Quality Control of Reactive Oxygen Species Measurement by Luminol-Dependent Chemiluminescence Assay

HIROSHI KOBAYASHI,* ENRIQUE GIL-GUZMAN,* AYMAN M. MAHRAN,* RAKESH K. SHARMA,* DAVID R. NELSON;† ANTHONY J. THOMAS JR,* and ASHOK AGARWAL*

From the *Center for Advanced Research in Human Reproduction, Infertility, and Sexual Function, Urological Research and Clinical Laboratories, Urological Institute, The Cleveland Clinic Foundation, Cleveland, OH 44195; and the †Department of Biostatistics and Epidemiology, The Cleveland Clinic Foundation, Cleveland, Ohio.

ABSTRACT: A total of 28 donor semen samples were used to evaluate the characteristics of laboratory variability in measuring reactive oxygen species (ROS). The objectives of this study were to assess the interassay (same sample observed on different days by the same observers) variability; interdonor, intraobserver (replications of the same sample on the same day) variability; and interobserver (multiple observers on the same day with the same sample) variability of the luminol-dependent chemiluminescence assay and to establish an optimal semen age and sperm concentration. Semen samples were collected from 6 healthy donors for 108 measures of ROS. ROS levels were measured by the assay using luminol as the probe. An additional assessment measured the effect of time (age of the sample) on ROS production in 12 donor samples at 60, 120, 180, and 240 minutes after the specimen was produced. Last, to evaluate the effect of sperm concentration on ROS production, ROS concentrations ranging from 1 to 120 × 10^6/mL. In the controls, the mean ROS level was 0.218 × 10^6 counted photons per minute; the interassay variability standard deviation (SD) was 0.077. The interobserver SD was 0.002 for an interobserver reliability of 97.5% (coefficient of variation [CV] = 0.9%). The intraobserver (between replication) SD was 0.001 for an intraobserver reliability of 98.7% (CV = 0.5%). The interassay SD was 0.005 for an interassay reliability of 93.8% (CV = 2.0%). There was no statistically significant interobserver, intraobserver, or interassay variation (P > .05). ROS levels decreased significantly with time; a dramatic decline in ROS production was seen in the specimens that were more than 60 minutes old (P < .001). ROS values decreased by 31% at 120 minutes and 62% at 180 minutes compared with the 60-minute-old specimens. A linear relationship was seen between the ROS levels and sperm concentration in 8 of the 10 samples analyzed (R^2 = .99). Our results demonstrate that the luminol-dependent chemiluminescence assay for ROS measurement is both accurate and reliable when the sperm concentration is greater than 1 × 10^6/mL and the samples are analyzed within the first hour after specimen collection.

Key words: Semen, spermatozoa, luminol, oxidative stress, male infertility.

J Androl 2001;22:568–574

The role of reactive oxygen species (ROS) in male infertility is now well recognized (Aitken and Clark-son, 1987; Iwasaki and Gagnon, 1992; Aitken and Fisher, 1994; Sikka et al, 1995; Sharma and Agarwal, 1996; Grieveau and de Lennou, 1997; Ochsendorf, 1999). Because spermatozoa contain a large amount of polyunsaturated fatty acids, they are vulnerable to ROS attack (Jones et al, 1979; Alvarez et al, 1987; Aitken, 1989; Rao et al, 1989; Aitken et al, 1993). Irrespective of the clinical diagnosis and semen characteristics, the presence of seminal oxidative stress in infertile men suggests that ROS plays a major role in the pathophysiology of male infertility (Pasqualotto et al, 2000).

ROS measurement appears to be a helpful tool in the initial evaluation and follow-up of infertile male patients because high oxidative stress seems to be strongly correlated with low fertility (Sharma et al, 1999). Also, a decrease in the functional competence of spermatozoa (sperm-oocyte fusion, acrosome reaction, and zona-free hamster oocyte penetration test) has been linked to high ROS levels (Aitken et al, 1989a; Gomez et al, 1996).

Numerous assays for ROS measurement have been introduced (Sharma and Agarwal, 1996; Ochsendorf, 1999). Measuring ROS directly in vivo is difficult because reactive oxidants are highly unstable and generally have very short-lived intermediates. Free radicals can be detected directly by electron paramagnetic resonance spectroscopy. However, to detect short-lived radicals such as alkoxyl or peroxyl radicals, the measurements must be performed at low temperatures—a condition that limits its use in vivo (Ochsendorf, 1999).

A wide variety of indirect methods of ROS measure-
ment have been designed (Weber, 1990; Ochsendorf, 1999). These assays detect oxidized end products. The chemiluminescence assay is one such method (Murphy and Sies, 1990; Aitken and Buckingham, 1992; Aitken et al, 1992). Two probes may be used with the assay: luminol and lucigenin. A luminol-mediated chemiluminescence signal in spermatozoa occurs when luminol oxidizes at the acrosomal level. Luminol undergoes an intracellular deoxygenation reaction mediated by a heterogeneous group of sperm peroxidases in the presence of hydrogen peroxide, whereas lucigenin is oxidized at the extracellular level by the superoxide anion (Aitken and Buckingham, 1992; Aitken et al, 1992).

The luminol assay is more advantageous for a number of reasons. It can measure H$_2$O$_2$, O$_2^-$, and OH$^-$ levels (Murphy and Sies, 1990; McKinney et al, 1996; Sharma and Agarwal, 1996), although it cannot distinguish these oxidants from one another. It can also measure the global level of ROS under physiological conditions, and it is easy to use. In addition, the assay can measure both extracellular and intracellular ROS, which means it has a high sensitivity (Sharma and Agarwal, 1996). Multiple studies have correlated high chemiluminescent signals using luminol as a probe with adverse effects on sperm function. The assay can be sensitized by adding horseradish peroxidase to the sperm suspension, thereby increasing the spontaneous luminescence levels commonly observed in healthy semen samples (Aitken et al, 1992).

The final chemiluminescent signal is the integrated sum of the partial signals generated by every spermatozoon. In this way, the amount of ROS measured is related to the ROS production capacity of each spermatozoon and the number of spermatozoa with this ability in a given sample. To optimize the ROS assay, several authors have maintained a constant sperm concentration (10–20 × 10$^6$ spermatozoa/mL). However, that is difficult to achieve, particularly in patients with oligospermia (Aitken et al, 1989b).

In addition, ROS generation is an energy-dependent reaction that requires large amounts of substrate (O$_2$; Griendeau et al, 1998) and metabolically active spermatozoa; both decline with time in semen samples. As a result, the accuracy of the assay may be influenced by the age of the semen sample. The presence of antioxidants in semen samples will also influence this time- and energy-dependent ROS generation.

For clinical purposes, it is important to have a reliable and reproducible method of ROS measurement. Also, strict quality control must be observed for this assay to be valid in a clinical laboratory setting. The objectives of this study were to assess the intra-assay, interassay, intraobserver, and interobserver variability of the luminol-dependent chemiluminescence assay and to standardize the administration guidelines by determining the optimal semen age and concentration.

**Materials and Methods**

**Subject Selection**

This study was approved by the Institutional Review Board of our hospital. Normal healthy men were screened and 10 of them were selected on the basis of normal semen analysis according to guidelines of the World Health Organization (WHO, 1992). A total of 28 donor semen samples were used to evaluate the characteristics of laboratory variability in measuring reactive oxygen species (ROS).

**Semen Collection and Assessment of Semen Variables**

Semen specimens were collected by masturbation after 48 to 72 hours of sexual abstinence. The specimens underwent complete liquefaction at 37°C for 20 minutes, and 5 μL of each specimen was loaded on a 20-μL Microcell chamber (Conception Technologies, San Diego, Calif) where it was analyzed for sperm concentration and motion characteristics using a computer-assisted semen analyzer (Cell-Track, version 4.24; Motion Analysis Corporation, Palo Alto, Calif).

**Quantitation of White Blood Cells**

The presence of white blood cells (WBCs) in each of the specimens was assessed using myeloperoxidase staining (Endtz test). A 20-μL volume of liquefied specimen was placed in a 2.0-mL cryogenic vial (Corning Coster Corp, Cambridge, Mass); 20 μL of phosphate-buffered saline (PBS; pH 7.0) and 40 μL of benzidine solution were added. The mixture was vortexed and allowed to sit at room temperature for 5 minutes. The peroxidase-positive WBCs turned brown, and these cells were counted in all 100 squares of the grid in a Makler chamber (Sefi Medical, Haifa, Israel). The results after correction for dilution were reported as 10$^6$ WBC/mL.

**Semenal Chemiluminescence Measurement**

After liquefaction, semen specimens were processed for ROS measurement. Briefly, samples were centrifuged at 300 × g for 7 minutes and the seminal plasma was removed. The sperm pellet was suspended in 3 mL of Dulbecco PBS solution (Irvine Scientific, Santa Ana, Calif) and washed again at 300 × g for 7 minutes. The sperm concentration was adjusted to 20 × 10$^6$/mL before ROS measurement. ROS formation was measured by a chemiluminescence assay using 5 μL of luminol (5 mM, 5-amino-2, 3-dihydro-1, 4-phthalazinedione, Sigma Chemical Company, St Louis, Mo). The luminometer measured chemiluminescence in the integration mode at 37°C for 15 minutes after luminol was added. ROS production was expressed as counted photons per minute (cpm) × 10$^6$. The study design for various measures is shown in Figure 1.
Figure 1. Flow chart demonstrating the measurement of interassay, intra-assay, interdonor, intraobserver, and interobserver variability in the amount of ROS production.

Intra-Assay Variation

To measure the intra-assay variation, a blank tube containing only 400 µL of PBS buffer and an assay control containing 5 µL of luminol working solution and 400 µL of PBS buffer were used. All the samples were measured in duplicate. To obtain the ROS level, the average control reading was subtracted from the average of the actual test values (i.e., sample luminescence = average sample reading – average background reading).

Interassay Variation

Interassay variability was evaluated by measuring ROS from the same assay on 3 separate days. In addition, a separate evaluation of temporal variation (intradonor variability) of ROS within donors was evaluated by measuring 3 different specimens of each donor during a 3-week period as described previously. The ROS measurements were performed using a Berthold (Autolumat LB 953; Wallac Inc, Gaithersburg, Md) luminometer. To establish the degree of variability related to the photomultiplier cell, the background and the luminescence changes due to the air-water phase transition were measured. The basal luminescence signal was measured simultaneously for 15 minutes on 3 different days using 7 empty polystyrene tubes and 7 tubes filled with deionized water. To assess the relevance of the amount of light present in the environment, the basal chemiluminescence was measured in 5 tubes previously exposed for 2 hours to a common light source and in 5 tubes stored in the dark.

Interobserver and Intraobserver Variation

The analysis of the sources of variability illustrated in Figure 1 was evaluated with 108 ROS measurements from 6 donors. In all the specimens, ROS production was measured in duplicate simultaneously by 3 different observers. The interobserver variation was obtained by analyzing the differences in the results produced by the 3 observers. The intraobserver variation was obtained by analyzing the differences between measurements of each observer.

Effect of Time on ROS Production

To assess the effect of time (age of the sample) on ROS production, 12 semen samples were analyzed for semen characteristics. Each sample was mixed with PBS (pH 7.4) and split into 4 equal aliquots. These samples were processed and analyzed for ROS production at 60, 120, 180, and 240 minutes after the samples were collected.

Effect of Washed Cellular Segments of Whole Semen on ROS Production

Ten semen samples were analyzed to assess the effect of washed cellular segments of whole semen on ROS production. After liquefaction, the sample was washed and the pellet was suspended in 1 mL of PBS buffer. The sperm concentrations were adjusted to range from 1 to $120 \times 10^6$ spermatozoa/mL. Each concentration was tested at the same time in duplicate.

To standardize the relationship between the sperm concentration and ROS levels, the results of each concentration were expressed as $10^6$ cpm/20–$10^6$ spermatozoa. The adjusted ROS value was as follows:

$$\text{Luminescence (cpm)} \times 20 \times 10^6$$

$$\text{Sperm concentration in the sample} = \text{ROS level}/20 \times 10^6 \text{ spermatozoa}$$

The differences between the actual value and the standardized value were calculated.

Statistical Analysis

Variance components were calculated using random effects analysis of variance (ANOVA) to compute reliability (interclass correlation). As opposed to “fixed effects ANOVA,” random effects ANOVA estimates the components of variance associated with each source of variability (e.g., between observers) and determines whether this variability is significantly greater than zero. Reliability is computed as the ratio of intra-assay variability to the total variability. Therefore, reliability measures near 100% indicates that almost all of the observed variability is associated with between sample differences, and not created by different observers, days, replications, etc. Comparisons between various background levels were tested using Student’s $t$ test, and a paired $t$ test was used to compare ROS measurements between different ages of the samples. The relationship between sperm
concentration and ROS, which was quantified as $R^2$ (the percentage of variance explained) within each sample, was calculated using linear regression analysis. Statistical significance was assessed at $P < .05$ with 2-tailed tests. The summary statistics are presented as means ± standard deviations (SDs). Data were analyzed by the SAS statistical software package (version 6.12; SAS Institute Inc, Cary, NC).

**Results**

**Intra-Assay and Interassay Variation**

The background luminescence was $2.36 ± 0.33$ cpm $\times 10^5$ when it was measured in deionized water alone. No differences in the background luminescence were seen in the empty tubes when air was used as the interface ($2.37 ± 0.37$ cpm $\times 10^5$) versus deionized water ($P = .34$). It is important to subtract daily background values from observed measurements to eliminate any potential variation. The photomultiplier cell coefficient of variation was 1.6%. Light exposure increased the background luminescence from $3.04 ± 0.03$ to $3.27 ± 0.05$ cpm $\times 10^5$ ($P < .001$). However, no difference was seen when the assay was performed in total darkness. The interassay SD was 0.005, which translated into an interassay reliability of 93.8% (coefficient of variation [CV] = 2.0%). The ROS level in the donor samples was $0.22 ± 0.08$ cpm $\times 10^6$, with a minimum of 0.02 cpm $\times 10^6$ and a maximum of 0.49 cpm $\times 10^6$. The interassay SD was 0.18 cpm $\times 10^6$; therefore, the coefficient of variation was 83.2%. This level of variability among donors was desirable as it allowed examination of the properties of measurement over a range of ROS levels.

**Effect of Time on ROS Production**

ROS levels decreased significantly with time; ROS production dramatically declined in the semen specimens that were more than 60 minutes old ($P < .001$). ROS values decreased by 31% at 120 minutes and 62% at 180 minutes compared with the 60-minute-old specimens (Figure 3).

**Effect of Washed Cellular Segments of Whole Semen on ROS Production**

Examination of relationships solely within individuals showed a linear relationship between ROS production and sperm concentration in 8 of 10 samples. However, due to
Figure 4. Overall relationship between sperm concentration and ROS levels in 10 donors. The within-donor coefficient of linearity was 0.993 (0.962–0.996); poor linearity was seen in 2 of the 10 samples analyzed (\(R^2 = 0.003\) and \(0.371\)).

Table 1. Effect of sperm concentration on ROS production indicating linearity (\(R^2\)) in 10 different semen specimens and their median and interquartile range (25th and 75th percentile)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Within Sample (R^2)</th>
<th>Mean Sperm Concentration, (\times 10^6/\text{mL})</th>
<th>Mean ROS, (\times 10^6) cpm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample 1</td>
<td>0.9878</td>
<td>51.6</td>
<td>0.118</td>
</tr>
<tr>
<td>Sample 2</td>
<td>0.9620</td>
<td>128.2</td>
<td>0.027</td>
</tr>
<tr>
<td>Sample 3*</td>
<td>0.0302*</td>
<td>51.3</td>
<td>0.130</td>
</tr>
<tr>
<td>Sample 4</td>
<td>0.9922</td>
<td>76.4</td>
<td>0.051</td>
</tr>
<tr>
<td>Sample 5</td>
<td>0.9959</td>
<td>64.6</td>
<td>0.037</td>
</tr>
<tr>
<td>Sample 6</td>
<td>0.9958</td>
<td>32.3</td>
<td>6.25</td>
</tr>
<tr>
<td>Sample 7</td>
<td>0.9980</td>
<td>22.8</td>
<td>3.088</td>
</tr>
<tr>
<td>Sample 8</td>
<td>0.9970</td>
<td>33.1</td>
<td>11.044</td>
</tr>
<tr>
<td>Sample 9</td>
<td>0.9945</td>
<td>64.6</td>
<td>0.037</td>
</tr>
<tr>
<td>Sample 10*</td>
<td>0.3709*</td>
<td>139.5</td>
<td>0.063</td>
</tr>
<tr>
<td>Median</td>
<td>0.9934</td>
<td>58.1</td>
<td>0.090</td>
</tr>
<tr>
<td>25th percentile</td>
<td>0.9620</td>
<td>37.6</td>
<td>0.041</td>
</tr>
<tr>
<td>75th percentile</td>
<td>0.9959</td>
<td>73.5</td>
<td>2.348</td>
</tr>
</tbody>
</table>

* Samples showed a lack of "within subject" linear relationship between concentration and ROS levels. ROS indicates reactive oxygen species; cpm, counted photons per minute.

Discussion

Given the growing evidence of the relationship between high oxidative stress and male infertility, and the clinical usefulness of the luminol-dependent chemiluminescence assay, it is important to ensure that the assay is reliable and accurate. In the present study, normal healthy donors were used as per WHO guidelines. We excluded all donors who demonstrated leukocytospermia (WBC levels of \(> 1 \times 10^6\) WBC/mL). A simple wash-and-resuspension technique was used in preparing the specimen for ROS measurement. However, this was not particularly true for the 2 patients with poor correlation between sperm concentration and ROS levels.

Increased by expressing the results as log (ROS + 1), or loading the leukocytic samples on a 2- or 3-layer density gradient media to remove the WBCs from the spermatozoal population. This will provide a highly motile sperm population devoid of leukocytes, and complete removal of the contaminating WBCs can be ensured by final separation using anti-CD45 coated paramagnetic (M-450) beads and confirming it by formyl-methionyl leucyl phenylalanine (FMLP) stimulation test. An FMLP test is specific to the receptors present only on the leukocytes and not on the spermatozoa. The absence of a positive response to FMLP stimulation can confirm the absence of WBCs in the fraction. The present study shows that the assay has a high degree of reproducibility and accuracy.
The small variations observed between the assays and between the observers confirm the reliability of this assay. Earlier reports (Hipler et al, 1998), have shown a close interassay and intra-assay correlation in the coefficients of variance when luminol and lucigenin were used. The Hipler study reported an intra-assay CV of 13% and an interassay variability of 15%. Similar results have been published in other studies (Aitken et al, 1989b). Our results show a clear improvement in the variability; our intra-assay variation was 0.5% and interassay variation was 2.0%. This improvement in the CV may be due to the standardization of the assay and because we minimized the age of the semen sample. Another possible explanation is that all our donors had low ROS levels.

One of the most important sources of variability in the ROS value is the age of the sample when it is analyzed. Our data suggest that 33% of the total variability of the assay was related to the time when the actual analysis was done. The metabolic pathways related to the ROS production have been explained in several reports, indicating that ROS production is an energy-dependent process. Our results indicate that ROS levels rapidly decrease with time. A continuous decline in sperm viability in association with a reduction in the amount of available substrate (O2) may explain this finding. The samples in our study were washed and resuspended in PBS. We did not examine the motility and the viability of the specimens at different time intervals in this study. Time-dependent (as short as 15 minutes) decrease in sperm motility was reported by Bell et al (1992) in specimens treated with hydrogen peroxide (0.01% and 0.5%), or with a xanthine-xanthine oxidase system for ROS production (de Lami-rande and Gagnon, 1992a, 1992b). These investigators reported a higher decrease in motility compared to little (10% to 20%) change using human spermatozoa, and a 13% decline in viability using mouse spermatozoa (Bairdi et al, 1997), or no change using equine spermatozoa (Baumber et al, 2000) under these incubation conditions. Sperm motility is a more sensitive indicator of oxidative stress. We did not study the time-dependent change in pH. Our results show that to maintain the reliability of the assay, the samples must be analyzed within the first hour after the specimen has been collected. In addition, the assay should be conducted in dark because room light affects the chemiluminescence.

Although we selected semen samples based on WHO criteria, the variation seen between subjects was extremely high in our study. Different ROS levels have been shown among subsets in samples fractionated by a Percoll gradient (Aitken et al, 1989a). This finding supports the theory that not all spermatozoa in a given sample have a similar capacity for ROS production. Chemiluminescence therefore depends on the percentage of spermatozoa that have a different capacity to produce ROS. Regardless of the high variation found between the subjects, a strong linear correlation “within-subject” was seen between the sperm concentration and the ROS levels in 83% (10 of 12) of the samples analyzed. It is remarkable to mention that in the 2 samples that did not show any linear relationship, low chemiluminescent signals were seen. Semen characteristics show a large variation within the same individual. We have also demonstrated physiological variation in ROS values within a given donor. A repeat specimen should be obtained for a donor showing large variation. If a specimen that tested negative for ROS becomes positive at a later time interval, it may be because of an underlying infection or other causes. If this occurs again on repeat evaluation, then that specimen should not be considered as a normal sample even though it may have normal semen parameters according to WHO criteria. To integrate this assay into the clinical laboratory, it is important to establish the range of ROS in a normal healthy population, using either a simple wash-and-resus pend method or sperm separation on a density gradient. If a sample is positive for ROS generation, then a repeat specimen from the patient at a short interval could be used to confirm the result. These 2 results may be averaged if the interval of collection between the specimens is short (ie, less than a week). This finding will enable the assay to be done in men with different clinical problems, such as oligospermia and hypospermia. Moreover, this will not only enable ROS to be measured accurately in patient specimens with low sperm concentrations, but it will also offer the possibility of standardizing both the procedure and the results in order to establish the normal values for the assay.

Despite the strong correlation between oxidative stress and male infertility, there is a lack of consensus in the literature regarding the normal values of ROS levels in human spermatozoa. The standardization of the assay is the first basic step to reach a consensus, and our study demonstrated a satisfactory correlation of the results with minimal differences in the actual values when expressed as $20 \times 10^6$ spermatozoa. On other hand, our study was limited in terms of the sensitivity of the assay, especially at very low sperm concentration (<5 $\times 10^6$/mL). The lack of correlation between sperm concentration and ROS levels in 2 patients with low chemiluminescent signals demonstrates that the assay is unable to measure low ROS levels. In patients with low chemiluminescence, the differences between the background luminescence and the sample values are related only with the intra-assay variation. In patients with low sperm concentrations and low luminescence (<0.1 $\times 10^6$ cpm), expressing the result as $20 \times 10^6$ spermatozoa may falsely elevate ROS values. We recommend that the ROS results be expressed as zero in such cases.

In conclusion, our results demonstrate that the luminol-
dependent chemiluminescence assay for ROS measurement is both accurate and reliable when the sperm concentration is greater than \(1 \times 10^6/\text{mL}\), and the samples are analyzed within the first hour after specimen collection.

**Acknowledgment**

Robin Verdi provided secretarial service and Karen Seifarth and Cheryl Wellstead assisted in the recruitment and scheduling of donors for the study.

**References**


