Dragon, Karen E. (CDC/NIOSH/EID)

LTFE Data
nalysis10-16-07.pdf

Sorry I missed that. Here is the corrected file.

-----Original Message-----
From: NIOSH Docket Office (CDC) [mailto:niocindocket@cdc.gov]
Sent: Tuesday, October 16, 2007 9:25 AM
To: Patty Scheib
Cc: Dragon, Karen E. (CDC/NIOSH/EID)
Subject: RE: Long-Term Field Evaluation (LTFE) NIOSH Docket Number (NIOSH - 101)

Ms. Scheib: The word "confidential" is still on pages 2-4. Is it ok for me to wite-out the word "confidential" or do you want to resend this submission without the word "confidential" on these three pages? Just let me know. Thanks,

Karen E. Dragon
NIOSH Docket Office Assistant
513/533-8303
ked2@cdc.gov

-----Original Message-----
From: Patty Scheib [mailto:pas@CSECorporation.com]
Sent: Thursday, October 11, 2007 5:57 PM
To: NIOSH Docket Office (CDC)
Cc: Dragon, Karen E. (CDC/NIOSH/EID); Scott Shearer
Subject: RE: Long-Term Field Evaluation (LTFE) NIOSH Docket Number (NIOSH - 101)

I have removed the "Confidential" from the document. These were intended to be for public review. This is the complete submission that I will resend. The first attachment has not been changed.

Thank you for bringing this to my attention.

-----Original Message-----
From: NIOSH Docket Office (CDC) [mailto:niocindocket@cdc.gov]
Sent: Thursday, October 11, 2007 9:51 AM
To: Patty Scheib
Cc: Dragon, Karen E. (CDC/NIOSH/EID)
Subject: RE: Long-Term Field Evaluation (LTFE) NIOSH Docket Number (NIOSH - 101)

Ms. Scheib: In reviewing the comments that you submitted I noticed at the bottom of the document LTFEDataAnalysis.pdf (Inter-office
correspondence) it is marked "CSE Corporation - Company Confidential" .
Since these comments will be posted on the NIOSH Docket website, I was not sure if this document was to be treated as confidential or if I have your permission to post them as comments submitted from you on behalf of CSE. Please let me know how you want me to handle this document. Thank you ....

Karen E. Dragon

1

NIOSH Docket Office Assistant
513/533-8303
ked2@cdc.gov

-----Original Message-----
From: Patty Scheib [mailto:pas@CSECorporation.com]
Sent: Monday, May 21, 2007 1:39 PM
To: NIOSH Docket Office (CDC)
Subject: FW: Long-Term Field Evaluation (LTFE) NIOSH Docket Number (NIOSH - 101)

I am seeking acknowledgement that this e-mail was received.

Thanks,

-----Original Message-----
From: Patty Scheib
Sent: Friday, May 18, 2007 6:08 PM
To: 'nioshdocket@cdc.gov'
Cc: Ed Murray; Scott Shearer; Tom Barrage
Subject: Long-Term Field Evaluation (LTFE) NIOSH Docket Number (NIOSH - 101)

Patricia A. Scheib
QC Manager
CSE Corporation
412-856-9200 X-1200

CSE

— Safety Works —————————————————————————————————

MEMO

TO: NIOSH Docket Office

FROM: Pat Scheib

DATE: 5/18/07

SUBJECT: Comments on the DRAFT Long Term Field Evaluation Program Concept; NIOSH Docket Number NIOSH-101

CSE wishes to commend NIOSH on program improvements that it intends to institute, particularly in the area of statistical design to the sampling program. We welcome this opportunity to offer comments and discuss the Long Term Field Evaluation ("LTFE") program and hope that this will lead to a continuing dialogue that will result in better programs and ultimately improve understanding of how our products perform in the field. We wish to offer the following comments on the draft documents. Comments are organized into 4 sections: 1) simulator instrumentation, 2) measurements and data collection, 3) the QA system to support the data collection, and 4)-7) comments specific to the draft documents themselves. CSE generally supports NIOSH's approach with respect to the other points in its Program Concept.

COMMENTS

1) Simulator Instrumentation

Since the proposed LTFE relies exclusively on the ABMS for data collection, it is important to verify the performance and accuracy of the simulator to ensure that data collected will be valid and reproducible.

In order to provide for complete transparency of the LTFE process and to enable others to conduct their own scientifically valid tests to reproduce the LTFE test results, SCSR equipment manufacturers should be provided with the design specifications, performance parameters, and configuration for the NIOSH ABMS. We cannot emphasize enough that having access to this information is essential to providing for a transparent, accurate and fair process of evaluating SCSR performance for the mining community.

CSE also suggests that NIOSH sponsor a forum to facilitate information exchange between NIOSH and all SCSR manufacturers on what appear to be certain idiosyncratic effects of the current ABMS configuration and to discuss options to resolve those issues.

Of particular importance when testing a chemical–generation-based breathing apparatus, such as the SR-100, is the water content of the gas that is used in the testing. Variation in the water content affects oxygen concentrations as water vapor is a critical component of the chemical reaction that generates the oxygen in the breathing apparatus. If wet bulb and dry bulb temperatures are to continue to be the indicator of water content of the air, then the wet bulb thermometer device must be subjected to scrutiny as to its capability, accuracy and precision. CSE feels that wet bulb response is not a fundamental property, but, rather, is specific to the system in which it exists. For example, wet bulb response will vary with different flow rates or amounts of water on the thermocouple.

CSE

Because of this, we expect that it will be difficult to standardize humidity responses between simulators using such a device. Please consider that manufacturers in the industry may wish to correlate their simulators to that used by NIOSH. The design used by NIOSH must therefore take into account the need for other simulators to be able to replicate NIOSHs results. For example, are these parts available for purchase? We have identified a fast response sensor that may be a suitable substitute for determining relative humidity of simulator systems that can be mass produced to industry specifications. We would be happy to provide you with more information on this alternative design.

Temperature control must also be examined when considering the relative humidity inside the simulator. Most of the ABMS system is not currently temperature controlled. For example, the water put into the Lung section is heated to 37° C. but the rest of the system is at a lower, variable, room temperature. Consider that the room temperature may vary 5° F. over a single day in a temperature controlled environment. Consider also the temperature variation within the system. This results in local areas of cooling and condensation resulting in droplets accumulating in the system as described in Procedure No. TEB-STP-XXX: at 5.18: "Blow out any condensed water from mouth port adapter sample lines" and page 5 at 5.27: "Drain mouth port adapter of condensed water." It is critical that temperature be accurately controlled during the testing as this will affect relative humidity inside the breathing circuit, which in turn can significantly affect performance of the breathing apparatus.

CSE also questions the assumption that the flow through NIOSHs ABMS simulator is constant over the 60 minute period required for the test. It must be remembered that all pumps are variable to some extent. NIOSH has not made public any information that would indicate the variability of the ABMS pump, however, the assumption that this rate is constant is the basis for calculations of the CO_2 values. CSE understands that the volume of the ABMS system is estimated by the assumption that the pump's flow rate is constant over a given period of time. This value is used to estimate a point on the CO_2 vs. time curve at which the start of inhalation is estimated to begin. Integration of the area under the CO_2 vs. time curve is used to give the CO_2 values. Since the CO_2 vs. time curve is rapidly changing at this point of switch-over from exhalation to inhalation, a small error in the selection of the point could result in a large variation in the CO_2 value calculated. Variation in the pump flow would result in variations of the time estimates taken from the ABMS system for this purpose of estimating the time to the detector. Therefore, pump variation could lead to higher CO_2 values because of the way in which the CO_2 values are integrated versus time.

A study of pump capability should be undertaken so that the error inherent in the calculation of CO_2 values is known. For example, if the time for the test gas to reach the detector was stored for multiple runs of the ABMS that have already been performed, this could serve as data to determine the variability of this time that is used to estimate CO_2 values. CSE recommends evaluation of how this variability affects the calculations of the CO_2. If it is determined that the variability found is of such a degree that it could result in inaccurate calculations of CO_2, then this must be addressed.

CSE also questions the gas exchange used by NIOSHs ABMS which introduces further unacceptable variability in the water vapor content of the breathing stream inside the ABMS. Gas is withdrawn from the system to simulate usage of oxygen and is then replaced by dry nitrogen that results in a change in temperature and humidity in the system. Dry gas will lower humidity in the breathing circuit immediately. This will retard the chemical reaction in the SCSR that is necessary to generate oxygen. If the temperature of the added gas is lower, the result could be a loss of water vapor due to condensation at the local site. The main effect would be from the introduction of dry gas at this step. NIOSH must control the water vapor content and temperature of the gas being introduced to the ABMS during gas make up operations for chemical oxygen generator SCSR units to be fairly evaluated. This is an area where ABMS conditions diverge greatly from the conditions during a trial with a human subject.

2) Measurements and Data Collection

CSE strongly endorses NIOSHs commitment, as stated in the proposal, to ensuring that LTFE testing is performed only on equipment that passes the manufacturer's inspection criteria: "NIOSH is using strict inspection procedure on collecting samples. No damaged units arguably failing manufacturers inspection criteria will be sampled." CSE has noted that, in past LTFE testing,some of the units tested did not meet inspection criteria.. This resulted in inaccurate test results. By law, if units do not pass inspection criteria, they are, no longer approved devices and must be removed from service. No valid data can be obtained from testing units in the LTFE that do not pass inspection criteria.

In the past, data collected during ABMS testing was processed and the raw data points used to create the test reports were not available to SCSR manufacturers. CSE strongly believes that all raw data from the testing should be retained to facilitate further analysis of test results and verification of the test methods when unexpected results occur. Such raw data should be made available to the manufacturer.

The temperature and moisture sensing elements of the ABMS should be calibrated as a part of process control. This is of particular interest to CSE since the performance of the oxygen generating SCSR is highly dependent on the water vapor content of the air in the system. Since moisture is critical to the performance of the chemical SCSR, the control of moisture must be assured.

Likewise, the moisture content of the air internal to the ABMS during tests must be closely monitored and controlled because moisture content is highly influenced by temperature. Please note, page 3 Procedure No. TEB-STP-XXX: at 5.18 "Blow out any condensed water from mouth port adapter sample lines" and page 5 at 5.27 "Drain mouth port adapter of condensed water. " Both procedures indicate that moisture is not well controlled or measured on the current ABMS apparatus. Running the ABMS in this way would not provide for accurate testing of an oxygen generating SCSR.

The LTFE Program Concept states on page 1: "Since that time, operation of the automated breathing and metabolic simulator (BMS) has been improved and the correlation to human performance established: further development of the correlation to human test subjects as part of the LTFE is no longer required." Later on page 3 the Program Concept states, "Limited human treadmill testing will be employed to establish a baseline correlation between certification testing, LTFE BMS testing and LTFE human subject testing." CSE is concerned that the correlation between the ABMS testing and LTFE human subject testing has not been established and that these differ in statistically significant ways, such as in CO_2 concentrations and dry bulb temperatures.
Refer to Sean Tremba Memo of May 9, 2007

CSE statistical analysis of the data provided on NIOSH's website suggests that there is greater variability in the machine test results than in the human subject tests. This is the opposite of what one would ordinarily expect. CSE's comparison of means indicates that NIOSH has not proven that the simulator is equivalent to a human test subject wearing an SCSR. Clearly, further testing and analysis is needed. For example, note the variability from run to run of the ABMS on minimum CO_2 concentrations and also on the inhalation dry bulb temperatures. The conclusions from the analysis performed by CSE on the ABMS test results versus the human test results under the LTFE conditions demonstrate that the two testing methods are not equivalent and that changes will need to be made in order to achieve equivalence.

One of the values that was found to be statistically different was the minimum inhaled CO_2. It is interesting to note that the early report on the simulator IC 9110, Bureau of Mines Information Circular; 1986, "Development of an Automated Breathing Metabolic Simulator", stated on page 15, "Two possible weak points in the DEEC Co design are the indirect control of metabolic flow rates and the sensitivity of the electronic calculation of average inhaled gas concentration.. . . . The calculation of average inhaled gas concentrations depends upon correct measurement of gas transport and response time that are used to delay

the integration of the area under the figurative gas concentration curves. If the gas transport or response times, which are measured in ms, change due to turbulence in the sample lines, the measured values will be incorrect." The discussion goes on to discuss sources of turbulence. CSE believes that turbulence could be (1) inherent in the system, (2) caused by water droplets forming in the lines (3) caused by introduction and/or (4) caused by removal of gas that is required to simulate respiration in the ABMS.

Current data on this correlation needs to be subjected to further statistical study to determine whether and how the two methods, ABMS and human treadmill at constant rate, are similar or different. If key differences are confirmed, changes must be made to the ABMS to correct variability. This exercise could pave the way to correlation of the two breathing simulators that are planned for use at NIOSH to support this program. In addition, CSE sees that such an analysis could be applied to CSE simulators or any and all simulators in use in the industry to further standardize the simulator testing. To assist in this effort, CSE has provided to NIOSH its analysis of minimum CO_2 concentration and inhalation temperatures, etc. that we have performed (Memo of May 9, 2007 attached to this report).

CSE is also concerned that the human test data represent too few participants to draw significant conclusions. NIOSH's Program Concept does not provide the number of tests used for the correlation thus far. The statement on page 3 of the LTFE Program Concept document seems to indicate that additional testing is planned. "Limited human treadmill testing will be employed to establish a baseline correlation between certification testing, LTFE BMS testing and LTFE human subject testing." Additional information is needed on the number of test subject runs that have contributed to the correlation to this point.

We also note that IC 9110, Bureau of Mines Information Circular; 1986, "Development of an Automated Breathing Metabolic Simulator", describes the capabilities of the DEEC ABMS on page 15 as follows, "The metabolism and breathing simulation of this simulator overall have proved accurate to within 5% of desired values." This statement needs to be further evaluated and explained in this rulemaking. For example, does it apply to all measurements or only to particular measurements made by the ABMS? What is the current level of capability/variability of the ABMS? See also the first paragraph under section 5 below.

3) Quality Assurance System

CSE enthusiastically supports NIOSHs stated intention of developing a quality system in accordance with ISO 9001 and ISO/IEC 17025 as was discussed at the NIOSH-NPPTL Respirator Manufacturers Meeting on April 10, 2007. The activities necessary to reach the level of these standards will assure that the laboratory and testing activities will produce technically valid data and results. The use of the International Standard will facilitate cooperation between your testing laboratory and other stakeholders such as the manufacturers, mines and miners. We believe it will assist in the exchange of information and the harmonization of standards and procedures. CSE notes that the current QA document referred to in Procedure NO. TEB-STP-XXXX at step 4.2, NIOSH Manual of Analytical Methods, Section C, Quality Assurance, has not been revised since 1/15/98. The IEC/ISO 17025 has been revised in 1999. A review of these procedures would be appropriate at this time.

NIOSH should specify what QA program will support the validity of the LTFE data. Refer to the NIOSH Laboratory Quality Assurance Program, page 10 at 4. QA Measurement, which states "Certain quality control checks should be performed with each sample set to further support the reported results on actual field samples."
Are any additional checks, blanks or standards, or duplicate samples considered for addition to this method? Some checks are a part of the Draft Procedure No. TEB-STP-XXX: at 5.9 to 5.14. More are needed.

NIOSH Laboratory Quality Assurance Program on page 15 at 8. Instrument Maintenance states: "Performance checks can be useful in documenting instrument performance over time and in detecting deviations. . . . An instrument's maintenance should be maintained in an 'instrument log book' and kept near that instrument or in a recognized location." Will a paper log book be kept? Will any computer files be a part of this record? We suggest a computer based system for easy retention and protection of records. This is more in keeping with the current practice of all electronic records. IC 9110, Bureau of Mines Information Circular; 1986, "Development of an Automated Breathing Metabolic Simulator", on page 15 describes the ABMS as having capability for "Disk storage of complete tests." CSE is requesting complete data be retained from each test run. See also the last paragraph in this section.

We are told that the equipment has been changed on page 1 step 3.3.1 of Procedure No. TEB-STP-XXX: and also on page 1 of the LTFE Program Concept document: "When used in monitoring and measurement of specified requirements, the ability of computer software to satisfy the intended application shall be confirmed. This shall be undertaken prior to initial use and reconfirmed as necessary." Per ISO 9001:2000 at 7.6. AND "Records shall be maintained of each item of equipment and its software significant to the tests and/or calibrations performed. Records shall include at least the following: h) any damage, malfunction, modification or repair to the equipment. ISO/IEC 17025: 1999(E) at 5.5.5." We understand that NIOSH does not currently use an ISO or 17025 quality assurance system but since interest was expressed in pursuing these goals, we suggest this part be incorporated at the current time to substantiate the testing results.

With regard to the statement at Page 2 at 4.1 which reads: "Prior to beginning any testing, all measuring equipment to be used must have been calibrated in accordance with the manufacturer's calibration procedure and schedule," please supply or publish the list of equipment used on the ABMS testing that is calibrated, the calibration intervals and the entity that performs the calibration.

If the ABMS equipment has been modified since its manufacture, how does this change the frequency of calibrations? We are told that the equipment has been changed on page 1 step 3.3.1 of this document and also on page 1 of the LTFE Program Concept document. SCSR manufacturers should be kept informed of all modifications to NIOSHs ABMS. In this regard, please see comments in the first section of this memorandum.

As stated in the Procedure No. TEB-STP-XXX: page 2 at 4.2 refers to the NIOSH Manual of Analytical Methods. This document describes a Laboratory Quality Assurance Program. Some valuable practices were found in this document dated 1/15/98. On page 8 "Each laboratory must set its own operating procedures and quality control practices, and document them in a Quality Assurance Manual." Please make this QA manual for the LTFE laboratory available for manufacturers.

Several other provisions of the NIOSH QA Manual are particularly relevant here: "In order to support the original data, an effective, complete, record-keeping system must be maintained. . . . Furthermore, if the appropriate quality control checks were performed with the analysis and documented, there can be no doubt regarding results." NIOSH Laboratory Quality Assurance Program on Page 9 at b. Sample Clerk, "The sample clerk's functions . . . receipt and log-in of sample. Field sample should be stored in a secure location under proper conditions (temperature, etc.) until analysis. Logging and tracking of samples in the laboratory is important so that history of these samples can be documented and processed in a timely manner. The sample clerk may also be given the responsibility of maintaining chain of custody documentation."

Page 15 at step 10 NIOSH Manual of Analytical Methods is also relevant. Quality Assurance Records: "Computerized record keeping systems should be backed up periodically. Archive copies of computer data require specialized storage conditions and these archive copies may not be reliable for extended periods of

time. Filing of records should be current and accurate. If rapid retrieval of data is not possible, then maintenance of quality assurance records loses its purpose." See also Page 7 Step 7 Record and Test Sheets: below.

4) Long Term Field Evaluation Program concepts

NIOSH proposes to institute a new testing requirement based on the ABMS test results stating "100% immediate corrective action (fix, user notice, recall, or rescind approval) is required." However, equivalence of the LTFE test parameters to those used for approval certification testing has not been established. Is NIOSH proposing to change the objective of the LTFE from research project to approval enforcement? CSE questions the appropriateness of doing so at this time as it has not been established that there is a direct correlation between the LTFE test parameters and those used for 42 CFR Part 84 approval testing. There also remain many questions, as discussed here, as to the accuracy of LTFE ABMS testing of oxygen generating SCSRs. Further, the values for the critical parameters noted in the LTFE Program Concept, specifically, for oxygen and CO2 are not the same as those required in approval testing.

5) Procedure No. TEB-STP-XXX: Determination of Capacity and Performance Tests of a SCSR using Automatic Breathing and Metabolic Simulator (ABMS) Standard Test Procedure.

Procedure No. TEB-STP-XXX: Page 2 at 4.2 Question: Has the precision and accuracy of the ABMS been determined? If so, where is this information published?

Procedure No. TEB-STP-XXX: On page 2 at step 5.3: The second sentence is incomplete.
Procedure No. TEB-STP-XXX: On page 3 at step 5.6. The flow rate is not completely stated.
Procedure No. TEB-STP-XXX: On page 4 at 5.24: There is a typographical error.
Procedure No. TEB-STP-XXX: On page 4 at 5.24.6: Please specify the units of pressure, vapor pressure, and temperature used in this equation.

Procedure No. TEB-STP-XXX: Page 7 Step 7 Record and Test Sheets: CSE suggests that records of data collected and calculated by the computer (one minute averages are the basis of the evaluation of stressors as per page 6 at step 6.3.1) for all runs is pertinent to supporting the validity of individual tests and of the LTFE testing program, since these will be the basis of investigations into product failures and shortcomings. The bulk of data for the testing is now computer generated. Currently, this data is not retained. We feel this is a serious shortcoming of the testing procedure.

What procedure is followed if the ABMS experiences a failure during a test run? (For example, a loss of power due to burn out of a component or pump, a leak or loss of power to the facility.) How would this test data be handled? What other indicators could signal that the instrumentation has passed out of the optimal condition to report valid results? All of these possibilities need to be addressed.

Procedure No. TEB-STP-XXX: Page 8 at step 9. Record of Change: Is this a record of changes to this procedure or the configuration of the instrument? Or the software used in the instrument? The item is left blank in the draft.

Procedure No. TEB-STP-XXX: NIOSH Manual of Analytical Methods page 5, contains a classification of NIOSH methods. Will this method be classified as Full or Partial evaluation method?

6) Attachments Reporting Form Initial Report

A code number should be assigned to each of the reasons for non-acceptance. Include a code for "Unable to locate SN unit". Print the code list on the form.

What if more than one reason is noted for SCSR units not accepted for the test? Will these be recorded? If all reasons for non-acceptance are not recorded, in what order will the criteria of acceptance be checked? This could make a big difference in what reasons for failure are reported. Those checked first will get reported more often than those checked later, lending bias to the inspection results.

Will units that fail inspection be collected at the mine site and removed? Collecting and removing non-acceptable units could become a logistical problem. But leaving the units at the mine site could raise safety concerns.

7) (Feb. 20, 2007) Sampling Plan and Statistical Considerations for the LTFE Program

On Page 3 at 3. subpoint 3) Will manufacturers have access to the information resulting from this collection such as # units in the population, # units sampled or collected, % of mines participating and size of operations(# of SCSRs at the location)? This would be helpful to know.

An annual report could list some of these statistics including % units that fail inspection criteria, % units not found at the mine site, % units that fail critical criteria, etc. This report could also describe how the sampling plan was changed (or not changed) as a result of the number of SCSRs actually collected. Such a report could comment on how well the sampling plan was implemented during the sampling year.

Date: May 9, 2007

TO: Distribution
FRM: Sean Tremba
SUBJ: Statistical Analysis of Data from the NIOSH LTFE Phase 9 Report

Summary:

NIOSH has made the claim that it has been verified that their ABMS produces test results similar to those obtained by human subjects. To check that claim, I have performed a statistical analysis of the data reported by NIOSH for their ABMS tests on the SR-100 and the Ocenco EBA 6.5. These analyses refute the claim that the ABMS has been proven to accurately simulate a human test subject.

Discussion:

Data were obtained from the most recent available NIOSH LTFE report, namely the LTFE Phase 9 report. Analyses were performed on the SR-100 data for those parameters for which the greatest difference between the ABMS and the human subjects were observed, namely the inhaled dry bulb temperature and the minimum carbon dioxide concentration. Standard statistical procedures were used for the analysis, namely a t test assuming unequal variances for comparison of the mean values and an f test to compare the variances of the data sets. The results of these analyses are shown in the tables below:

T-Test for minimum inhaled CO_2:

	ABMS Tests	Human Tests
Mean	1.13	0.80
Variance	0.597	0.147
Standard Deviation	0.773	0.383
No. of Observations	87	7
Degrees of Freedom	10	
T-Statistic	1.978	
P-Value	0.0761	

F-Test for minimum inhaled CO_2:

	ABMS Tests	Human Tests
Mean	1.13	0.80
Variance	0.597	0.147
Standard Deviation	0.773	0.383
No. of Observations	87	7
Degrees of freedom	86	6
F-Statistic	4.072	
P-Value	0.0402	

T-Test for inhaled dry bulb temperature:

	ABMS Tests	Human Tests
Mean	46.7	40.1
Variance	16.66	7.81
Standard Deviation	4.08	2.79
No. of Observations	87	7
Degrees of Freedom	8	
T-statistic	5.766	
P-Value	0.00042	

F-Test for inhaled dry bulb temperature

	ABMS Tests	Human Tests
Mean	46.7	40.1
Variance	16.66	7.81
Standard Deviation	4.08	2.79
No. of Observations	87	7
Degrees of Freedom	86	6
F-statistic	2.134	
P-Value	0.1703	

The above analyses show that, at least for the SR-100, significant differences exist between the results obtained from human test subjects and simulations of human test subjects using the NIOSH ABMS. Both the t test and the f test work by first making the assumption that the data from the data sets under comparison have an identical distribution. This assumption, known as the null hypothesis, is then tested by utilizing the relevant statistical test to determine whether the data indicate that the null hypothesis should be rejected. The p-value indicated in the above tables is the probability that a difference at least as large as the observed difference would occur under the assumption that the two data sets are identically distributed. Therefore, a low value for this p-value indicates that the null hypothesis should be rejected and gives the confidence level for the conclusion that the data sets are different.

To compare means, a t-test assuming unequal variances is used. The t-statistic is calculated by subtracting one mean from the other and dividing by a value that represents a pooled standard deviation for the two data sets. The distribution of this statistic has been extensively studied and is well known, so a p-value can be calculated from this distribution. A p-value of 0.0761 was obtained for the CO_2 data. Typically a p-value of 0.05 or lower is used to indicate statistical significance. However, it is still possible to conclude that, to a confidence level of 92.39%, that the mean values of the inhaled CO_2 are different in the human tests than they are in the ABMS tests. This is contrary to the claim that there is no difference between the ABMS testing and the human testing.

Even more interesting is the result of the f-test on the CO_2 data. The f-statistic is calculated simply by taking the ratio of the variances of the two data sets. Like the t-test, this distribution has been studied and tabulated, allowing the determination of a p-value. It has been claimed that the ABMS testing will be more reproducible than human subject testing. However, the f-test on the CO_2 data indicates otherwise. The variance of the temperatures measured during ABMS testing is greater than that measured during human treadmill tests to a statistically significant degree. The p-value of 0.0402 indicates that a difference exists to a confidence level of almost 96%.

The analyses of the temperature data leads to a similar conclusion. The confidence level for asserting a difference in the variances is only 82.97%, less than that typically needed to state that a statistically significant difference exists, but suggesting that it has not been demonstrated that the ABMS test simulates human tests accurately. The p-value of 0.00042 for the t-test indicates that to a very high degree of confidence (99.958%), it can be claimed that the ABMS test produces results that are different from the results obtained by testing human subjects.

While the above data show that a significant difference between ABMS testing and human testing exists for the SR-100, it could be claimed that this difference is due to the SR-100 and not due to a real difference between the ABMS simulation and a human subject test. Therefore, an analysis of the data from LTFE Phase 9 obtained from testing the Ocenco EBA 6.5 was performed. The data analyzed were the inhaled oxygen content and the inhaled dry bulb temperature. Results from these analyses appear in the tables below:

T-Test for Ocenco EBA 6.5 Inhaled Dry Bulb Temperatures:

	ABMS Testing	Human Testing
Mean	45.0	37.8
Variance	2.71	2.25
Standard Deviation	1.64	1.50
No. of Observations	52	4
Degrees of Freedom	4	
T-statistic	9.248	
P-Value	0.00076	

F-Test for Ocenco EBA 6.5 Inhaled Dry Bulb Temperatures:

	ABMS Testing	Human Testing
Mean	45.0	37.8
Variance	2.71	2.25
Standard Deviation	1.64	1.50
No. of Observations	52	4
Degrees of Freedom	51	3
F-statistic	1.203	
P-value	0.517	

T-test for Ocenco EBA 6.5 Inhaled Oxygen Content:

	ABMS Testing	Human Testing
Mean	59.2	79.5
Variance	156.5	59.7
Standard Deviation	12.51	7.72
No. of Observations	52	4
Degrees of Freedom	4	
T-statistic	4.809	
P-value	0.0086	

F-Test for Ocenco EBA 6.5 Inhaled Oxygen Content:

	ABMS Testing	Human Testing
Mean	59.2	79.5
Variance	156.5	59.7
Standard Deviation	12.51	7.72
No. of Observations	52	4
Degrees of Freedom	51	3
F-statistic	2.623	
P-value	0.2331	

It should be noted that for both of these parameters, the mean values obtained from the ABMS testing are statistically significantly different to a very high confidence level (99.924% for the temperature data and 99.14% for the oxygen data). The small number of human tests performed makes the confidence level of this difference even more striking since a larger difference is required to obtain a given confidence level when the number of observations is smaller. It is possible that the lack of statistical significance for the variance difference in the oxygen data is caused by the small number of human tests. In other words, more human tests might demonstrate that a significant difference in the variance does in fact exist. There is no evidence of a variance difference in the temperature data.

It should be noted that both the t-test and the f-test work on the assumption that the underlying distribution of the data sets is a normal distribution. For most scientific work, such an assumption is typically justified. However, both of these tests are fairly robust in terms of dealing with deviations from normality. These tests give reasonably accurate p-values even when the deviation of the underlying data from normality is fairly large. Further, the data reported in the NIOSH LTFE reports are average values for the entire test. It can be shown mathematically that the distribution of the average of a number of values tends to normality as the number of values averaged increases. This is true regardless of whether the individual values are normally distributed. Therefore, average values have a distribution closer to normal than the individual measurements. Thus, it is reasonable to conclude that the t-tests and f-tests give reasonably accurate p-values for the data reported in the LTFE Phase 9 report.

Conclusion:

The claim that there is no significant difference between the test results obtained using the NIOSH ABMS and the results of human subject testing is refuted by statistical analysis of the data reported in the LTFE Phase 9 report. It is worthy to note that while this claim could be refuted by statistical analysis, standard statistical tests cannot be used to establish this claim. This is because the result of a statistical test is either that the null hypothesis should be rejected or that there is insufficient evidence to reject the null hypothesis. Therefore, even statistically insignificant observed differences between the ABMS data and the human test data do not indicate that no difference in the tests exists. Thus, even without the statistically significant differences that have been observed, it cannot be stated that the ABMS test results have been demonstrated to be identical to the human test results. Of course, the observed differences are sufficient to refute this claim. Further, because the data used in these analyses are the averages over entire test runs, the differences in parameters measured instantaneously would be even greater than the differences seen in the data reported in the LTFE Phase 9 report.

Distribution:
Scott Shearer, Tom Barrage, Pat Schieb, Ed Murray