Asian Journal of Andrology (2019) 21, 438–444
www.asiaandro.com; www.ajandrology.com

Open Access

ORIGINAL ARTICLE

Sperm Biology

Critical evaluation of two models of flow cytometers for the assessment of sperm DNA fragmentation: an appeal for performance verification

Rakesh Sharma[1,*], Sajal Gupta[1,*], Ralf Henkel[1,2], Ashok Agarwal[1]

Lack of standardized, reproducible protocols and reference values is among the challenges faced when using new or upgraded versions of instruments in reproductive laboratories and flow cytometry. Terminal deoxynucleotidyl transferase dUTP nick end labeling (TUNEL) assay combined with flow cytometry routinely used for diagnostic measurement of sperm DNA fragmentation (SDF) is a unique example. Any change in the setting of the standard instrument, including upgrades of hardware or software, can lead to different results and may affect clinicians' decision for treatment. Therefore, we compared TUNEL results of SDF obtained from a standard (C6) flow cytometer with a newer version of the same instrument (C6 Plus) and examined the cutoff, sensitivity, and specificity without calibration (adjustment) and after adjustment. Identical sperm preparation and matched acquisition settings were used to examine the performance of two flow cytometers. The strength of agreement of the results between the two observers was also assessed. After adjustment of the settings, overall concordance became high and the two cytometers showed 100% positive and negative predictive value with 100% area under the curve. The overall correlation coefficient observed between C6 and C6 Plus was highly significant ($P < 0.0001$; $r = 0.992$; 95% confidence interval [CI]: 0.982–0.997). After adjustment, the two cytometers showed very high precision of 98% and accuracy of >99%. The interobserver agreement on C6 flow cytometer for the two observers was 0.801 ± 0.062 and 0.746 ± 0.044 for C6 Plus. We demonstrated a strong agreement between the samples tested on the two flow cytometers after calibration and established the robustness of both instruments.

Asian Journal of Andrology (2019) **21**, 438–444; doi: 10.4103/aja.aja_109_18; published online: 8 January 2019

Keywords: DNA fragmentation; flow cytometer; oxidative stress; sperm; TUNEL assay

INTRODUCTION

Male factor accounts for approximately 20% of cases of male infertility.[1] Sperm DNA integrity testing has been proposed to be a test with promising potential to complement standard semen analysis.[2] Sperm DNA fragmentation (SDF) has an adverse effect on fertility and causes failure to conceive,[3] longer time to pregnancy,[4] poor outcome following stimulated intrauterine insemination,[5,6] impaired embryo development,[7] higher miscarriage rates,[8] and increased pregnancy loss after both *in vitro* fertilization (IVF) and intracytoplasmic sperm injection.[9] Furthermore, it may have far-reaching consequences for reproductive outcome. Although all SDF tests currently suffer from the common pitfall that the nature and type of DNA damage are unclear, numerous studies have illustrated the prognostic value of SDF tests irrespective of the testing method used.[10,11] In spite of the above data, separate reports from the American Society for Reproductive Medicine Practice Committee (2008),[12] the European Society for Human Reproduction and Embryology,[13] and the British Fertility Society[14] have all concluded that, at the present time, there is insufficient evidence for sperm DNA testing to be introduced as part of clinical laboratory work-up with the need for further research being identified.

Laboratories having large testing volumes for male infertility and IVF clinics may acquire more than one instrument from the same vendor but of different models or with upgraded software. Although the instrument may be certified by the manufacturer to perform as per specifications listed, very rarely is diagnostic equipment tested for the comparability in their performance with clinical testing guidelines in Andrology and IVF laboratory setting. Differences in the results obtained with various models of the same instrument may differentially categorize a patient as normal or abnormal because of changes in the technical settings of the instrument. There are no guidelines from the "Clinical Laboratory Improvement Amendments of 1988, American College of Pathologists, or Joint Commission on Accreditation Health Care Organization or State Licensing Agencies" on the performance verification following hardware or software upgrades of instruments used in the clinical laboratories for diagnostic purposes. Therefore, the accuracy of the results, even if the same methodology is used, should be established by comparing the performance of two different instruments under identical assay conditions to minimize interassay variability. Before using a new instrument in the clinical laboratory, the concordance of two different models of the same instrument needs

[1]American Center for Reproductive Medicine, Cleveland Clinic, Cleveland, OH 44195, USA; [2]Department of Medical Bioscience, University of the Western Cape, Bellville 7535, South Africa.
*These authors contributed equally to this work.
Correspondence: Dr. A Agarwal (agarwaa@ccf.org)
Received: 12 May 2018; Accepted: 21 October 2018

to be established by performance verification of the new model versus the standard model.

Terminal deoxynucleotidyl transferase deoxyuridine triphosphate nucleotide (dUTP) nick end labeling (TUNEL) is a promising assay for the quantification of SDF. It is a direct test that measures both single and double DNA-strand breaks through the enzymatic incorporation of dUTP at the 3'-OH terminal end of a DNA strand.[15] There are several kits available for the detection of DNA fragmentation, and all of them are capable of quantifying DNA damage in spermatozoa.[16,17] None of these kits were originally developed for measuring DNA fragmentation in the spermatozoa. The combination of propidium iodide (PI) with TUNEL flow cytometry has made this test very promising. However, the threshold values correlating with male fertility vary in the threshold and range from 20% to 35%.[18–20] In an earlier study, we reported the standardization of the TUNEL protocol on identical C6 bench-top flow cytometers across two laboratories on different continents.[21] This study was conducted under strictly identical controlled conditions where we demonstrated the precision and the accuracy of the TUNEL assay across laboratories. Recently, a new model of the bench-top flow cytometer called C6 Plus was introduced in the market with some improvements over the current C6 flow cytometer model. Some of the features of the BD Accuri C6 Plus (BD Bioscience, San Diego, CA, USA) include (1) redesigned red laser that is securely anchored to the instrument to minimize misalignment of the laser, (2) the fluidics that are enhanced and more sensitive in C6 Plus versus C6 to any issues related to the sheath fluid, (3) upgraded software that is compatible with 64-bit processor, and (4) built-in software with addition of the cyto screening and tracking CS&T RUO beads (BD Bioscience) which are dyed with fluorochromes that are excited by the cytometer's lasers. These beads are standardized to ensure instrument performance is consistent. The beads automatically characterize, track, and report the performance measurement of the C6 Plus flow cytometer. Therefore, the goals of this study were to (1) compare the concordance correlation, sensitivity, and specificity of the SDF results from standard (C6) flow cytometer with a newer version (C6 Plus) of the same instrument with unadjusted setting and after adjustment with the standard C6, (2) establish the analytical validity of a new model of flow cytometer utilizing the current flow cytometer as the standard, (3) compare the inter- and intra-observer variability on both flow cytometers, and (4) establish the agreement between two observers: precision and accuracy of the newer model of the flow cytometer.

PARTICIPANTS AND METHODS

Participants and study design

Following approval from the Cleveland Clinic institutional review board (Cleveland, OH, USA), four healthy male volunteers (donors) of proven and unproven fertility were selected on the basis of normal semen quality according to the guidelines of the World Health Organization (WHO) 5th edition.[22] In addition, semen samples from 14 infertile men were also included for preparing the spermatozoa for SDF testing. Both donors and patients provided several samples, and informed consent was obtained from each participant. All samples were prepared under identical conditions (sample preparation, staining, and analysis). The comparison of the unadjusted and adjusted settings of C6 Plus flow cytometer for the four aims of the study is shown in **Supplementary Figure 1**.

Semen specimens were collected by masturbation and 6 µl of a well-liquefied aliquot was loaded on a 20-µm depth Leija cell chamber (Spectrum Technologies, Healdsburg, CA, USA) for the evaluation of manual sperm concentration by following the WHO 5th edition criteria.[22] An aliquot containing 2.5×10^6 sperm per ml was used to assess SDF using the TUNEL assay as described below.

Preparation of samples and SDF measurement by TUNEL assay

Sample preparation was performed by washing 2.5×10^6 spermatozoa twice in phosphate-buffered saline (PBS; Irvine Scientific, Santa Ana, CA, USA). Samples were centrifuged at 300*g* (Eppendorf North America, Hauppauge, NY, USA) to remove seminal plasma and re-suspended in 3.7% (*w/v*) paraformaldehyde (PFA; Sigma Aldrich, St. Louis, MO, USA) prepared in PBS.

Spermatozoa (2.5×10^6 ml^{-1}) were aliquoted and prepared into negative and positive controls. Negative spermatozoa controls were prepared by omitting terminal deoxytransferase (TdT) enzyme from the staining step. Positive spermatozoa controls were prepared by inducing SDF with 2% (*v/v*) hydrogen peroxide (Sigma-Aldrich) and incubated for 1 h at 50°C in a heating block. At the end of the incubation, the spermatozoa sample was centrifuged for 7 min at 300*g* to remove the hydrogen peroxide. After washing twice with PBS, the supernatant was discarded and 1 ml of PFA was added to the sperm pellet.[23] Before the staining procedure, the test, negative, and positive samples were centrifuged at 400 *g* for 7 min. PFA was removed and the sample was re-suspended in 70% (*v/v*) ice-cold ethanol, kept at −20°C, and batched until analyses.

SDF was evaluated using a TUNEL assay with an Apo-Direct™ kit (Pharmingen, San Diego, CA, USA). Also included in the kit were the negative and positive kit controls. These are not spermatozoa but human lymphoma cell line suspended in ethanol. All frozen test, positive, and negative spermatozoa control samples stored in ethanol were brought to room temperature. After removal of ethanol, kit controls and test samples were washed two times with "wash buffer" to remove ethanol. After removal of the buffer, 50 µl of freshly prepared staining solution was added to each aliquot as per the instructions from the manufacturer. The staining solution contains reaction buffer, TdT enzyme, fluorescein isothiocyanate-tagged deoxyuridine triphosphate nucleotides (FITC-dUTP), and distilled water. All specimens were incubated in the dark for 60 min. Excess stain was removed by adding 1 ml of "rinse buffer" and centrifuged twice. The samples were re-suspended in 0.5 ml of PI/RNase solution and incubated for 30 min followed by flow cytometric analysis.[23]

BD Accuri C6 flow cytometer analysis

All fluorescence signals of labeled spermatozoa were analyzed by the BD Accuri C6 Flow Cytometer and C6 Plus Flow Cytometer (Becton Dickinson, San Jose, CA, USA). A total of 10 000 spermatozoa was counted with fluidic set at "slow" with flow rate of <100 cells per second. The laser excitation is provided at two wavelengths of 488 nm supplied by a solid blue laser at 20 mW and 640 nm powered by 14.7-mW diode red laser. Green fluorescence (480–530 nm) is measured in the FL-1 channel and red fluorescence (640 nm) in the FL-2 channel. The percentage of positive cells (TUNEL-positive) was calculated on a 1023-channel scale from the flow cytometer software. The analysis was conducted with a similar strategy on both the C6 and the C6 Plus flow cytometer, and dot plots were generated by the BD Accuri software (BD Biosciences, Ann Arbor, MI, USA).

TUNEL/PI data acquisition

The same set of sample was run simultaneously on the C6 and C6 Plus flow cytometers. We used the BD Accuri workspace (BD Biosciences, San Jose, CA, USA) and gating strategy which was standardized using three plots and sequential gating as described in our earlier study.[21] The following plots were used:

1. Forward scatter versus side scatter or "Plot a" – Gate was drawn to exclude the small debris and larger nonsperm cells as well as seminal cells. The gating helped select only spermatozoa, which had the expected size or G1 population. This was achieved with gating spermatozoa stained with PI with a flame-shaped gate.[24] The flame-shaped gate was established in the forward scatter (FSC) versus side scatter (SSC) plot to select the sperm population. This helps remove the large cells and nonsperm cells as well as the debris
2. PI fluorescence and FSC or "Plot b" – The PI positive events were gated within the cells belonging to the G1 population by application of the PI-positive gating (G2). The PI-positive event gating application selected the PI-positive spermatozoa and excluded the debris and M450 apoptotic bodies. The M450 bodies are devoid of nuclei and are not stained by PI stain. A very small percentage of the M450 bodies are positive for TUNEL staining, and these M450 bodies are known to interfere with TUNEL assay. The combination of TUNEL with PI staining helps exclude the M450 bodies[25]
3. PI-fluorescence and FITC-fluorescence positive gates or "Plot c" – The upper right quadrant delineated all the TUNEL as well as PI-positive samples. Results were expressed as percentage of sperm with DNA fragmentation.[23]

Samples were run in parallel and analyzed on the C6 flow cytometer (designated as the standard) and the C6 Plus (designated as the test flow cytometer). A minor modification in the gates in FSC versus SSC plots was done for the C6 Plus to accommodate better location of the cell population.

Standardization of the test flow cytometer

The "acquisition workspace" was standardized for the C6 flow cytometer (standard instrument) as described by us in an earlier publication.[21] The gate settings of C6 Plus (test flow cytometer) were compared with those of the C6 flow cytometer. The same standardized acquisition space was imported into the C6 Plus cytometer. The settings for the new model C6 plus cytometer have to be calibrated against the standard instrument before clinical use. This requires adjustment of the acquisition workspace for comparable results. In the first assay, acquisition was done on the C6 Plus with the unadjusted standardized workspace and also with the modified or adjusted workspace. The two sets of SDF values obtained from the different workspaces for C6 Plus were compared with the SDF values obtained from C6 machine. To accomplish this, we used the dual strategy for data analysis as described briefly in our earlier publication.[26]

Unadjusted and adjusted C6 Plus versus standard C6 flow cytometer variation

Semen samples ($n = 23$) were read in parallel on the newer and updated C6 Plus version (with unadjusted) settings and with acquisition settings that were adjusted. These samples were also run and data were acquired with the standard C6 flow cytometer. The concordance coefficient correlation, precision, accuracy, cutoff, sensitivity, and specificity were examined in the unadjusted settings and after adjustment in the C6 Plus versus C6 flow cytometer for comparison of the results.

Inter- and Intra-observer variation

Inter- and intra-observer variability was examined by analyzing the differences in the results produced by the two observers. The interobserver variability was assessed by the likelihood that an observer's designated value (for either observer) for a given measurement was within a specified value (either by absolute difference or percentage difference) from the mean designated value of the two observers.

Inter- and Intra-assay variation and precision of the instrument

Inter- and intra-assay variation was measured using the negative and positive controls provided in assay kit. These are human lymphoma cells. In addition, we used negative and positive internal controls that consisted of semen samples from the patient population. The intraobserver variability was calculated by determining the degree to which an individual measurement differed from the final assigned TUNEL measurement.

Statistical analyses

Comparison of the C6 and C6 Plus flow cytometer was done with the MedCalc Statistical Software, version 18.2.1 (MedCalc Software bvba, Ostend, Belgium; https://www.medcalc.org; 2017). The C6 Plus readings were examined with standard C6 settings before adjustment and after adjustment of these settings. Spearman rank correlation; Chi-square tests for cross-tables, concordance correlation (to determine concordance, precision (Pearson ρ), and accuracy (bias correction factor C_b); Passing-Bablok regression and Bland-Altman plots (for direct comparison of two measurement methods); and receiver operating characteristic (ROC)-curve analyses (to determine sensitivity, specificity, positive and negative predictive value, and the diagnostic cutoff of a method) were used to compare the SDF of the newer C6 Plus version with the results from the standard C6 flow cytometer. For the ROC-curve analysis, the discriminant variable was the identification of a patient with the C6 flow cytometer as the cutoff was previously established.

The Tukey's test was used for outlier detection.[27] The Kolmogorov–Smirnov for normal distribution was used for assessing whether the data were normally distributed or not. Accordingly, Pearson Chi-square or Spearman rank correlations were calculated. Concordance of the two instruments was assessed and the methods were compared by the Passing-Bablok regression analysis. Bland-Altman plots were generated for different categories including spermatozoa from patient samples, donor samples, and internal sperm-positive and sperm-negative control samples for comparison of the C6 versus C6 Plus flow cytometer.

RESULTS

Representative figures showing the plots with unadjusted and adjusted acquisition settings for positive sample on C6 Plus are shown in **Figure 1.** The acquisition plots were obtained using unadjusted gates for acquisition of the data (**Figure 1a–1c**). The gates were shifted to include the population of interest (**Figure 1d**). The acquisition plots (**Figure 1d–1f**) were obtained with data acquisition utilizing the adjusted settings for the positive sample. The percentage of sperm positive for DNA fragmentation is shown in the upper right quadrant (**Figure 1f**). Similarly, the dot plots obtained on the C6 Plus with the adjusted gates are shown in **Figure 2a–2c**. These plots show the data for negative sample, used for setting the quadrant. Dot plots were acquired with adjusted gate settings on C6 Plus (**Figure 2d–2f**). The upper right quadrant in acquisition plot in **Figure 2c** shows negative DFI value for the sample. Plots in **Figure 2e** and **2f** show the acquisition for a positive sperm sample run in duplicate. The duplicate readings show similar DFI values. A significant ($P < 0.0001$) difference was seen in the correlation between the adjusted and unadjusted C6 Plus with the standardized C6. Of the 23 measurements, the unadjusted C6 Plus failed to detect correctly 5 (21.7%) samples, 7 (30.4%) were true negative, and 11 (47.8%) true positive (**Figure 3a** and **3b**). Lower concordance correlation (0.6155 *vs* 0.9775), precision ($\rho = 0.7479$ *vs* $\rho = 0.9934$), and accuracy ($C_b = 0.8230$ *vs* $C_b = 0.9840$) were seen with

unadjusted C6 Plus. Unadjusted C6 Plus had low sensitivity (94.4%) and specificity (80.0%) and an area under curve (AUC) of 0.90. The positive predictive value was 94.4% and the negative predictive value was 80.0%. In addition, the cutoff for percentage SDF was significantly lower and only 7.5% instead of the established cutoff of 17.0% (**Figure 4a**). With adjusted settings, 100% sensitivity and specificity were seen for the adjusted C6 Plus cytometer. The cutoff for SDF was greater than 17% with 100% positive and negative predictive values and AUC of 1.00. C6 Plus cytometer could predict TUNEL results with the same accuracy as C6 cytometer (**Figure 4b**).

A total of 147 TUNEL measurements were examined on both flow cytometers. These included measurements from negative and positive controls from the kit, spermatozoa from donor and patient with negative and positive controls (internal spermatozoa controls), and donor and patient test samples. There were significant differences in the SDF values for the first set of comparisons with unadjusted gate setting for the C6 Plus. Following adjustment, the SDF values were comparable between the C6 and C6 Plus with high precision and accuracy.

The SDF ranged from 0 to 51.4% on the C6 flow cytometer with a mean value of 15.1% (95% CI: 10.1–15.7). Similarly, on the C6 Plus flow cytometer, SDF ranged from 0 to 47.5% with a mean value of 14.8% (95% CI: 9.9–15.2). The overall correlation coefficient observed between the C6 and C6 Plus was 0.984 (**Supplementary Figure 2**).

Standardization of the test flow cytometer

Twenty-three measurements were obtained twice on the C6 Plus flow cytometer with two different "acquisition workspaces," and the same measurements were obtained on C6 cytometer with the standardized workspace. The two sets of SDF values were obtained with "unadjusted acquisition workspace" and "adjusted acquisition workspace" on the C6 Plus flow cytometer.

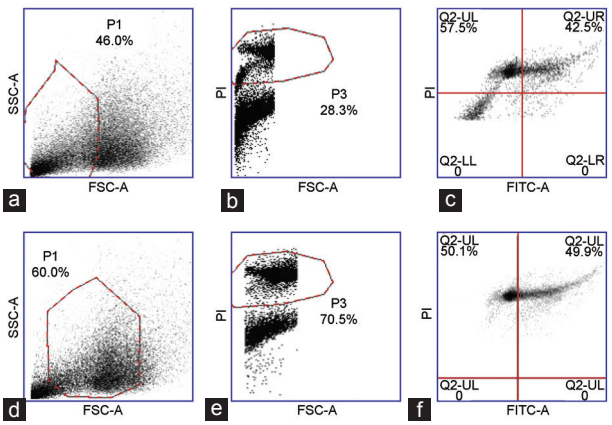


Figure 1: Representative dot plots on the C6 Plus flow cytometer showing comparison of data acquisition utilizing unadjusted and adjusted settings with data acquisition images of positive sample with (**a**) unadjusted flame shaped gate settings on C6 Plus flow cytometer (FSC-A *vs* SSC-A), (**b**) unadjusted settings on C6 Plus flow cytometer (FSC-A *vs* PI), (**c**) unadjusted settings on C6 Plus flow cytometer (FITC-A *vs* PI), (**d**) adjusted settings on C6 Plus flow cytometer (FSC-A *vs* SSC-A) including the sperm and the apoptotic bodies. (**e**) Data plots of positive sperm sample with the adjusted settings on C6 Plus flow cytometer (FSC-A *vs* PI) including the PI-positive events within the gate and excluding the M450 apoptotic bodies. (**f**) Data acquisition images of positive sperm sample with the adjusted settings on C6 Plus flow cytometer (FITC-A *vs* PI): the TUNEL/PI-positive events are depicted in the right upper quadrant and this estimates the TUNEL positivity. FITC-A: fluorescein isothiocynate-area; SSC-A: side scatter area; PI: propidium iodide; UL: upper left; UR: upper right; LL: lower left; LR: lower right; P1: gate P1; P3: gate P3; Q2: quadrant 2.

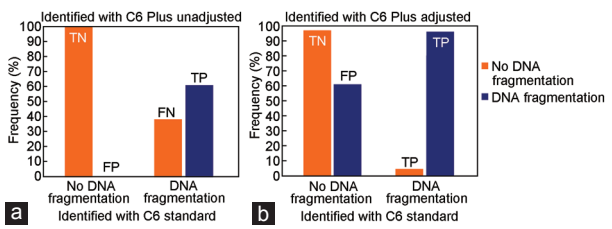


Figure 3: Percentage chart showing extent of DNA fragmentation in (**a**) measurements without DNA fragmentation (orange) and DNA fragmentation (blue) on C6 Plus without adjustment when compared with C6 standard. In measurements without DNA fragmentation, it could not differentiate the false-positive measurements and all were identified as true-negative measurements. However, in measurements with DNA fragmentation, it incorrectly identified 40% of measurements as false negative and only 60% as true positive, (**b**) same measurements on C6 Plus after adjustment. The number of false -positive measurements without DNA fragmentation was significantly low. Majority of the measurements were correctly identified as true positive in case of DNA fragmentation. TN: true negative; FP: false positive; FN: false negative; TP: true positive.

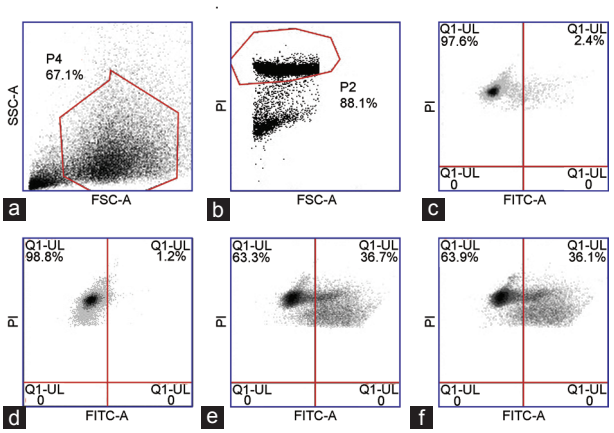


Figure 2: Dot plots obtained on the C6 Plus with the modified or adjusted gates for negative control sperm sample. The first plot **a** depicts the FSC-A versus SSC-A plot with flame shaped gate that captures the spermatozoa and apoptotic bodies, plot **b** is a PI-A/FSC-A plot with PI positive events. This gate includes all spermatozoa and excludes the M450 apoptotic bodies, (**c**) is the PI-A/FITC-A plot for the negative sperm control which is used to set the quadrants, (**d**) depicts the PI-A versus FITC-A gate plot of a negative test sample. The positive sperm sample was run in duplicate, and shows SDF of (**e**) 36.7% and (**f**) 36.1 %. The SDF values of the test sample run in duplicate reflect the very low intra-assay variability. FITC-A: fluorescein isothiocynate-area; SSC-A: side scatter-area; PI: propidium iodide; SDF: sperm DNA fragmentation; UL: upper left; UR: upper right; LL: lower left; LR: lower right; P2: gate P2; P4: gate P4; Q1: quadrant 1.

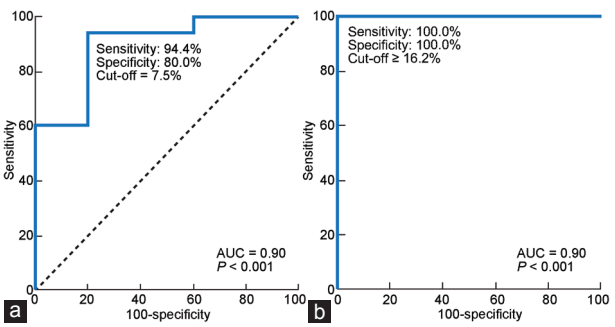


Figure 4: Receiver operating characteristic curve showing sensitivity, specificity, and cutoff for C6 Plus (**a**) in unadjusted SDF and (**b**) after adjustment with the standard C6 flow cytometer. AUC: area under curve; SDF: sperm DNA fragmentation.

There was a significant positive correlation of SDF values obtained with unadjusted workspace on C6 Plus cytometer (r = 0.854; 95% CI: 0.682–0.937; P < 0.0001) with those of the standardized workspace for the C6 cytometer. When the SDF values acquired from the adjusted workspace in the C6 Plus were correlated with values from standardized workspace for C6 flow cytometer (**Figure 5**), the resultant correlation was highly significant (r = 0.992; 95% CI: 0.982–0.997; P < 0.0001) and even significantly stronger than the one for the correlation of the unadjusted workspace on C6 Plus with the standardized workspace for C6 cytometer (P < 0.0001).

Methods comparison

After direct correlation of the percentage values of SDF, a significant correlation (r = 0.984; P < 0.0001) was observed (**Figure 6**). However, despite this correlation, patients could not be correctly identified if the evaluation acquisition template for the C6 Plus flow cytometer was not adjusted. This discrepancy is shown in a frequency chart indicating the percentages of true-positive and -negative values and false-positive and -negative values (**Figure 3**). This difference is also reflected in a lower AUC, cutoff, specificity, and sensitivity after ROC-curve analysis (**Figure 4**).

After Passing-Bablok linear regression, the Cusum test for linearity showed no deviation from linearity (P = 0.99 and P = 0.78, respectively) in both cases (**Figure 6**). The differences between the methods were much smaller for the C6 Plus adjusted and the C6 standard method than those for the C6 Plus unadjusted and the C6 standard method. The smaller differences are obvious from the Bland-Altman plots (**Figure 5**). Whereas the mean difference for the comparison of C6 Plus unadjusted with C6 standard method is 11.8 with limits of agreement of −19.1 and 42.7, it was −3.7 with limits of agreement of −11.1 and 3.6 for the comparison C6 Plus adjusted with C6 standard.

Precision and accuracy of two flow cytometers

This consisted of all 147 measurements from experiments 2 and 4 as shown in **Supplementary Figure 1**. Owing to the wider 95% CI of the readings (0.3552–0.7872) for the correlation between the data of the C6 Plus unadjusted and C6 standard than those of the correlation between C6 Plus adjusted and C6 standard (0.9554–0.9887), the concordance correlation coefficient (0.6155), precision (ρ = 0.7479), and accuracy (C_b = 0.8230) were much lower for the unadjusted than for the adjusted C6 Plus instrument (concordance correlation coefficient: 0.9775; precision: ρ = 0.9934; accuracy: C_b = 0.9840). Thus, the two flow cytometers showed a strong concordance with a very high precision of 98% and an accuracy of >99%.

Intra- and inter-observer variability

Intra- and inter-observer variability was examined on 54 TUNEL measurements by the two flow cytometers. Two highly experienced observers conducted the tests in a blinded fashion. The differences in inter- and intra-observer calculations were obtained from these values. There was a strong agreement in the values obtained from the two observers on the C6 flow cytometer from 29 TUNEL measurements, with a rank correlation coefficient of 0.922 (P < 0.0001) and 95% CI of 0.83–0.97. Similarly, the interobserver variability on C6 Plus (25 TUNEL measurements) was small and showed a strong correlation (r = 0.993; P < 0.0001) with 95% CI of 0.985–0.997.

Differences in the values were within the acceptable range. No significant differences were observed between the duplicates in any of the experiments performed. The two observers also demonstrated a high intraobserver agreement, both on C6 for observer 1 (r = 0.948; 95% CI: 0.885–0.997; P < 0.0001) and observer 2 (r = 0.941, 95% CI: 0.877–0.972; P < 0.0001). The interobserver agreement (weighted kappa ± standard error) on the C6 flow cytometer for the two observers was 0.801 ± 0.062 and on C6 Plus flow cytometer was 0.746 ± 0.044.

DISCUSSION

Sperm DNA integrity is included in the clinical assessment of male infertility in many Andrology centers. The most common tests of DNA integrity are the sperm chromatin structure assay (SCSA), TUNEL assay, sperm chromatin dispersion (SCD) assay, and Comet assay. Presently, the routine use of SDF tests in male factor evaluation is generally not supported by professional societies.[12–14] However, the potential role of SDF has been acknowledged in the latest American Society for Reproductive Medicine (ASRM),[28] American Urological Association (AUA),[29] and European Association of Urology (EAU) guidelines on male infertility.[30] The uniqueness of SDF testing in providing assessment of the genetic content of the male gamete could be considered as complementary to conventional semen analysis in improving male factor evaluation. The practice recommendations by Agarwal *et al.*[31] represent the first evidence-based attempt to identify the circumstances in which SDF testing should be of the greatest clinical value. There are several adequately powered studies that have highlighted the negative impact of high SDF on certain conditions such as varicocele and unexplained infertility and on ART outcomes; these studies have further elaborated on the significance of SDF assessment as an important parameter for these select conditions.

One of the reasons for the variability in the reference values is the fact that different tests measure different endpoints. Yet, of more importance is the lack of standardized tests with clearly established reference values. The TUNEL assay combined with the PI staining is one of the more robust techniques available for SDF assessment.

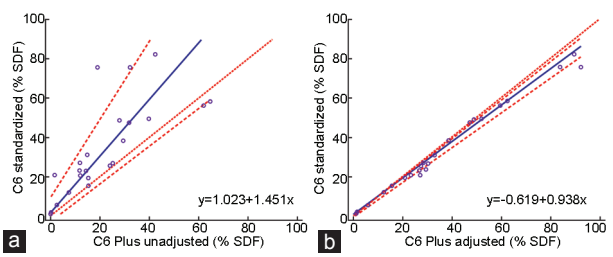


Figure 5: Bland-Altman plots for (**a**) unadjusted C6 Plus versus C6 standardized and (**b**) adjusted C6 Plus versus C6 standardized. SDF: sperm DNA fragmentation; s.d.: standard deviation.

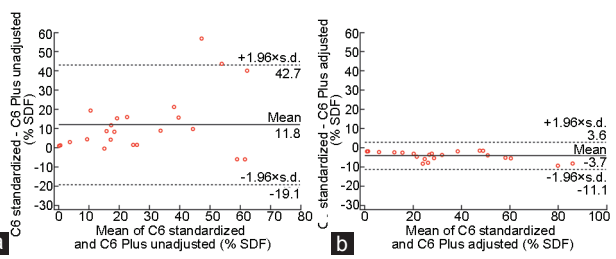


Figure 6: Passing-Bablok regression analysis showing (**a**) unadjusted C6 Plus versus C6 standardized (the wider deviation of the values from one another can clearly be seen), and (**b**) adjusted C6 Plus versus C6 standardized (the deviation in the values is much narrower). Red dashed line shows the confidence interval of the regression line expected averages of SDF; red dotted line is the identity line and blue is the regression line of the measured averages. SDF: sperm DNA fragmentation; s.d.: standard deviation.

Paired with flow cytometry, the TUNEL assay is becoming increasingly popular for testing SDF. The assay is gaining more acceptance by Andrology laboratories across the globe because it is a direct assay, when compared with other tests such as SCSA, SCD, and Comet assays, which are indirect tests and are dependent on acid or alkali denaturation for the assessment of the chromatin structural integrity. Furthermore, the assessment of 10 000 spermatozoa is rapidly conducted by flow cytometry.

Studies from our laboratory have established the precision and accuracy of the TUNEL-PI assay.[21,31,32] In our earlier study, we established the cutoff values of SDF for infertile men and highlighted the ability of SDF as a biomarker to discriminate between normal (normozoospermic) and infertile men.[19] The precision and accuracy of a test have important implications for the clinical validity of the test. SDF as assessed by TUNEL assay has been reported to be a test with high precision and reliability in a study conducted in two laboratories in two continents.[21]

We have demonstrated significant differences in the cutoff, precision, accuracy, sensitivity, and specificity if the upgraded flow cytometer is not calibrated against the standard flow cytometer. With adjusted settings, C6 Plus can predict normal or abnormal SDF in the same way as the C6. The results of the study establish the precision, accuracy, and reproducibility of SDF on both models of the bench-top flow cytometer. A central highlight of the current study is that the evaluation of SDF measurements was conducted on and with identical samples and specimen preparation conditions as well as a validated template for measuring the percentage of SDF. Furthermore, we have utilized semen samples with both normal and abnormal parameter values from donors and patients to maintain the heterogeneity of the samples tested and therefore be representative of different types of patient samples. Consequently, this can lead to a better predictive capacity of the assay in both the natural and assisted fertility setting.

The kit controls provided for DNA fragmentation contain a cell population other than spermatozoa, which have a different gate setting. The strength of the study is that the performance verification was done for samples including normozoospermic samples, samples from infertile men, as well as both positive and negative controls from sperm samples. A limitation of our study was that we did not include any subset of subjects with specific clinical conditions such as varicocele or subjects who were smokers and more likely to generate higher levels of reactive oxygen species (ROS) and therefore higher DNA fragmentation. It is also important to recognize that the assay itself has some inherent intra-assay variability, and the intraobserver variability described in this process includes an element of intra-assay variability. This variability can be due to differences in the negative and positive kit controls or negative and positive sperm samples used as controls with each batch.

A lack of systematic bias between the performances of the two instruments validates the new instrument against the standard, and the new model can be then utilized in the laboratory for clinical testing. It also helps provide compelling evidence that the instrument(s) have been validated with standardized lab protocols. This documentation can also be made available to various accrediting agencies that may request such information for the specialized instruments in any given laboratory. Furthermore, recommendations can be made to help increase the confidence of the results generated using these instruments. The results obtained can be relied upon once the protocol has been standardized and validated for any instrument that is upgraded with newer software or replaced with a newer model.

CONCLUSION

We recommend that all new or updated diagnostic flow cytometers should have settings and adjustments validated to obtain comparable diagnostic results. Furthermore, the study also establishes the reproducibility of the two bench-top flow cytometers when used with the standardized staining protocol and offers clear-cut evidence of robustness of the TUNEL results obtained by the two machines. To prevent misreporting clinical results, the adjustment and optimization of gate settings as well as revalidation are essential for all other diagnostic equipment used in Andrology and/or IVF laboratories, with routine hardware or software upgrades or newer models before use in clinical laboratories.

AUTHOR CONTRIBUTIONS

AA, SG, and RS planned the experiments. RS and SG analyzed the collected data. RH performed the statistical analyses. RS and SG were involved in the manuscript preparation. RS, SG, AA, and RH critically reviewed the manuscript. All authors read and approved the final manuscript.

COMPETING INTERESTS

All authors declare no competing interests

ACKNOWLEDGMENTS

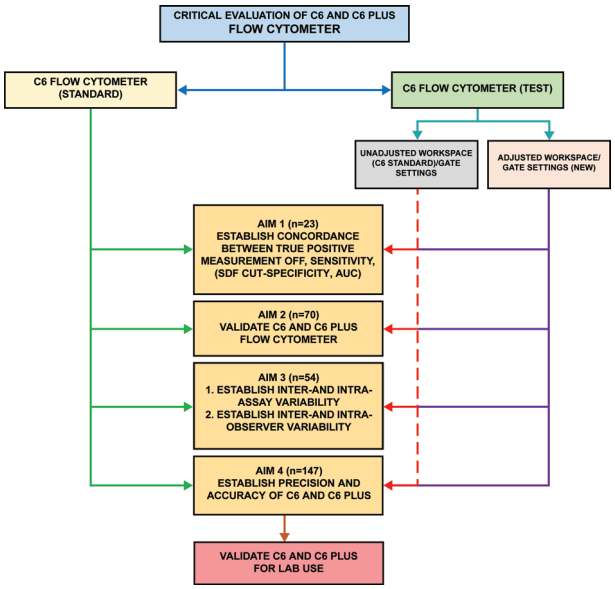
The authors are grateful to the Andrology Center technologists for scheduling the study subjects. Financial support was provided by the American Center for Reproductive Medicine, Cleveland Clinic.

Supplementary Information is linked to the online version of the paper on the *Asian Journal of Andrology* website.

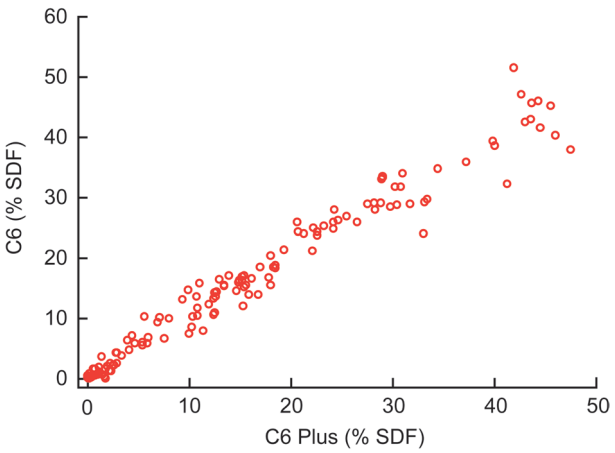
REFERENCES

1 Thonneau P, Marchand S, Tallec A, Ferial ML, Ducot B, *et al*. Incidence and main causes of infertility in a resident population (1,850,000) of three French regions (1988-1989). *Hum Reprod* 1991; 6: 811–6.
2 Aitken RJ, De Iuliis GN. Origins and consequences of DNA damage in male germ cells. *Reprod Biomed Online* 2007; 14: 727–33.
3 Evenson DP, Wixon R. Data analysis of two *in vivo* fertility studies using Sperm Chromatin Structure Assay-derived DNA fragmentation index vs. pregnancy outcome. *Fertil Steril* 2008; 90: 1229–31.
4 Spano M, Bonde JP, Hjollund HI, Kolstad HA, Cordelli E, *et al*. Sperm chromatin damage impairs human fertility. The Danish First Pregnancy Planner Study Team. *Fertil Steril* 2000; 73: 43–50.
5 Duran EH, Morshedi M, Taylor S, Oehninger S. Sperm DNA quality predicts intrauterine insemination outcome: a prospective cohort study. *Hum Reprod* 2002; 17: 3122–8.
6 Bungum M, Humaidan P, Axmon A, Spano M, Bungum L, *et al*. Sperm DNA integrity assessment in prediction of assisted reproduction technology outcome. *Hum Reprod* 2007; 22: 174–9.
7 Morris ID, Ilott S, Dixon L, Brison DR. The spectrum of DNA damage in human sperm assessed by single cell gel electrophoresis (Comet assay) and its relationship to fertilization and embryo development. *Hum Reprod* 2002; 17: 990–8.
8 Robinson L, Gallos ID, Conner SJ, Rajkhowa M, Miller D, *et al*. The effect of sperm DNA fragmentation on miscarriage rates: a systematic review and meta-analysis. *Hum Reprod* 2012; 27: 2908–17.
9 Zini A, Boman JM, Belzile E, Ciampi A. Sperm DNA damage is associated with an increased risk of pregnancy loss after IVF and ICSI: systematic review and meta-analysis. *Hum Reprod* 2008; 23: 2663–8.
10 Agarwal A, Cho CL, Esteves SC. Should we evaluate and treat sperm DNA fragmentation? *Curr Opin Obstet Gynecol* 2016; 28: 164–71.
11 Gosálvez J, López-Fernández C, Fernández JL, Esteves SC, Johnston SD, *et al*. Unpacking the mysteries of sperm DNA fragmentation ten frequently asked questions. *J Reprod Biotechnol Fertil* 2015; 4: 1–16.
12 Practice Committee of American Society for Reproductive Medicine. The clinical utility of sperm DNA integrity testing. *Fertil Steril* 2008; 90: S178–80.
13 Barratt CL, Aitken RJ, Bjorndahl L, Carrell DT, de Boer P, *et al*. Sperm DNA: organization, protection and vulnerability: from basic science to clinical applications – a position report. *Hum Reprod* 2010; 25: 824–38.

14 Tomlinson M, Lewis S, Morroll D. British Fertility Society. Sperm quality and its relationship to natural and assisted conception: British Fertility Society guidelines for practice. *Hum Fertil* (*Camb*) 2013; 16: 175–93.
15 Mitchell LA, De Iuliis GN, Aitken RJ. The TUNEL assay consistently underestimates DNA damage in human spermatozoa and is influenced by DNA compaction and cell vitality: development of an improved methodology. *Int J Androl* 2011; 34: 2–13.
16 Sharma R, Agarwal A. Laboratory evaluation of sperm chromatin: TUNEL assay. In: Zini A, Agarwal A, editors. Sperm chromatin: biological and clinical applications in male infertility and assisted reproduction. Ch. 14. New York: Springer Science Business Media, LLC; 2011. p201–15.
17 Ribeiro SC, Muratori M, De Geyter M, De Geyter C. TUNEL labeling with BrdUTP/anti-BrdUTP greatly underestimates the level of sperm DNA fragmentation in semen evaluation. *PLoS One* 2017; 12: e0181802.
18 Sergerie M, Laforest G, Bujan L, Bissonnette F, Bleau G. Sperm DNA fragmentation: threshold value in male fertility. *Hum Reprod* 2005; 20: 3446–51.
19 Sharma RK, Sabanegh E, Mahfouz R, Gupta S, Thiyagarajan A, *et al*. TUNEL as a test for sperm DNA damage in the evaluation of male infertility. *Urology* 2010; 76: 1380–6.
20 Dominguez-Fandos D, Camejo MI, Ballesca JL, Oliva R. Human sperm DNA fragmentation: correlation of TUNEL results as assessed by flow cytometry and optical microscopy. *Cytometry Part A* 2007; 71A: 1011–8.
21 Ribeiro S, Sharma R, Gupta S, Cakar Z, De Geyter C, *et al*. Inter- and intra-laboratory standardization of TUNEL assay for assessment of sperm DNA fragmentation. *Andrology* 2017; 5: 477–85.
22 World Health Organization. WHO Laboratory Manual for the Examination and Processing of Human Sperm. 5th ed. Geneva: World Health Organization; 2010.
23 Agarwal A, Gupta S, Sharma R. Measurement of DNA fragmentation in spermatozoa by TUNEL assay using bench top flow cytometer. In: Agarwal A, Gupta S, Sharma R, editors. Andrological evaluation of male infertility: a laboratory guide. Ch. 24. New York: Springer Publishing; 2016. p181–203.
24 Muratori M, Marchiani S, Tamburrino L, Tocci V, Failli P, *et al*. Nuclear staining identifies two populations of human sperm with different DNA fragmentation extent and relationship with semen parameters. *Hum Reprod* 2008; 23: 1035–43.
25 Muratori M, Tamburrino L, Marchiani S, Cambi M, Olivito B, *et al*. Investigation on the origin of sperm DNA fragmentation: role of apoptosis, immaturity and oxidative stress. *Mol Med* 2015; 21: 109–22.
26 Gupta S, Sharma R, Agarwal A. Inter-and intra-laboratory standardization of tunel assay for assessment of sperm DNA fragmentation. *Curr Protoc Toxicol* 2017; 74: 16.11.1–16.11.22.
27 Tukey JW. Tukey Exploratory Data Analysis. Reading: Addison Weseley; 1977.
28 Practice Committee of the American Society for Reproductive Medicine. Diagnostic evaluation of the infertile male: a committee opinion. *Fertil Steril* 2015; 103: e18–25.
29 Jarow J, Sigman M, Kolettis PN, Lipshultz L, McClure D, *et al*. The optimal evaluation of the infertile male: best practice statement reviewed and validity confirmed 2011 in American Urological Association; 2011.
30 Jungwirth A, Diemer T, Kopa Z, Krausz S, Minhas S, *et al*. Guidelines on male infertility. Available from: https://www.uroweb.org/guideline/male-infertility/. [Last accessed on 2018 Nov 09].
31 Agarwal A, Majzoub A, Esteves SC, Ko E, Ramasamy R, *et al*. Clinical utility of sperm DNA fragmentation testing: practice recommendations based on clinical scenarios. *Transl Androl Urol* 2016b; 5: 935–50.
32 Sharma R, Ahmad G, Esteves SC, Agarwal A. Terminal deoxynucleotidyl transferase dUTP nick end labeling (TUNEL) assay using bench top flow cytometer for evaluation of sperm DNA fragmentation in fertility laboratories: protocol, reference values, and quality control. *J Assisted Reprod Genetics* 2016; 33: 291–300.



Supplementary Figure 1: Flow diagram showing the study objectives on the standard C6 flow cytometer and the unadjusted and adjusted C6 Plus flow cytometer.



Supplementary Figure 2: Scatter plot of the correlation between C6 and C6 Plus measurements of the percentage SDF ($r = 0.984$; $P < 0.0001$).