of Intra precision	Temperature Humidity Inter operator variability Pipette uncertainty values Reagent variability Calibration lot number variability. Inter precision variability

7.1 How to calculate Uncertainty when using an Uncertainty Budget.

Calculate the SD of the intra precision

•SD intra precision = A

•Calculate the SD of the inter precision

•SD inter precision = B

Add the 2 uncertainties, in order to get rid of a - or + Square them, add them and calculate the Square Root

u=$\sqrt{A^2 + B^2}$

Expanded uncertainty

Uncertainty is calculated as 1 SD, 1 SD gives 68% confidence on the Gaussian Curve, it is reasonable to multiply the uncertainty by 2 to attain a confidence level of 95%. This is referred to the coverage factor and is represented by K

7.2 Calculating Expanded Uncertainty.

Uncertainty is expressed as $\mathbf{u=\sqrt{A^2 + B^2}}$

Expanded uncertainty is expressed as **U=2xu**

7.3 INTRODUCTION TO TYPE A & B UNCERTAINTIES AND CONSIDERATIONS

This uncertainty budget seeks to systematically examine and document all sources of measurement uncertainty associated with performing a routine semen analysis. This is done by examining i) readily measurable and statistically describable variability – so called Type A uncertainties, and ii) Type B uncertainties, which are less easily quantified but are known to effect the outcome of the test.

Routine semen analysis is made up of several components (ejaculate volume, sperm concentration, sperm motility and sperm morphology) and such it is prudent to list Type A and Type B uncertainties for each.

Ideally, it should then be possible to mathematically combine all the measurement uncertainties associated with measuring a particular parameter to derive an overall level of uncertainty for measuring that parameter. This document also seeks to examine whether such an approach is applicable to routine semen analysis.

Please note that a fundamental assumption is made at the outset when performing a semen analysis in that the sample being analysed has indeed been provided by the patient. Unlike taking a venous blood sample, the requirement (in all but rare exceptions eg electro-ejaculation or surgical sperm retrieval) for the sample to be produced by masturbation means that its provenance cannot be guaranteed with certainty.

7.3.1 Ejaculate volume (measured by weight)

Type A uncertainties	Type B uncertainties
Volume measurement by weight Imported uncertainty (eg equipment calibration)	Incomplete sample collection - *we rely upon the patient to report sample spillage* Abstinence period - *we rely upon the patient to report abstinence period* Intra-patient variability – *it is well recognised that a patient will produce samples of varying volume* Retrograde ejaculation – *the patient may produce an incomplete sample due to partial retrograde ejaculation*

7.3.2 Sperm concentration (measured manually)

Type A uncertainties	Type B uncertainties
Intra-sample variability Inter and intra-operator variability Inter and intra-chamber variability Imported uncertainty (eg equipment calibration)	Incomplete sample collection - *we rely upon the patient to report sample spillage. The first fraction of the ejaculate contains the majority* Abstinence period - *we rely upon the patient to report abstinence period* Intra-patient variability – *it is well recognised that a patient will produce samples of varying concentration* Retrograde ejaculation – *the patient may produce an incomplete sample due to partial retrograde ejaculation* Sample heterogeneity – *sperm are suspended in seminal fluid which is a viscous, heterogeneous medium. Vigorous sample mixing is not recommended. Furthermore, the sample undergoes an enzymatically driven process of liquefaction which significantly affects sample viscosity and consistency.*

7.3.3 Sperm concentration (measured using CASA)

Type A uncertainties	Type B uncertainties
Intra-sample variability Inter and intra-Sperminator variability Manual vs CASA concentration measurement Imported uncertainty (eg equipment calibration)	Incomplete sample collection - *we rely upon the patient to report sample spillage. The first fraction of the ejaculate contains the majority* Abstinence period - *we rely upon the patient to report abstinence period* Intra-patient variability – *it is well recognised that a patient will produce samples of varying concentration* Retrograde ejaculation – *the patient may produce an incomplete sample due to partial retrograde ejaculation* Sample heterogeneity – *sperm are suspended in seminal fluid which is a viscous, heterogeneous medium. Vigorous sample mixing is not recommended. Furthermore, the sample undergoes an enzymatically driven process of liquefaction which significantly affects sample viscosity and consistency.*

7.3.4 Sperm motility (measured manually)

Type A uncertainties	Type B uncertainties
Intra-sample variability Inter and intra observer variability Effect of temperature Imported uncertainty (eg equipment calibration)	Incomplete sample collection - *we rely upon the patient to report sample spillage. The first fraction of the ejaculate contains the majority* Abstinence period - *we rely upon the patient to report abstinence period* Intra-patient variability – *it is well recognised that a patient will produce samples of varying motility* Retrograde ejaculation – *the patient may produce an incomplete sample due to partial retrograde ejaculation* Sample heterogeneity – *sperm are suspended in seminal fluid which is a viscous, heterogeneous medium. Vigorous sample mixing is not recommended. Furthermore, the sample undergoes an enzymatically driven process of liquefaction which significantly affects sample viscosity and consistency.*

	Sample handling - *we rely on the patient to protect the sample from extremes of temperature from production to delivery to the laboratory* Ejaculation to analysis interval – *we rely on the patient to accurately record the time that the sample was produced to facilitate its analysis within 60 mins of ejaculation* *Slide and room temperature - whilst we attempt to control the temperature of the analysis itself by using heated stages the effects of ambient temperature should be recognized* Presence of antisperm anitbodies and associated effects.

7.3.5 Sperm motility (measured using CASA)

Type A uncertainties	Type B uncertainties
Intra-sample variability Inter and intra-Sperminator variability Manual vs CASA motility measurement Effect of temperature Imported uncertainty (eg equipment calibration)	Incomplete sample collection - *we rely upon the patient to report sample spillage. The first fraction of the ejaculate contains the majority* Abstinence period - *we rely upon the patient to report abstinence period* Intra-patient variability – *it is well recognised that a patient will produce samples of varying motility* Retrograde ejaculation – *the patient may produce an incomplete sample due to partial retrograde ejaculation* Sample heterogeneity – *sperm are suspended in seminal fluid which is a viscous, heterogeneous medium. Vigorous sample mixing is not recommended. Furthermore, the sample undergoes an enzymatically driven process of liquefaction which significantly affects sample viscosity and consistency.* Sample handling - *we rely on the patient to protect the sample from extremes of temperature from production to delivery to the laboratory* Ejaculation to analysis interval – *we rely on the patient to accurately record the time that the sample was produced to facilitate its analysis within 60 mins of ejaculation*

	Slide and room temperature - *whilst we attempt to control the temperature of the analysis itself by using heated stages the effects of ambient temperature should be recognized* Presence of antisperm anitbodies and associated effects.

7.3.6 Sperm morphology

Type A uncertainties	Type B uncertainties
Intra and inter-sample variability Intra and inter-observer variability	Incomplete sample collection - *we rely upon the patient to report sample spillage. The first fraction of the ejaculate contains the majority* Abstinence period - *we rely upon the patient to report abstinence period* Intra-patient variability – *it is well recognised that a patient will produce samples of varying morphology* Retrograde ejaculation – *the patient may produce an incomplete sample due to partial retrograde ejaculation* Sample heterogeneity – *sperm are suspended in seminal fluid which is a viscous, heterogeneous medium. Vigorous sample mixing is not recommended. Furthermore, the sample undergoes an enzymatically driven process of liquefaction which significantly affects sample viscosity and consistency.* Changing perception of a morphologically 'normal' sperm to meet defined reference ranges. Accepted approach that 'unknown' = 'abnormal' Fixation artefacts.

8. POST VASECTOMY ANALYSIS

Further considerations must be made with regards to post vasectomy semen analysis (PVSA). The numbers of sperm that are detectable in these particular samples are extremely low. Critical levels for the clinician are 100,000 sperm/ ml and no sperm in the ejaculate. Therefore the accuracy of numbers of sperm present should be shown with respect of measurement uncertainty.

The Neubauer haemocytometers are shown to have accuracy above the number of 56,000 sperm per ml and limitations regarding this are discussed above.

With regards to the limitations of technicians stating 'no sperm seen in ejaculate', tests were performed to determine our actual detection limits. Serial dilutions of Accubeads were formulated.

The table below shows the levels whereby Accubeads have been detected on a 10ul wet prep and secondly, with large volume fixed depth slides (our two tests in which technicians may report stating 'no sperm seen in ejaculate'), and the corresponding concentration for a number of staff in the department. This has determined our limits of detection. For instance, if we say there is no sperm in the ejaculate; it actually means that there are less than 500 sperm /ml, which is acceptable.

Concentration Sperm/ml	Staff member	1	2	3	4	5	6
A	≈ 62						
B	≈ 125						
C	≈ 250	X				X	
D	≈ 500	X	X	X	X	X	X
E	≈ 1,000	X	X	X	X	X	X
F	≈ 2,000	X	X	X	X	X	X
G	≈ 4,000	X	X	X	X	X	X
H	≈ 8,000	X	X	X	X	X	X
I	≈15,625	X	X	X	X	X	X
J	≈ 31,250	X	X	X	X	X	X
K	≈62,500	X	X	X	X	X	X
L	≈125,000	X	X	X	X	X	X
M	≈ 250,000	X	X	X	X	X	X
N	≈ 500,000	X	X	X	X	X	X
O	≈1M	X	X	X	X	X	X

*Variability observed with 6 observers determining lower detection limits of sperm per ml on a 10ul wet preparation. (*Completed October 2020*)*

X indicates sperm identified

Concentration Sperm/ml	Staff member	1	2	3	4	5	6
A	≈ 62						
B	≈ 125						
C	≈ 250	X	X	X			
D	≈ 500	X	X	X	X	X	X
E	≈ 1,000	X	X	X	X	X	X
F	≈ 2,000	X	X	X	X	X	X
G	≈ 4,000	X	X	X	X	X	X
H	≈ 8,000	X	X	X	X	X	X
I	≈15,625	X	X	X	X	X	X
J	≈ 31,250	X	X	X	X	X	X
K	≈62,500	X	X	X	X	X	X
L	≈125,000	X	X	X	X	X	X
M	≈ 250,000	X	X	X	X	X	X
N	≈ 500,000	X	X	X	X	X	X
O	≈1M	X	X	X	X	X	X

*Variability observed with 6 observers determining lower detection limits of sperm per ml on a large volume fixed depth slide. (*Completed October 2020*)*

X indicates sperm identified

Therefore, the limits of detection for post vasectomy analysis are less than 500 sperm /ml when assessing samples using either the 'coverslip and centrifugation' method or the large volume fixed depth slides method.

NB. These measurements have been reassessed recently as criteria in Section 5 have been met ie staff changes and a new product implemented.

9. CONCLUSIONS

Sections above describe that it is indeed possible to derive some statistical descriptors of some of the parameters routinely measured as part of a semen analysis (ie the Type A uncertainties).

However, what is overwhelmingly clear is that the Type B uncertainties are predominant and of such scale and relevance that the determination of an uncertainty budget for a routine semen analysis (or indeed its individual components) is meaningless.

Nevertheless, a semen analysis remains the only first-line test available to assess male fertility. As such it is incumbent upon those performing semen analyses to provide information to users about the overall reliability of the test result and to assist them in their interpretation particularly in relation to the uncertainties that might be associated with the result.

Whilst this document shows that it is not possible to quantify (in conventional terms) the measurement uncertainty associated with a semen analysis it does not lessen the importance of drawing to the attention of service users to where uncertainties in semen analysis arise, and perhaps

most importantly, how these might be reduced. The Hewitt Fertility Centre Andrology Laboratories seek to achieve this through the 'Measurement of uncertainty in semen analysis – information for users' document.

10. AS A USER OF THE DIAGNOSTIC ANDROLOGY SERVICE – WHAT DO I NEED TO DO?

The simple answer to this question is nothing. The way in which semen analysis testing is performed and the inherent problems and difficulties therein remain unchanged. Similarly, the uncertainty measurement associated with performing a routine semen analysis has, and will always be present to a greater or lesser degree.

As the provider of your semen analysis testing we would simply ask, having taken the time to read this document, that you consider its content when interpreting a semen analysis result within the clinical environment.

10.1 Summary of considerations & advice

For your convenience, we have provided below a summary which seeks to draw together the main points when considering measurement uncertainty and semen analysis.

- **It is essential that patients be strongly advised to follow instructions regarding sample collection and abstinence to reduce the uncertainty that this can introduce.**
- **An interval of more than an hour between ejaculation and analysis may lead to a reduction in sperm motility – this will be highlighted on the report.**
- **Results from samples which are not fully liquefied may not be truly representative of the sample's quality – this will be highlighted on the report.**
- **Men will produce samples of very variable quality due to normal biological variation. As such, a diagnosis of sub-normality should not be made on a single semen sample.**
- **The measurement of sperm concentration (either manually or by computer) is associated with a high degree of measurement uncertainty and this should be taken into account when interpreting semen analysis results, particularly at the limits of normality.**
- **The measurement of sperm motility (either manually or by computer) is associated with a very high degree of measurement uncertainty and this should be taken into account when interpreting semen analysis results, particularly at the limits of normality.**
- **The measurement of sperm morphology is associated with a very high degree of measurement uncertainty and this should be taken into account when interpreting semen analysis results. Measurement of sperm morphology is of considerable value in identifying gross morphological abnormalities.**

11. ACKNOWLEDGEMENTS

The Hewitt Fertility Centre is grateful to Dr Mathew Tomlinson (Consultant Reproductive Scientist) from the Fertility Unit at Nottingham University Hospital for kindly providing some of the data used within this document and for his intellectual input.