LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-TR-678880

Integrated Data Collection Analysis (IDCA) Program-Statistical Analysis of RDX Standard Data Sets

M. M. Sandstrom, G. W. Brown, D. N. Preston, C. J. Pollard, K. F. Warner, D. N. Sorensen, D. L. Remmers, J. J. Phillips, T. J. Shelley, J. A. Reyes, P. C. Hsu, J. G. Reynolds

October 30, 2015

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Integrated Data Collection Analysis (IDCA) Program — Statistical Analysis of RDX Standard Data Sets

Mary M. Sandstrom[1], Geoffrey W. Brown[1], Daniel N. Preston[1], Colin J. Pollard[1], Kirstin F. Warner[2], Daniel N. Sorensen[2], Daniel L. Remmers[2], Jason J. Phillips[3], Timothy J. Shelley[4], Jose A. Reyes[5], Peter C. Hsu[6], and John G. Reynolds[6*]

[1]Los Alamos National Laboratory, Los Alamos, NM USA
[2]Indian Head Division, Naval Surface Warfare Center, Indian Head, MD USA
[3]Sandia National Laboratories, Albuquerque, NM USA
[4]Air Force Research Laboratory, Tyndall Air Force Base, FL USA
[5]Applied Research Associates, Tyndall Air Force Base, FL USA
[6]Lawrence Livermore National Laboratory, Livermore, CA USA

ABSTRACT

The Integrated Data Collection Analysis (IDCA) program is conducting a Proficiency Test for Small-Scale Safety and Thermal (SSST) testing of homemade explosives (HMEs). Described here are statistical analyses of the results for impact, friction, electrostatic discharge, and differential scanning calorimetry analysis of the RDX Type II Class 5 standard. The material was tested as a well-characterized standard several times during the proficiency study to assess differences among participants and the range of results that may arise for well-behaved explosive materials. The analyses show that there are detectable differences among the results from IDCA participants. While these differences are statistically significant, most of them can be disregarded for comparison purposes to assess potential variability when laboratories attempt to measure identical samples using methods assumed to be nominally the same. The results presented in this report include the average sensitivity results for the IDCA participants and the ranges of values obtained. The ranges represent variation about the mean values of the tests of between 26% and 42%. The magnitude of this variation is attributed to differences in operator, method, and environment as well as the use of different instruments that are also of varying age. The results appear to be a good representation of the broader safety testing community based on the range of methods, instruments, and environments included in the IDCA Proficiency Test.

The overall IDCA effort is funded by the Department of Homeland Security (DHS). The testing performers involved are Lawrence Livermore National Laboratory (LLNL), Los Alamos National Laboratory (LANL), Indian Head Division, Naval Surface Warfare Center, (NSWC IHD), Sandia National Laboratories (SNL), and Air Force Research Laboratory (AFRL/RXQL).

Keywords: Small-scale safety testing, proficiency test, impact-, friction-, spark discharge-, thermal testing, round-robin test, safety testing protocols, HME, RDX, statistical analysis.

IDCA Explosives Safety Testing of Homemade Explosives
Integrated Data Collection Analysis Program

1 INTRODUCTION

The IDCA Proficiency Test was designed to assist the explosives community in comparing and perhaps standardizing inter-laboratory Small-Scale Safety and Thermal (SSST) testing for improvised explosive materials (homemade explosives or HMEs) and aligning these procedures with comparable testing for typical military explosives[1]. The materials for the Proficiency Test have been selected because their properties invoke challenging experimental issues when dealing with HMEs. Many of these challenges are not normally encountered with standard military or industrial explosives, which are the materials that were used to develop the routine safety tests. For HMEs, to a large extent, the issues are centered on the physical forms and stability of the improvised materials. Details of the results of proficiency tests for the chosen materials are documented in IDCA reports—RDX first testing[2], RDX second testing[3], RDX testing comparison[4], $KClO_3$/sugar (separated with a 100 mesh sieve)[5], $KClO_3$/sugar (as received)[6], $KClO_3$/Dodecane[7], $KClO_4$/Dodecane[8], $KClO_4$/Al[9], $KClO_4$/Carbon[10], $NaClO_3$/sugar[11], PETN[12] and Methods[32].

Evaluation of the results of SSST testing of unknown materials is generally done as a relative process, where an understood standard is tested alongside the HME. In many cases, the standard employed is PETN or RDX. The standard is obtained in a high purity, narrow particle size range, and measured frequently. The performance of the standard is well documented on the same equipment (at the testing laboratory), and is used as the benchmark. The sensitivity to external stimuli and reactivity of the HME (or any energetic material) are then evaluated relative to the standard.

Most of the results from SSST testing of HMEs are not analyzed any further than this. The results are then considered in-house. This approach has worked very well for military explosives and has been a validated method for developing safe handling practices. However, there has never been a validation of this method for HMEs. Although it is generally recognized that these SSST practices are acceptable for HME testing, it must always be kept in mind that HMEs have different compositional qualities and reactivities than conventional military explosives.

The IDCA is attempting to evaluate SSST testing methods as applied to HMEs. In addition, the IDCA is attempting to understand, at least in part, the laboratory-to-laboratory variation that is expected when examining the HMEs. The IDCA team has taken several steps to make this inter-laboratory data comparison easier to analyze. Each participating laboratory uses materials from the same batches and follows the same procedures for synthesis, formulation, and preparation. In addition, although the Proficiency test allows for laboratory-to-laboratory testing differences, efforts have been made to align the SSST testing equipment configurations and procedures to be as similar as possible, without significantly compromising the standard conditions under which each laboratory routinely conducts their testing.

The first and basic step in the Proficiency test is to have representative data on a standard material to allow for basic performance comparisons. Class 5 Type II RDX was chosen as the primary standard, and Class 4 PETN was chosen as a secondary material. RDX was tested in triplicate several times throughout the IDCA Proficiency test. In this report all of the RDX results are analyzed to determine statistical differences among participants, average values, expected ranges, percent variability, dependence on method or environment, and possible causes for the differences that are observed

The testing performers in this work are Lawrence Livermore National Laboratory (LLNL), Los Alamos National Laboratory (LANL), and Indian Head Division, Naval Surface Warfare Center, (NSWC IHD), Sandia National Laboratories (SNL), and Tyndall Air Force Research Laboratory (AFRL/RXQL).

2 EXPERIMENTAL

General information. All samples were prepared according to IDCA methods on drying and mixing procedures[3,4]. The sample was dried in an oven at 60°C for 16 h, then cooled and stored in a desiccator until use. The RDX used in this effort is Class 5 Type II RDX and was obtained from the Holston Army Ammunition Plant batch # HOL89D675-081 and provided to the participating laboratories by IHD[5]. RDX Type II is made by the acetic anhydride (Bachman) process and generally contains ~ 10-wt % HMX as a by-product[14]. The composition of the HOL89D675-081 material was determined to be 90% RDX and 10% HMX by High Performance Liquid Chromatography[6]. Laser Diffraction (Light Scattering method using Microtracs Model FRA9200) gave a particle size distribution of 7.8 to 104.7 micron with a maximum at 31.1 microns[6,7]. The Military Specification for RDX Type II Class 5 is that a minimum of 97-wt % of the materials passes through a 325-mesh (44 μm[15]) sieve fraction[16]. More details on the characterization of this material are in the RDX Set 1[2] and Set 2 reports[17].

Table 1 summarizes the SSST testing conditions used by the participants. The SSST testing data for the individual participants was obtained from the following IDCA Data reports: Small Scale Safety Test Report for RDX (LLNL)[38], Small Scale Safety Test Report for RDX (second calibration) (LLNL)[33], Small Scale Safety Test Report for RDX (third in a series) (LLNL)[8], Small Scale Safety Test Report for RDX (4th in a series) revised for 1-kg striker data (LLNL)[9], RDX 50188_rev 0 (LANL)[39], 50188 I RDX Second Run (LANL)[34], 50188 P RDX Third time (LANL)[10], 50188 V RDX 4th Time (LANL)[11], RDX Report Run #1 (IHD)[40], RDX Report Run #2 (IHD)[35], RDX Report Run #3 (IHD)[12], RDX report Run #4 (IHD)[37], SNL Small-Scale Sensitivity Testing Report RDX (SNL)[36], and RDX first time (AFRL)[41].

Table 1. Summary of conditions for the analysis of RDX (All = LANL, LLNL, IHD, SNL, AFRL)

Impact Testing
1. Sample size—LLNL, IHD, SNL, AFRL, 35 ± 2 mg; LANL, 40 ± 2 mg
2. Preparation of samples—All, dried per IDCA drying methods[3]
3. Sample form—All, loose powder; LLNL, pressed
4. Powder sample configuration—All, conical pile; LLNL pellet also
5. Apparatus—LLNL, LANL, IHD, ERL Type 12*; SNL, AFRL, MBOM with Type 12 tooling*
6. Sandpaper—All (180-grit garnet); LANL (150-grit garnet); LLNL (120-grit Si/Carbide)
7. Sandpaper size—LLNL, IHD, SNL, AFRL, 1 inch square; LANL, 1.25 inch diameter disk dimpled
8. Drop hammer weight—All, 2.5 kg
9. Striker weight—IHD, SNL, AFRL, 2.5 kg; LANL, 0.8 kg; LLNL 2.5 and 1.0 kg
10. Positive detection—LANL, LLNL, microphones with electronic interpretation as well as observation; IHD, SNL, AFRL, observation
11. Data analysis—All, modified Bruceton and TIL before and above threshold; LANL, Neyer also

Friction analysis
1. Sample size—All, ~5 mg, but not weighed
2. Preparation of samples—All, dried per IDCA procedures[3]
3. Sample form—All, powder
4. Sample configuration—All, small circle form
5. Apparatus—LANL, LLNL, IHD, SNL, BAM; IHD, AFRL, ABL
6. Positive detection—All, by observation
7. Room Lights—LANL, SNL, on; LLNL off; IHD, BAM on, ABL off; AFRL, off
8. Data analysis—LLNL modified Bruceton (log-scale spacing) and TIL; LANL, modified Bruceton (linear spacing) and TIL; IHD, Neyer and TIL

ESD
1. Sample size—All ~5 mg, but not weighed
2. Preparation of samples—All, dried per IDCA drying methods[3]
3. Sample form—All, powder
4. Tape cover—LANL, scotch tape; LLNL, Mylar; IHD, SNL, AFRL, none
5. Sample configuration—All, cover the bottom of sample holder
6. Apparatus—All, ABL; LLNL, custom built*
7. Positive detection—All observation
8. Data analysis methods—All, TIL

Differential Scanning Calorimetry
1. Sample size—LLNL, LANL, IHD, AFRL, ~ <1 mg

2. Preparation of samples—LLNL, LANL, IHD, AFRL, dried per IDCA procedures[3]
3. Sample holder—LLNL, LANL, IHD, AFRL, pin-hole; LLNL, IHD, hermetically sealed
4. Scan rate—LLNL, LANL, IHD, AFRL, 10°C/min
5. Range—LLNL, LANL, IHD, AFRL, 40 to 400°C
6. Sample holder hole size—LANL, IHD, AFRL, 75 μm; LLNL 50 μm
7. Instruments—LANL, AFRL, TA Instruments Q2000; LLNL, TA Instruments 2920; IHD, TA Instruments Q1000*

Footnotes: *Test apparatus, *Impact:* LANL, LLNL, IHD—ERL Type 12 Drop Weight Sensitivity Apparatus, AFRL, SNL— MBOM modified for ERL Type 12 Drop Weight; *Friction*: LANL, LLNL, IHD, SNL—BAM Friction Apparatus, LANL, IHD, AFRL—ABL Friction Apparatus; *Spark:* LLNL, LANL, IHD, AFRL, SNL—ABL Electrostatic Discharge Apparatus, LLNL—custom-built Electrostatic Discharge Apparatus; *Differential Scanning Calorimetry:* LANL—TA Instruments Q1000, Q2000, LLNL—TA Instruments 2910, 2920, Setaram Sensys DSC, IHD—TA Instruments Model 910, 2910, Q1000, AFRL—TA Instruments Q2000.

Data Analysis Methods

Bruceton Up-Down Testing

The most useful way to characterize the sensitivity of a material is by measuring the parameters of the statistical distribution that describes its response to an external stimulus. The only way to measure these parameters is by testing at various stimulus levels and interpreting the sequences of various Go and No-Go events based on a statistical model. The method used often for explosives is the Bruceton Up-Down method developed in the 1940's [18]. In this method, the explosive is tested at some initial stimulus level. If a Go is observed, the stimulus level for the next test is decreased by one step but if a No-Go is observed, the stimulus level is increased by one step. This Up and Down step adjustment continues for a predetermined number of tests to build statistics for the reaction probabilities at a few levels near the mean. If the steps are evenly and linearly spaced with respect to the explosive's response, if the response is Gaussian, and if the step spacing is close to the standard deviation, then the statistical results can be analyzed with simple algebraic formulas to determine an estimated mean and standard deviation of the probed distribution.

Neyer D-optimal Testing

The Bruceton Up-Down method concentrates testing around the mean of the distribution (50% level) but does not provide an optimum determination of the standard deviation. An alternative method was developed by Neyer in 1994 [Neyer 1994] using a maximum likelihood approach that concentrates testing at the +/- 1σ levels. The test design is "D-optimal", meaning that it maximizes the determinant of the information matrix associated with the results. In practice, the testing is carried out via commerical software that carries out analysis and test level changes during testing. The software fits a Gaussian distribution to the final set of test results, providing an estimate of the mean and standard deviation.

Threshold Initiation Level Testing

Threshold Initiation Level testing, or TIL determination, is defined in this context as the method of determining the highest stimulus level at which some predetermined number of No-Go events are observed without any Go events occurring. In the data sets below the predetermined number of No-Go events is often 10 although occasionally 20 is used. Practically, testing is carried out at a chosen level until either a Go event occurs or the predetermined number of No-Go events is reached. If a Go occurs, the stimulus is decreased one step and the testing is repeated. If the result was all No-Go events, the stimulus is increased one step and testing is repeated. The "TIL" level or "TIL 0" level is defined as the level at which all No-Go events were observed while at least one Go event was observed at the next highest level. This next highest level can be defined as the "TIL +" level and used for comparison purposes as well. There is no obvious statistical distribution parameter associated with these TIL levels although for a 0/10 result, the TIL 0 level will be an estimate of an upper bound on the 10% reaction probability level.

Analysis of Variance

Analysis of statistical measurements from different laboratories can be formally evaluated through Analysis of Variance (ANOVA) [Devore 2012]. ANOVA is a standard method for assessing agreement among different measurements of mean values by comparing the standard deviation within a set of measurements to the standard deviation of the set averages. For data sets that are statistically equivalent, the standard deviations computed in these two different ways will be similar and their ratio will follow a statistical distribution with known characteristics. If one data set is statistically different from the others, the standard deviations computed in these two different ways will differ and their ratio will vary from the expected distribution by an amount that is characterized here by what is called a p-value. The p-value ultimately represents the probability that claiming there is a difference between the sets of results will be in error. For example, if p=0.05, then we could claim that there is a significant difference in the data sets with a 5% chance that the assessment is incorrect. This is called a Type I error in statistical texts. As the p-value gets larger, there is a greater chance of Type I error and so it is accurate to say that the data sets do not differ. For example, if p=0.95, we could claim that the results were different but would have 95% chance of being wrong – the natural interpretation would be to say that the data sets were the same. For very small p-values, the chance of a Type I error is very small and it is accurate to say that the data sets do differ In this report, the ANOVA treatment of the impact data was carried out using MiniTab 16, a commercially available software package [Minitab].

Tukey and Fisher Comparisons

The data set or sets responsible for disagreement in ANOVA can be determined using Tukey or Fisher comparison methods [Devore 2012, Fisher 1935]. In either of these methods, pairwise comparisons of the individual data sets are made and assessed against statistical distributions that are expected to describe their behavior. For the Tukey test, the distribution is called the "studentized range" distribution and for the Fisher test, the F-distribution is used. Disagreement among pairs in the set of results is used to assign groups of results that can be considered to be in agreement. These groups can be used to describe average results and identify outliers.

3 RESULTS

3.1 Impact testing results for RDX Type II Class 5

All participants evaluated the RDX impact sensitivity using the Bruceton Up-Down method[18,19] to estimate the mean and standard deviation of the response function. These results are presented in Table 2. In addition to the Bruceton method, LANL also evaluated the RDX using a D-optimal maximum likelihood method[20] implemented in commercial software [Neyer 1994]. Those results are presented in Table 3. Notable differences between laboratories from Table 1 for impact testing are variation in sandpaper type, amount of sample, striker mass, and the methods for detection of a reaction.

Table 2. Impact testing results for RDX Type II Class 5

Lab (grit)[1]	Set	Striker, kg	Test Date	T, °C	RH, %[2]	DH_{50}, cm[3]	s, cm[4]	s, log unit[4]
LLNL (120^{p})[5]	1	2.5	11/19/09	24	18	28.8	2.8	0.042
LLNL (120)	1	2.5	02/08/10	23	22	24.2	0.8	0.015
LLNL (120)	1	2.5	02/16/10	23	23	24.0	1.9	0.035
LLNL (120^{p})[5]	2	2.5	9/8/10	23.9	32	34.0	4.63	0.059
LLNL (180)	2	2.5	9/9/10	23.9	30	22.9	2.22	0.042
LLNL (180)	2	2.5	9/13/10	22.8	23	20.7	4.56	0.095
LLNL (120)	3	2.5	4/24/11	23.9	18	24.8	3.09	0.054
LLNL (180)	3	2.5	5/4/11	23.9	18	22.8	4.65	0.088
LLNL (180)	3	2.5	5/4/11	23.9	18	21.4	2.02	0.041
LLNL (180)	4	2.5	5/25/11	23.9	20	22.1	2.29	0.045
LLNL (180)	4	2.5	5/25/11	23.3	21	23.3	1.88	0.035
LLNL (180)	4	2.5	5/27/11	23.3	22	24.8	3.90	0.068
LLNL (180)	4	1.0	11/28/11	23.9	21	26.0	9.10	0.149
LLNL (180)	4	1.0	11/28/11	23.9	21	26.5	2.14	0.035
LANL (150)	1	0.8	11/23/09	21	17	26.5	1.2	0.019
LANL (150)	1	0.8	11/23/09	22	16	25.5	1.1	0.019
LANL (150)	1	0.8	11/23/09	22	16	24.2	1.5	0.027
LANL (180)	2	0.8	12/06/10	22.3	< 16	22.0	1.52	0.030
LANL (180)	2	0.8	12/09/10	21.7	< 16	20.3	2.30	0.049
LANL (180)	2	0.8	12/10/10	21.7	< 16	20.0	2.26	0.049
LANL (180)	3	0.8	4/6/11	22.9	< 10	23.3	1.45	0.027
LANL (180)	3	0.8	4/12/11	21.8	< 10	23.1	2.56	0.048
LANL (180)	3	0.8	4/12/11	21.7	< 10	23.1	1.60	0.030
LANL (180)	4	0.8	5/10/11	23.1	< 10	19.6	2.81	0.062
LANL (180)	4	0.8	5/11/11	20.4	< 10	17.7	4.82	0.117
LANL (180)	4	0.8	5/12/11	21.2	< 10	19.2	6.44	0.143
IHD (180)	1	2.5	11/24/09	26	38	22	8.3	0.16
IHD (180)	1	2.5	01/11/10	26	38	19	8.1	0.18
IHD (180)	1	2.5	01/20/10	26	40	18	10.9	0.25
IHD (180)	1	2.5	01/20/10	26	40	18	4.6	0.11
IHD (180)	2	2.5	3/8/11	28	40	17	4.76	0.12
IHD (180)	2	2.5	3/9/11	24	43	21	1.94	0.04
IHD (180)	2	2.5	3/8/11	29	43	15	3.13	0.09
IHD (180)	3	2.5	1/4/12	24	42	24	3.33	0.06
IHD (180)	3	2.5	2/15/12	26	43	21	2.91	0.07
IHD (180)	3	2.5	4/11/12	27	41	22	2.54	0.05
IHD (180)	4	2.5	9/24/12	19	45	18	2.91	0.07
IHD (180)	4	2.5	10/26/12	24	44	18	3.33	0.08
IHD (180)	4	2.5	1/9/13	22	41	19	2.63	0.06
AFRL (180)	1	2.5	4/29/10	22	43	15.1	3.70	0.10
AFRL (180)	1	2.5	4/29/10	23	45	13.1	3.70	0.17
AFRL (180)	1	2.5	5/4/10	27	57	17.6	4.70	0.09
SNL (180)	2	2.5	5/8/12	21.7	29.9	22.2	0.8	0.016
SNL (180)	2	2.5	5/10/12	20.0	28.2	22.6	1.5	0.023
SNL (180)	2	2.5	5/15/12	22.5	33.6	25.1	1.2	0.021

1. Value in parenthesis is grit size of sandpaper (180 is 180 garnet dry 150 is garnet dry and 120 is 120 Si/Carbide wet); 2 relative humidity; 3. DH_{50}, in cm, from a modified Bruceton method, height for 50% probability of reaction (DH_{50}); 4. Standard deviation; 5. p = pressed into pellet

Table 3. Impact testing results for RDX Type II Class 5 (Neyer or D-Optimal Method)

Lab[1,5]	Set	Test Date	T, °C	RH, %[2]	DH_{50}, cm[3]	s, cm[4]	s, log unit[4]
LANL (150)	1	12/24/09	20	17	24.0	3.3	0.06
LANL (150)	1	12/24/09	20	17	24.4	3.4	0.06
LANL (150)	1	12/24/09	20	17	23.7	2.7	0.05
LANL (150)	1	4/8/10	24.2	<10	26.7	5.6	0.09
LANL (180)	2	4/8/10	24.2	<10	20.4	3.3	0.07
LANL (180)	2	12/06/10	21.8	< 10	23.2	2.5	0.047
LANL (180)	2	12/09/10	21.8	< 10	21.2	2.3	0.047
LANL (180)	2	12/10/10	21.7	< 10	20.1	1.3	0.028
LANL (180)	3	4/6/11	22.5	< 10	20.6	3.7	0.079
LANL (180)	3	4/12/11	22.1	< 10	23.3	1.0	0.019
LANL (180)	3	4/12/11	21.8	< 10	21.3	1.5	0.031
LANL (180)	4	5/10/11	22.9	< 10	18.7	5.6	0.134
LANL (180)	4	5/11/11	20.4	< 10	21.9	2.8	0.056
LANL (180)	4	5/11/11	20.2	< 10	20.1	5.8	0.129
AFRL (180)	2	3/27/12	22.8	45	10.2	3.2	0.14

1. Value in parenthesis is grit size of sandpaper (180 is 180 garnet dry 150 is 150 is garnet dry); 2 relative humidity; 3. DH_{50}, in cm, from a Neyer D-Optimal method, height for 50% probability of reaction; 4. Standard deviation; 5. 0.8 kg Striker weight.

3.2 Analysis of Impact Testing Results

For the statistical analysis of results, the goals are to:

- Determine whether all labs or a subset of labs appear to be making "equivalent" measurements,
- Determine the expected range of values that might be observed by any laboratory,
- Evaluate possible dependence of the results on method or environment variables,
- Identify causes for any lab-to-lab variation.

Equivalency Characterization

Statistically, the question in this type of comparison is whether the impact test results from different participants are in agreement. In other words, given a few test results from each laboratory, is there a statistically significant difference among the participants or does the pooled set of results appear to arise from natural sampling error that would occur with repeated identical measurements of a single system? Since the RDX material supplied to the participants was from one batch, this analysis probes variation in the test methods and testing environment.

Due to significant differences in the test method details between laboratories, only the DH_{50} values are compared – the standard deviations produced by Bruceton analysis are not analyzed. The standard deviations are dependent on the number of tests, the step size, and the type of spacing (linear vs logarithmic), many of which varied among the participants.

The impact results in Table 2 can be visually evaluated using box plots with the data divided into sets differentiated by the combination of testing laboratory, sandpaper type (as indicated by grit size) and evaluation method (Bruceton or Neyer). Figure 1 shows these plots. Box plots are constructed so that the shaded region represents the middle 50% of the data and the horizontal line is the median of the

DH_{50} values. The vertical lines extend to the maximum and minimum DH_{50} values. The mean of the data set is at the midpoint of the shaded area.

In Figure 1, the two data points for LLNL 1 kg strikers were removed from the LLNL 180-grit data set. Striker mass was constant within each laboratory and is not indicated to differentiate sets of results. Any set that is not labeled as Neyer was evaluated with the Bruceton method. Visual inspection suggests that the results range from symmetric to skewed and that a subset or subsets of the different groupings are likely in agreement with each other, based on overlap of the shaded regions and to some extent the max/min bars. The AFRL 180 data appears to be significantly separated from the rest.

The p-value resulting from the ANOVA treatment of the RDX DH_{50} impact data was 0.000. Based on this at least one of the data sets represented in Figure 1 was statistically different than the others and that there is less than 0.1% chance that this assessment is in error.

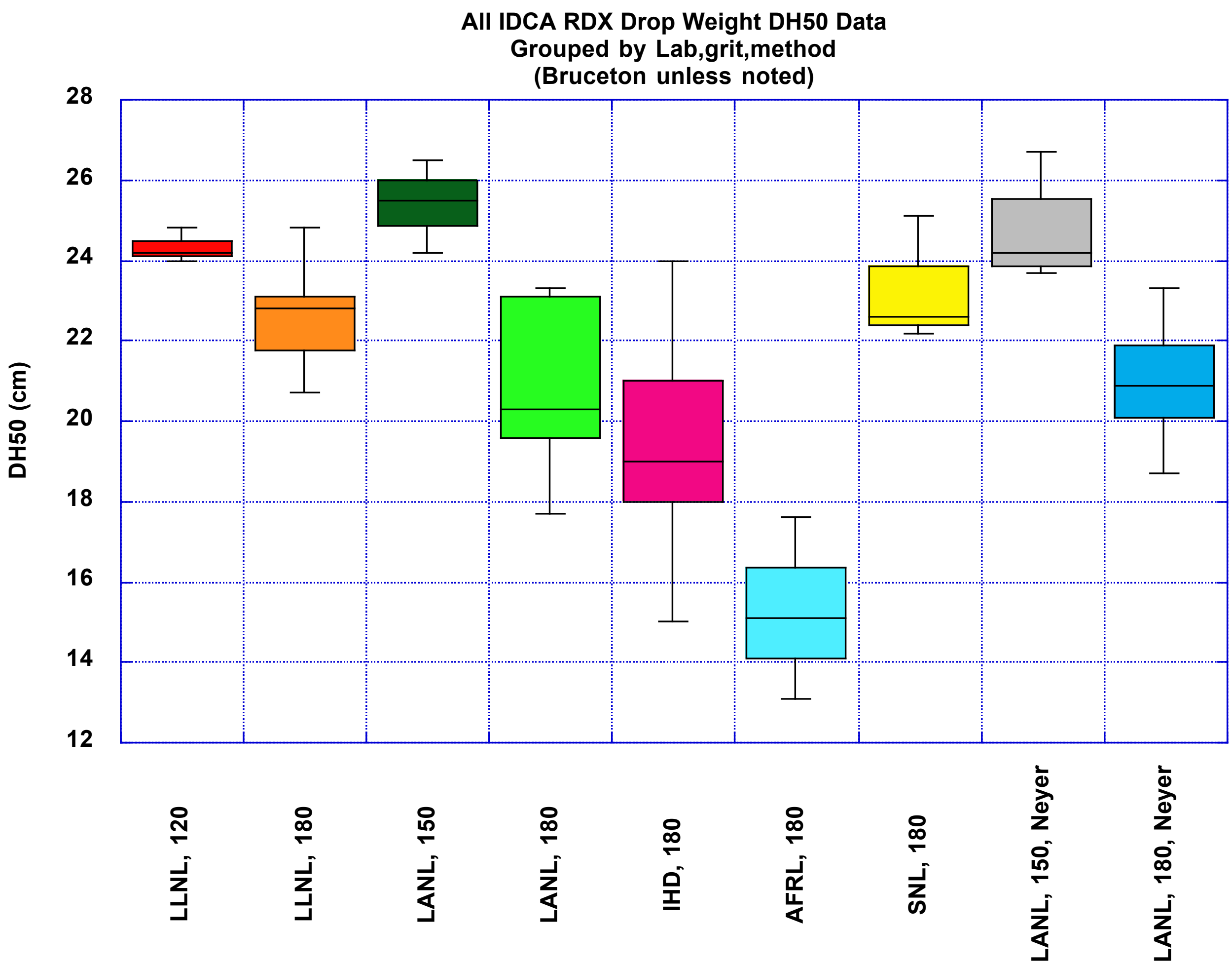

Figure 1. Box plot of the DH_{50} values grouped by participant sandpaper grit size and data reduction method. Details are explained in the text.

Tukey and Fisher comparison analysis results at a 95% confidence level are shown in Table 4. Overall, many of the individual data sets are in agreement but, as expected, the AFRL 180 data set is different from the rest of the sets no matter how they are grouped. [I think we should we show these evaluations in the appendix.]

Table 4. Groupings resulting from Tukey and Fisher Comparison Tests of RDX Impact Data.

Tukey	**Data sets in grouping**[1]
Subgroup 1	LANL (150/B), LANL (150/N), LLNL (120/B), SNL (180/B), LLNL (180/B)
Subgroup 2	LANL (150/N), LLNL (120/B), SNL (180/B), LLNL (180/B), LANL (180/N)
Subgroup 3	LLNL (120/B), SNL (180/B), LLNL (180/B), LANL (180/N), LANL (180/B)
Subgroup 4	SNL (180/B), LANL (180/N), LANL (180/B), IHD (180/B)
Subgroup 5	AFRL (180/B)
Fisher	
Subgroup 1	LANL (150/B), LANL (150/N), LLNL (120/N), SNL (180/N), LLNL (180/B)
Subgroup 2	LANL (180/B), SNL (180/B), LANL (180/N)
Subgroup 3	IHD (180/B), LANL (180/B)
Subgroup 5	AFRL (180/B)

1. Values in parentheses indicate type of sandpaper/analysis method (120 is 120-grit Si/C wet/dry sandpaper, 150 is 150-grit garnet sandpaper, 180 is 180-grit garnet sandpaper, B = Bruceton method, N= Neyer D-Optimal method);

Expected Range of Observations

One useful outcome of a Proficiency Test is an assessment of the expected range of values that might be obtained by other laboratories carrying out nominally the same measurements in the future. This can be used as verification that future laboratories are capable of making this type of measurement adequately, hence the name "Proficiency". In the IDCA context, the range of observations is probably more appropriately interpreted as the variability in the observations that may be expected. This can be used as a lower bound for expected variability in materials that are not as well behaved or as well characterized as RDX.

Figure 2 shows the individual DH_{50} impact results from Bruceton or Neyer analysis. Included are the LLNL data taken using the 1-kg striker and 180-grit sandpaper and the AFRL data taken with 180-grit sandpaper, but not the LLNL data from pressed pellets. Each point is the mean value for a particular data set and the error bars are the standard deviation calculated in the Bruceton or Neyer methods. The standard deviations are not easily compared for reasons noted above. They are, however, often larger than the scatter observed in repeated measurements that produce the DH_{50} values and so they are a good representation of a worst-case estimate.

The range is illustrated in Figure 2 using green horizontal lines that pass through the maximum and minimum observed means. This range is 13.1 (AFRL 180-grit sandpaper) to 26.7 (LANL 150-grit sandpaper, Neyer analysis) cm. Because there were not very many tests leading to some of the data points in Figure 2, it is appropriate to take into account the standard deviations, and report a broadened range of the expected values. As an estimate, using the average standard deviation of all of the measurements, which is around 3 cm, broadens the range to 10.1 to 29.5 cm. This broadened range is highlighted using the blue horizontal lines. Based on these results, for a future laboratory using any instrument from very old to brand new and methods of detection ranging from operator to microphone, should expect results for this particular RDX to fall between 10.1 and 29.5 cm [let us compare this to the RS RDX study by Doherty et al. in the discussion].

The mean of the values presented in Figure 2 is 21.0 cm and half of the unbroadened range is 6.7 cm or approximately a relative 30% of the mean, setting an expectation to observe percentage variability in testing other materials that have higher or lower mean DH_{50} values to have this value.

Based on the ANOVA results presented above, it is appropriate to evaluate the same ranges and percent variability after removing the AFRL 180 data since it is statistically different from all of the other sets based on both Tukey and Fisher comparisons. With the AFRL data removed, the range of means is 15 to 26.5 cm with an average of 21.5 cm and a percent variability of 27%. The broadened range would be 12 to 29.5 cm.

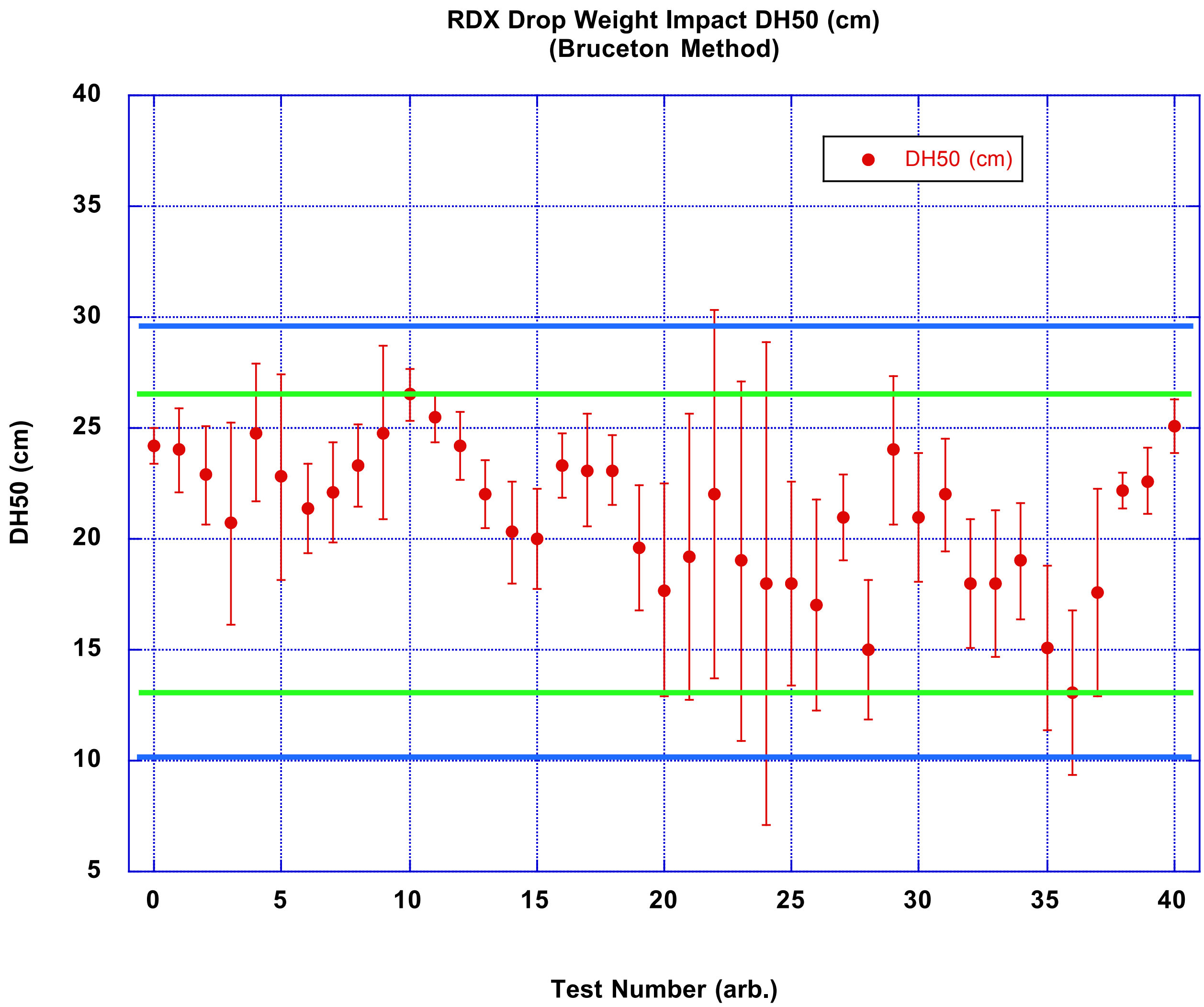

Figure 2. Comparison of individual DH_{50} evaluations for all labs with estimated errors determined from the Bruceton or Neyer standard deviation values. Test number is arbitrary and for display purposes only (implies no order of testing).

Dependence on Method or Environment Variables

When replicate measurements are available, it is possible to compare the results against other parameters and look for relations that suggest an influence due to a variable in the test method or in the local environment. For the RDX impact data set this is complicated by the variability among laboratories. Fortunately, moving between any adjacent Tukey subgroups 1 through 4 includes all laboratories except AFRL most of the time and so it is appropriate to use all of the data except the AFRL 180 set to examine dependence on method or environment variables. The variables that were tracked by each lab include striker mass, sandpaper type (identified by grit size but includes other sandpaper properties), temperature, and relative humidity. Figure 3 shows individual DH_{50} values (determined by Bruceton or Neyer methods) as a function of specific variables. None of the variables showed an influence on the DH_{50} with the possible exception of sandpaper type, which has been shown to matter in other studies of other materials. The data set is not large enough to be conclusive about the dependence at this point.

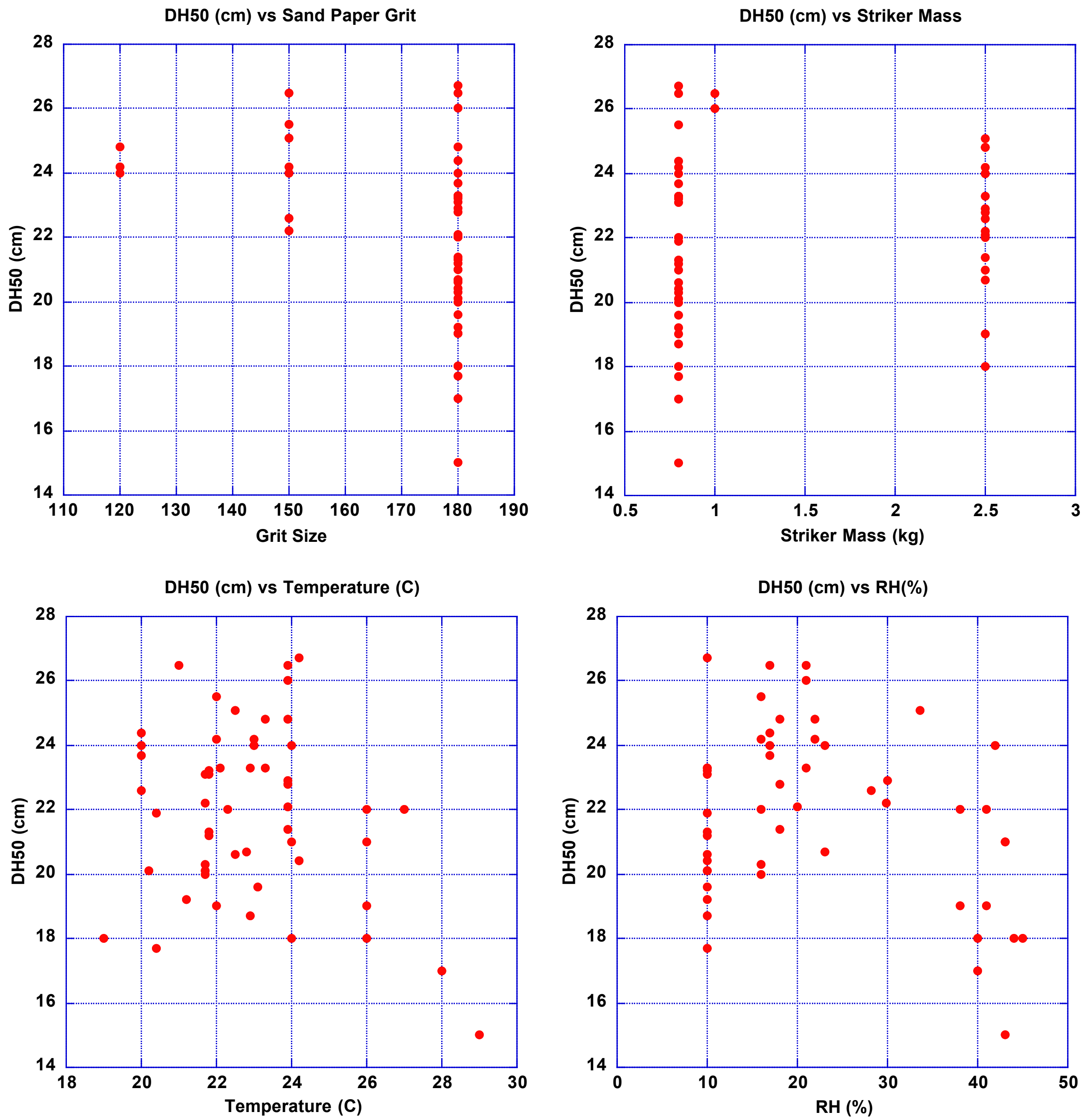

Figure 3. Comparison of DH_{50} with various method and environment variables.

3.3 Analysis of BAM Friction testing results.

Table 5 shows all of the RDX BAM Friction testing performed by LANL, LLNL, IHD, and SNL (AFRL does not have BAM friction). The notable differences in test methods for BAM friction testing are the methods for positive detection and the environment surrounding the instrument. LANL, LLNL, IHD and SNL performed data analysis using the threshold initiation level method (TIL)[21]. LANL, LLNL and IHD also used a modified Bruceton method[18,19] and IHD used the Neyer method[20] on Data set 2 because their data did not meet Bruceton criteria (analysis performed by LANL). SNL did not carry out a Bruceton evaluation with their instrument.

Table 5. Average BAM Friction Testing results for RDX Type II Class 5

Lab	Set	Test Date	T, °C	RH, %[1]	TIL, kg[2]	TIL+, kg[3]	F_{50}, kg[4]	s, kg[5]	s, log unit[5]
LLNL	1	11/23/09	22.8	18	0/10 @ 19.2	1/10 @ 21.6	25.4	3.2	0.054
LLNL	1	02/09/10	22.8	23	0/10 @ 21.6	1/10 @ 24.0	24.6	2.8	0.050
LLNL	1	02/16/10	22.8	30	0/10 @ 16.8	1/10 @ 19.2	26.1	4.2	0.070
LLNL	2	9/08/10	23.9	26	0/10 @ 16.0	1/10 @ 16.8	23.1	1.86	0.035
LLNL	2	9/09/10	23.9	31	0/10 @ 16.8	1/10 @ 18.0	25.4	3.17	0.054
LLNL	2	9/09/10	23.9	31	0/10 @ 16.8	1/10 @ 19.2	26.0	3.00	0.050
LLNL	3	4/28/11	23.9	20	NA[6]	NA[6]	19.8	3.58	0.078
LLNL	3	5/3/11	23.9	15	NA[6]	NA[6]	23.2	5.27	0.098
LLNL	3	5/4/11	23.9	13	NA[6]	NA[6]	20.3	1.97	0.042
LLNL	4	5/25/11	23.9	23	0/10 @ 16.0	1/10 @16.8	20.6	2.76	0.058
LLNL	4	5/26/11	23.9	20	0/10 @ 16.8	1/10 @ 17.4	25.0	1.56	0.027
LLNL	4	5/27/11	21.7	24	0/10 @ 16.0	1/10 @ 16.8	21.1	1.31	0.042
LANL	1	11/23/09	22.0	16.0	NA[6]	NA[6]	20.8	3.4	0.07
LANL	1	11/24/09	20.0	17.0	NA[6]	NA[6]	23.0	2.1	0.04
LANL	1	11/24/09	21.0	17.0	NA[6]	NA[6]	18.7	5.2	0.12
LANL	1	01/11/10	19.1	< 10	0/10 @ 19.2	1/4 @ 21.6	NA[7]	NA[7]	NA[7]
LANL	2	12/06/10	22.1	< 10	0/10 @ 9.6	1/8 @ 12.0	NA[7]	NA[7]	NA[7]
LANL	2	12/08/10	21.1	< 10	0/10 @ 12.0	1/3 @ 14.4	NA[7]	NA[7]	NA[7]
LANL	2	12/08/10	22.1	< 10	0/10 @ 9.6	1/5 @ 12.0	NA[7]	NA[7]	NA[7]
LANL	2	12/0610	22.2	< 10	NA[6]	NA[6]	15.1	3.6	0.106
LANL	2	12/08/10	20.8	< 10	NA[6]	NA[6]	16.7	2.3	0.060
LANL	2	12/08/10	20.8	< 10	NA[6]	NA[6]	17.1	1.8	0.046
LANL	3	4/11/11	22.0	< 10	NA[6]	NA[6]	14.9	1.73	0.051
LANL	3	4/11/11	21.8	< 10	NA[6]	NA[6]	16.7	2.96	0.078
LANL	3	4/11/11	21.8	< 10	NA[6]	NA[6]	15.1	1.73	0.086
LANL	3	4/11/11	21.8	< 10	0/10 @ 12.2	1/5 @ 14.7	NA[7]	NA[7]	NA[7]
LANL	3	4/11/11	21.8	< 10	0/10 @ 12.2	1/1 @ 14.7	NA[7]	NA[7]	NA[7]
LANL	3	4/11/11	21.9	< 10	0/10 @ 9.8	1/9 @ 12.2	NA[7]	NA[7]	NA[7]
LANL	4	5/10/11	21.8	< 10	NA[6]	NA[6]	19.4	3.7	0.084
LANL	4	5/10/11	23.4	< 10	NA[6]	NA[6]	20.4	1.3	0.028
LANL	4	5/10/11	22.6	< 10	NA[6]	NA[6]	21.4	1.5	0.030
LANL	4	5/11/11	23.4	< 10	0/10 @ 12.2	1/10 @ 14.7	NA[7]	NA[7]	NA[7]
LANL	4	5/11/11	23.4	< 10	0/10 @ 14.7	1/4 @ 17.1	NA[7]	NA[7]	NA[7]
LANL	4	5/11/11	23.5	< 10	0/10 @ 14.7	1/6 @ 17.1	NA[7]	NA[7]	NA[7]
IHD	1	11/25/09	26	37	0/10 @ 14.7	1/3 @ 16.3	NA[7]	NA[7]	NA[7]
IHD	1	01/25/10	27	49	0/10 @ 14.7	1/6 @ 16.3	NA[7]	NA[7]	NA[7]
IHD	1	01/25/10	27	46	0/10 @ 16.3	1/2 @ 18.4	NA[7]	NA[7]	NA[7]
IHD	1	01/25/10	27	48	0/10 @ 14.7	1/4 @ 16.3	NA[7]	NA[7]	NA[7]
IHD	2	3/31/11	23	40	0/10 @ 11.0	1/4 @ 12.2	NA[7]	NA[7]	NA[7]
IHD	2	2/23/11	26	40	0/10 @ 12.2	1/5 @ 14.7	NA[7]	NA[7]	NA[7]
IHD	2	4/22/11	22	40	0/10 @ 12.2	1/5 @ 14.7	NA[7]	NA[7]	NA[7]
IHD[8]	2	4/11/11	NA[9]	NA[9]	NA[6]	NA[6]	31.6	7.0	0.098
IHD[8]	2	4/11/11	NA[9]	NA[9]	NA[6]	NA[6]	24.9	12.0	0.228
IHD[8]	2	4/11/11	NA[9]	NA[9]	NA[6]	NA[6]	26.9	23.7	0.600
IHD	3	1/3/12	26	42	0/10 @ 12.2	1/3 @14.7	NA[7]	NA[7]	NA[7]

Lab	Set	Test Date	T, °C	RH, %[1]	TIL, kg[2]	TIL+, kg[3]	F_{50}, kg[4]	s, kg[5]	s, log unit[5]
IHD	3	2/16/12	27	43	0/10 @ 11.0	1/1 @ 12.2	NA[7]	NA[7]	NA[7]
IHD	3	4/11/12	28	40	0/10 @ 11.0	1/3 @ 12.2	NA[7]	NA[7]	NA[7]
IHD	3	6/8/12	20	41	NA[6]	NA[6]	19	2.3	0.053
IHD	3	6/8/12	20	42	NA[6]	NA[6]	20	2.9	0.063
IHD	3	6/8/12	20	42	NA[6]	NA[6]	19	3.0	0.069
IHD	4	11/1/12	24	42	0/10 @ 12.2	1/2 @ 14.7	NA[7]	NA[7]	NA[7]
IHD	4	12/18/12	24	40	0/10 @ 11.1	1/6 @ 12.2	NA[7]	NA[7]	NA[7]
IHD	4	2/8/13	25	40	0/10 @ 12.2	1/6 @ 14.7	NA[7]	NA[7]	NA[7]
IHD	4	3/8/13	25	40	NA[6]	NA[6]	16.6	2.4	0.063
IHD	4	3/8/13	25	40	NA[6]	NA[6]	17.6	3.2	0.080
IHD	4	3/8/13	25	40	NA[6]	NA[6]	18.2	3.0	0.072
SNL	2	5/8/12	22.2	31.0	0/20 @ 16.8	1/14 @ 18.0	NA	NA	NA
SNL	2	5/9/12	22.2	28.1	0/20 @ 16.0	1/2 @ 16.8	NA	NA	NA
SNL	2	5/10/12	20.4	31.3	0/20 @ 16.0	1/7 @ 16.8	NA	NA	NA

1. Relative humidity; 2. Threshold Initiation Level (TIL) is the load (kg) at which zero reaction out of twenty or fewer trials with at least one reaction out of twenty or fewer trials at the next higher load level; 3. Next level where positive initiation is detected; 4. F_{50}, in kg, is by a modified Bruceton method, weight for 50% probability of reaction; 5. Standard deviation; 6. Not applicable, separate measurements performed for modified Bruceton analysis; 7. Not applicable, separate measurement performed for TIL; 8. Modified Neyer analysis; 9. Not measured. LLNL uses log-spacing and LANL uses liner spacing for the Bruceton up and down method experimentation and data analysis.

Equivalency Characterization

Figure 4 shows the BAM friction data for RDX as averaged for each participant as presented as box plots. Box plots are constructed so that the shaded region represents 50% of the data and the horizontal line is the median of the F_{50} values. The vertical lines show the maximum and minimum F_{50} values. The mean of the data set is at the midpoint of the shaded area. The smaller number of sets visually might imply better agreement between the participants as compared to the impact (DH_{50}) results, however the bulk of the LANL and LLNL data are still significantly offset from each other.

TIL data is more difficult to compare because not all the participants used equal numbers of test events (trails and No-Go events) to evaluate the threshold levels. As a result, directly comparing the TIL values and the level above TIL (TIL+) better illustrate the trends in the data. Figure 5 compares the TIL and TIL+ values determined by BAM friction for LANL, LLNL, IHD and SNL data. The first four point types are TIL (TIL 0) data and the last four are TIL+ (TIL 1) data.

Visually, for the TIL values, although there is overlap of the data sets, the LLNL data points appear higher than the other, implying that LLNL found the RDX less sensitive than the other participants. The same holds true for the TIL+ values, LLNL appears to find the RDX less sensitive than the other participants.

Table 6 shows the ANOVA results for F_{50} and both TIL data sets along with the various subgroups determined by Tukey and Fisher analysis methods that are in apparent agreement. For the F_{50} determination, LANL appears to differ from both LLNL and IHD with a p-value of 0.000 although IHD forms a subgroup with either. For TIL values, the participants differ with a p-value of 0.000 and the subgroups each include at least two participants. For TIL+, the participants differ with a p-value of 0.001 and subgroups are similar to those for TIL 0. In both TIL 0 and TIL+, SNL is on the borderline between two groups and may be included in both. These p-values results indicate that the data sets are all different with a 0.1% chance that this assessment is a Type I error.

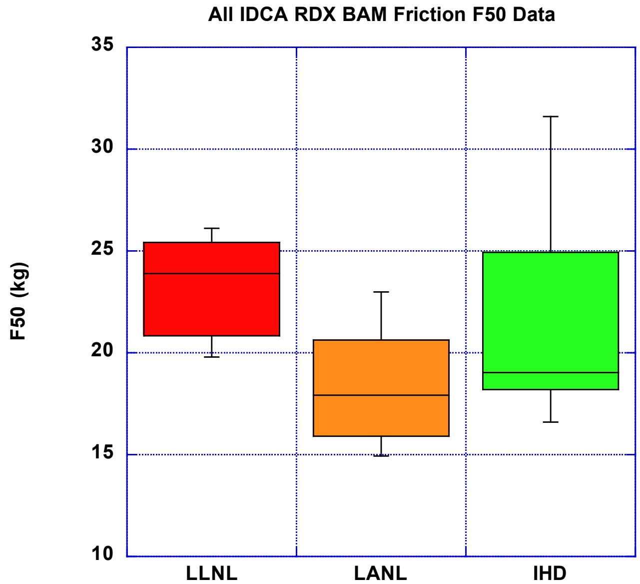

Figure 4. Comparison of RDX F_{50} results represented as box plots. All data acquired using BAM friction apparatus.

Table 6. ANOVA Results for BAM Friction Testing

	ANOVA p-value	Tukey subgroups	Fisher subgroups
F_{50}	0.000	1. IHD, LLNL 2. LANL, IHD	1. IHD, LLNL 2. LANL, IHD
TIL	0.000	1. LLNL, SNL 2. SNL, IHD, LANL	1. LLNL, SNL 2. IHD, LANL
TIL+	0.001	1. LLNL, SNL 2. SNL, IHD, LANL	1. LLNL, SNL 2. SNL, IHD, LANL

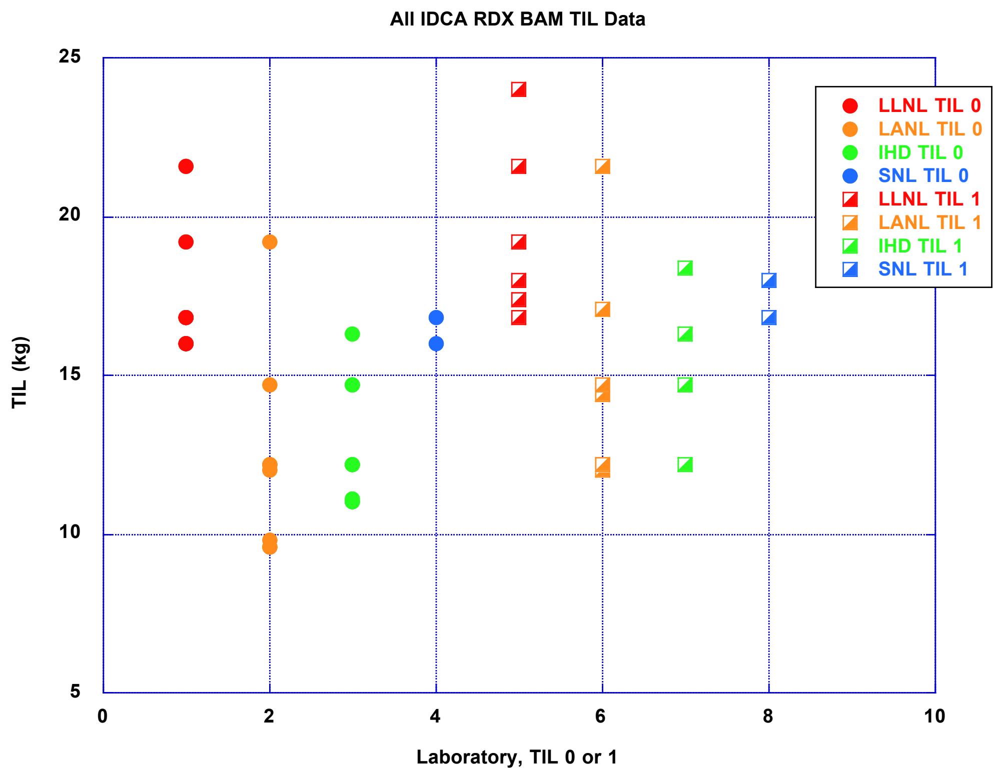

Figure 5. Comparison of TIL and TIL+ levels from BAM friction data for RDX. The X axis numbering is for display purposes only—does not imply testing order.

Expected Range of Observations

For the BAM friction F_{50} data, because there were only three participants contributing, all of the data was evaluated together to assign a range of expected values from 14.9 to 31.6 kg. The mean value of all of the F_{50} measurements is 21.0 kg so that half of the range represents a variability of 40 %. The standard deviation of the F_{50} measurements is only 4 kg, which is 19% of the mean. The 40% variability is roughly equal to two standard deviations.

For the TIL 0 data, the range runs from 9.6 to 21.6 kg. The mean TIL 0 value is 14.2 kg with a standard deviation of 3 kg. Using half of the range, the expected variability among the participants is 6 kg, or about 42% of the mean, and again about 2 standard deviations.

For the TIL+ data, the range is from 12 to 24 kg. The mean TIL+ value is 16.0 kg with a standard deviation of 2.9 kg. Using half of the range, the expected variability among the participants is also 6 kg, or about 38% of the mean, and again close to 2 standard deviations.

Dependence on Method or Environment Variables

There are sufficient data to be able to compare the results against other parameters and look for relations that suggest an influence due to a variable in the test method or in the local environment. The variables considered were temperature and humidity. Because all of the instruments and friction pin/plate surfaces were the same (not accounting for aging due to use), no other variables were considered. Figure 6 shows the comparison of the F_{50}, TIL, or TIL+ as a function of temperature and relative humidity No dependence on either variable is evident in the figure. This implies that there is no dependence of these measurements on method or environment variables except for those variables that are inherent to each laboratory and therefore captured in the ANOVA analyses. For BAM, these inherent variables are the operator, background environment (including insulation), and method of reaction determination, which are all interrelated.

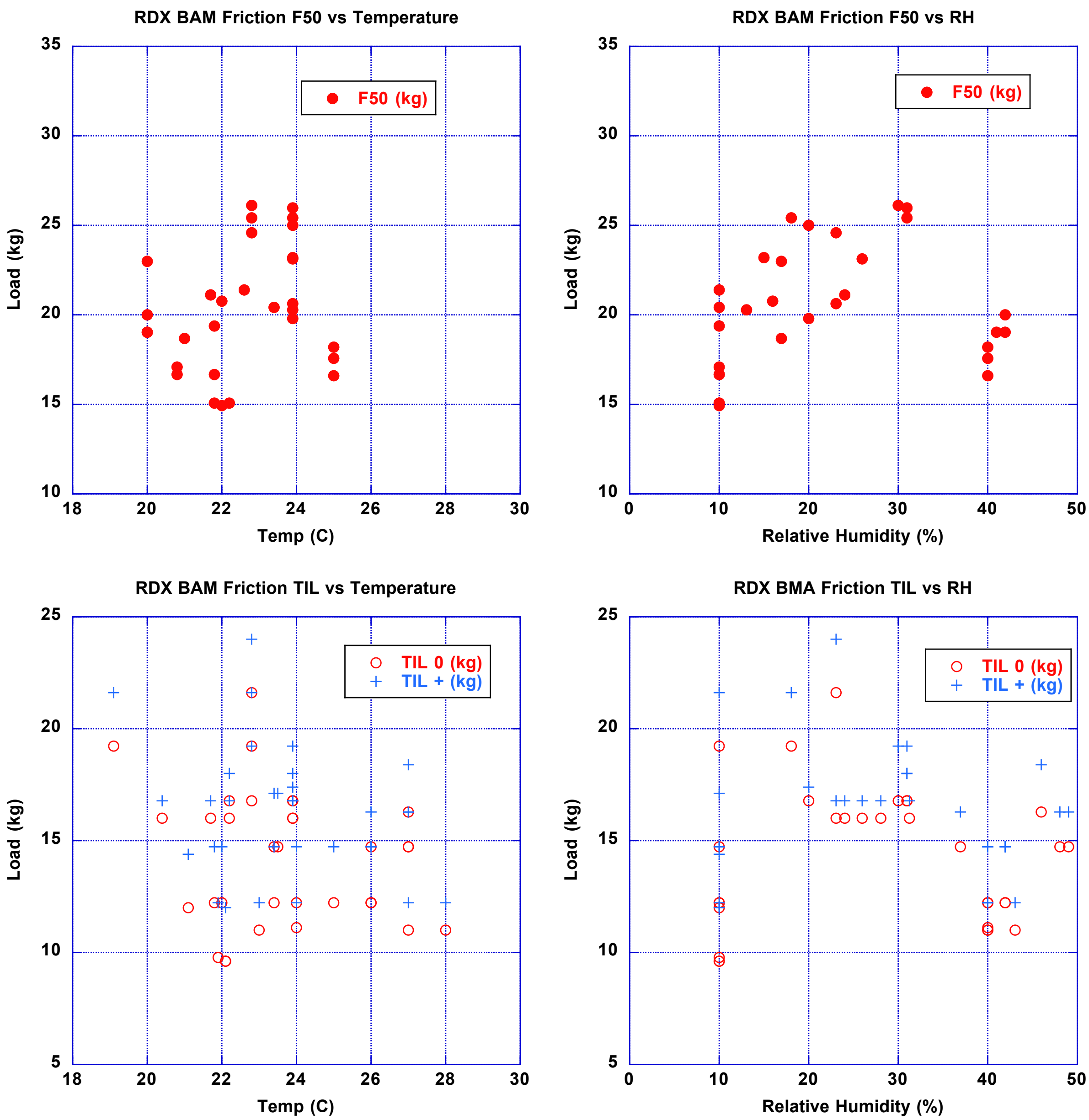

Figure 6. Comparison of F_{50} with temperature and humidity.

3.4 Electrostatic discharge testing of RDX Type II Class 5

Electrostatic Discharge (ESD) testing of the RDX Type II Class 5 was performed by LLNL, LANL, and IHD. Table 7 shows all results. Differences in the testing procedures are shown in Table 2, and the notable differences are the use of tape and what covers the sample. All participants performed data analysis using the threshold initiation level method (TIL)[21]. LLNL used a custom built ESD system with a 510-Ω resistor in line to simulate a human body for Set 1 and Set 2. LLNL also used a new ABL system for Set 3 and Set 4. Other participants used older ABL systems.

Table 7. Electrostatic discharge testing results for RDX Type II Class 5

Lab	Set	Ω	Test Date	T, °C	RH, %[1]	TIL, Joule[2]	TIL+, Joule[3]
LLNL	1	510[4]	11/18/09	22.8	18	0/10 @ 1.0	0/10 @ 1.0[4]
LLNL	1	510[4]	02/08/10	22.8	23	0/10 @ 1.0	0/10 @ 1.0[4]
LLNL	1	510[4]	02/16/10	22.8	30	0/10 @ 1.0	0/10 @ 1.0[4]
LLNL	2	510[4]	9/08/10	23.9	26	0/10 @ 1.0	0/10 @ 1.0
LLNL	2	510[4]	9/08/10	23.9	32	0/10 @ 1.0	0/10 @ 1.0
LLNL	2	510[4]	9/10/10	23.9	29	0/10 @ 1.0	0/10 @ 1.0
LLNL	3	0[5]	4/20/11	23.9	21	0/10 @ 0.038	1/2 @ 0.063
LLNL	3	0[5]	4/26/11	23.9	16	0/10 @ 0.038	1/3 @ 0.063
LLNL	3	0[5]	4/26/11	23.9	16	0/10 @ 0.038	1/3 @ 0.063
LLNL	4	0[5]	4/26/11	23.3	22	0/10 @ 0.038	1/3 @ 0.063
LLNL	4	0[5]	4/26/11	24.4	20	0/10 @ 0.038	1/2 @ 0.063
LLNL	4	0[5]	4/26/11	23.3	21	0/10 @ 0.038	1/6 @ 0.063
LANL	1	0[5]	11/24/09	20	17	0/20 @ 0.025	2/11 @ 0.0625
LANL	1	0[5]	11/24/09	19	17	0/20 @ 0.025	2/7 @ 0.0625
LANL	1	0[5]	11/24/09	19	17	0/20 @ 0.025	2/7 @ 0.0625
LANL	2	0[5]	12/06/10	22.2	< 10	0/20 @ 0.025	1/17 @ 0.0625
LANL	2	0[5]	12/08/10	21.0	< 10	0/20 @ 0.0625	1/1 @ 0.125
LANL	2	0[5]	12/08/10	20.9	< 10	0/20 @ 0.025	1/13 @ 0.0625
LANL	3	0[5]	4/11/11	22.3	< 10	0/20 @ 0.025	1/9 @ 0.0625
LANL	3	0[5]	4/11/11	21.9	< 10	0/20 @ 0.025	1/3 @ 0.0625
LANL	3	0[5]	4/11/11	22.0	< 10	0/20 @ 0.025	2/16 @ 0.0625
LANL	4	0[5]	5/5/11	23.4	< 10	0/20 @ 0.025	1/12 @ 0.0625
LANL	4	0[5]	5/5/11	23.6	< 10	0/20 @ 0.025	1/10 @ 0.0625
LANL	4	0[5]	5/5/11	22.9	< 10	0/20 @ 0.025	1/7 @ 0.0625
IHD	1	0[5]	11/24/09	26	36	0/20 @ 0.095	1/7 @ 0.165
IHD	1	0[5]	01/15/10	27	40	0/20 @ 0.095	1/7 @ 0.165
IHD	1	0[5]	01/15/10	27	40	0/20 @ 0.095	1/14 @ 0.165
IHD	1	0[5]	01/19/10	27	40	0/20 @ 0.095	1/12 @ 0.165
IHD	2	0[5]	3/10/11	24	42	0/20 @ 0.037	1/4 @ 0.095
IHD	2	0[5]	3/10/11	24	42	0/20 @ 0.037	1/3 @ 0.095
IHD	2	0[5]	3/16/11	24	42	0/20 @ 0.037	1/16 @ 0.095
IHD	3	0[5]	11/20/11	28	42	0/20 @ 0.037	1/7 @ 0.095
IHD	3	0[5]	1/4/12	23	40	0/20 @ 0.095	1/8 @ 0.165
IHD	3	0[5]	2/16/12	26	42	0/20 @ 0.095	1/8 @ 0.165
IHD	4	0[5]	10/31/12	23	42	0/20 @ 0.095	1/4 @ 0.165
IHD	4	0[5]	12/11/12	22	42	0/20 @ 0.095	1/5 @ 0.165
IHD	4	0[5]	2/8/13	25	40	0/20 @ 0.095	1/4 @ 0.165

1. Relative humidity; 2. Threshold Initiation Level (TIL) is the load (joules) at which zero reaction out of twenty or fewer trials with at least one reaction out of twenty or fewer trials at the next higher load level; 3. Next level where positive initiation is detected; 4. LLNL used a custom built ESD with a 510-Ω resistor in the discharge unit to mimic the human body; 5. ABL ESD with 0-Ω resistance.

The ESD results from the participating labs are much more coarse grained than the data sets for impact or friction because of the step levels used in the testing and it is not informative to create scatter or box plots for this data set as a result. ANOVA analysis is also not useful because of the discreteness and clustering of the TIL and TIL+ levels shown in Table 7. The main obvious difference in the full ESD data set is the LLNL subgroup from the custom instrument with integrated 510-Ω resistor. Figure 7 shows a comparison of average TIL values with standard deviations illustrates a secondary difference, which is the higher TIL or TLI+ values [describe how the averages were calculated) from IHD compared to the corresponding values from LANL and LLNL.

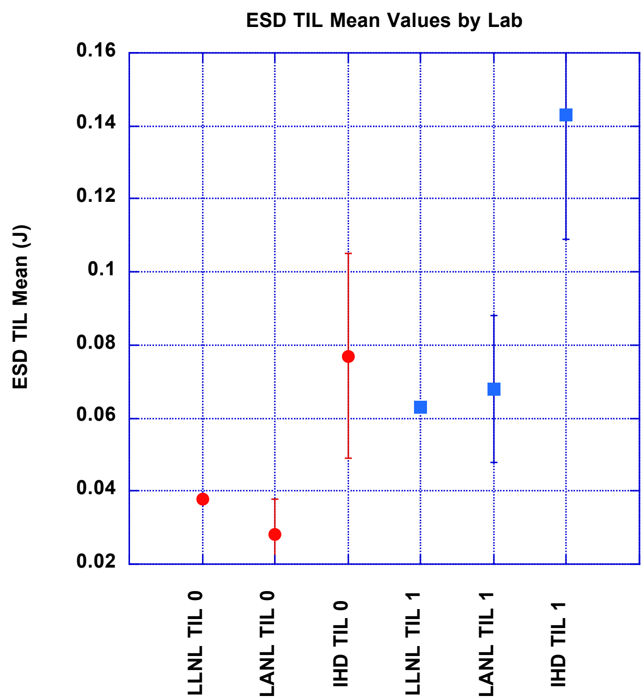

Figure 7. Average TIL and TIL+ values for ESD results. LLNL results using the custom instrument with a 510-Ω resistor are not included.

Figure 7 shows the degree of separation between IHD and other participants. In each case, the standard deviations do not even overlap. This shows that the measurements made by IHD are not equivalent to those made by the other participants.

For this data set, there is a possible link between an environment variable and the results. Examination of Table 7 and Figure 7 shows that the IHD results are higher than those of the other participants and were also obtained at roughly twice the relative humidity. Without more testing at different humidity levels, it is not possible to definitively say that this leads to the higher TIL values, but it is an understandable correlation since electrostatic effects are greatly influenced by humidity[42].

Assuming that the IHD results are due to humidity variation, then it is natural to group all participants together to assess the range of possible values that might be obtained by other laboratories attempting to carry out nominally the same type of ESD testing without a tightly controlled laboratory environment. In this case, the TIL range is from 0.025 to 0.095 J with an average of 0.051 J and a standard deviation of 0.030 J. Using either the range or the standard deviation implies greater than 50% variability, which is not very useful as a metric. For TIL+, the range is from 0.0625 to 0.165 J with a mean of 0.099 J and a standard deviation of 0.046 J. The range and standard deviation again imply a very large variability. For reporting and comparison purposes, the range itself is a more useful metric to assess any future measurements.

3.5 Thermal testing (DSC) of RDX Type II Class 5

Differential Scanning Calorimetry (DSC) was performed on the RDX Type II Class 5 by LLNL, LANL, and IHD. All participating laboratories used different versions of the DSC by TA Instruments. Results were obtained at a 10°C/min heating rate.

Table 8 shows the DSC results from RDX Sets 1 through 4 and Figure 7 shows a typical DSC scan using the pinhole hermetic pans and one type of sealed pan. The principal features of the DSC examinations are essentially the same from all participants—two overlapping low temperature endothermic features near 200°C and a major exothermic feature near 240°C. LLNL and IHD were able to examine the RDX using both open sample holders and sealed sample holders.

Table 8. Differential Scanning Calorimetry results for RDX Type II Class 5, 10°C/min heating rate

Lab	Set	Sample Holder	Test Date	Endothermic, onset/minimum, °C (ΔH, J/g)	Exothermic, onset[1]/maximum, °C (ΔH, J/g)
LLNL	1	Pinhole[2]	12/01/09	187.5/189.0, 199.2 (143)	203/241.1 (2281)[16]`
LLNL	1	Pinhole[2]	02/04/10	187.8/189.1, 199.3 (139)	203/240.7 (2299)[16]
LLNL	1	Pinhole[2]	02/04/10	187.8/189.1, 198.8 (136)	203/241.5 (2316)[16]
LLNL[15]	2	Pinhole[2]	8/27/10	187.3/188.3, ~199[4] (126)	213.13/240.1 (2432)
LLNL[15]	2	Pinhole[2]	8/27/10	187.5/188.6, ~200[4] (129)	215.61/240.6 (2419)
LLNL[15]	2	Pinhole[2]	8/27/10	187.4/188.4, ~199[4] (135)	217.91/238.7 (2399)
LLNL	3	Pinhole[2]	4/1/11	187.8/188.9, 199.2 (140)	217.1/242.4 (2353)
LLNL	3	Pinhole[2]	4/1/11	187.8/189.2, 199.3 (154)	218.4/242.0 (1890)
LLNL	3	Pinhole[2]	4/1/11	187.8/189.1, 199.5 (181)	218.4/243.5 (1927)
LLNL	4	Pinhole[2]	5/23/11	187.7/ 189.0, 199.2 (141)	218.6/242.4 (2195)
LLNL[15]	4	Pinhole[2]	5/23/11	187.8/189.0, 199.2 (145)	218.7/243.5 (2186)
LLNL	4	Pinhole[2]	5/23/11	187.7/188.9, 199.2 (130)	217.8/242.6 (2227)
LANL	1	Pinhole[5]	11/17/09	188.0/189.1, 199.6 (137)	218.8[3]/242.8 (2205)[16]
LANL	1	Pinhole[5]	11/24/09	188.1/189.6, 200.7 (135)	220.9[3]/242.8 (2260)[16]
LANL	1	Pinhole[5]	11/24/09	188.0/189.2, 199.9 (135)	224.8[3]/242.1 (2246)[16]
LANL	2	Pinhole[5]	12/02/10	188.2/189.7, 200.5 (129)	217.03/242.4 (2091)
LANL	2	Pinhole[5]	12/09/10	188.2/189.6, 200.8 (131)	219.23/243.0 (2138)
LANL	2	Pinhole[5]	12/15/10	188.0/189.2, 199.3 (140)	218.03/242.1 (2300)
LANL[14]	3	Pinhole[5]	4/12/11	188.6/189.8, 200.5 (137)	219.0/242.1 (2148)
LANL	3	Pinhole[5]	4/12/11	188.2/189.8, 200.1 (135)	218.8/243.0 (2097)

LANL	3	Pinhole[5]	4/12/11	188.2/189.9, 200.4 (130)	218.7/241.2 (2148)
LANL	4	Pinhole[5]	5/10/11	188.1/189.6, 200.2 (136)	215.8/242.2 (2204)
LANL	4	Pinhole[5]	5/10/11	188.3/189.5, 199.8 (115)	219.8/243.3 (2017)
LANL	4	Pinhole[5]	5/10/11	188.2/ 189.6, 200.6 (120)	220.3/244.2 (1947)
IHD	1	Pinhole[5]	11/25/09	188.0/189.2, 199.8 (120)	217.7/242.4 (1947)
IHD	1	Pinhole[5]	11/25/09	187.8/189.1, 199.4 (122)	218.0/242.3 (2034)
IHD	1	Pinhole[5]	11/25/09	188.0/189.4, 199.5 (127)	219.2/241.9 (2141)
IHD[6]	2	Pinhole[5]	9/29/09	187.7/189.2, 199.3 (107)	210.93/240.2 (1375)[6,17]
IHD[6]	2	Pinhole[5]	9/29/09	188.2/189.5, 199.8 (96)	201.83/244.2 (1038)[6,17]
IHD	3	Pinhole[5]	9/7/12	187.7/189.2, 199.7 (140)	217.0/241.7 (2312)
IHD	3	Pinhole[5]	9/7/12	188.1/189.6, 199.4 (128)	214.1/241.5 (2219)
IHD[18]	3	Pinhole[5]	9/7/12	187.4/188.7, 198.6 (92)	213.8/239.8 (2257)
IHD	4	Pinhole[5]	9/28/12	188.0/189.7, 199.7 (142)	217.0/242.0 (2313)
IHD	4	Pinhole[5]	9/28/12	188.1/189.6, 199.8 (137)	215.4/243.0 (2385)
IHD	4	Pinhole[5]	9/28/12	188.0/189.5, 200.1 (146)	216.0/241.8 (2237)
LLNL	1	Sealed[3]	12/01/09	187.4/188.9, 199.2 (125)	205/233.5 (3024)
LLNL	1	Sealed[3]	02/04/10	187.7/188.9, 198.8 (144)	205/235.6 (2880)
LLNL	1	Sealed[3]	02/04/10	187.6/189.1, 198.8 (125)	203/233.7 (2998)
LLNL	2	Sealed[3]	8/27/10	187.3/188.3, ~199[4] (126)	215.63/238.0 (3517)
LLNL	2	Sealed[3]	8/27/10	187.3/188.3, ~199[4] (132)	214.63/231.2 (3478)
LLNL	2	Sealed[3]	8/27/10	187.4/188.3, ~199[4] (114)	215.23/230.6 (3805)
LLNL	3	Sealed[3]	3/31/11	187.8/189.1, 199.4 (137)	220.0/244.0 (2003)
LLNL	3	Sealed[3]	3/3111	187.8/189.1, 199.2 (138)	217.4/237.3 (3168)
LLNL	4	Sealed[3]	5/19/11	187.8/188.9, 199.2 (143)	217.7/233.0 (3385)
LLNL	4	Sealed[3]	5/20/11	187.7/188.9, ~199[4] (138)	216.1/238.2 (2612)
LLNL	4	Sealed[3]	5/24/11	187.6/188.8, 198.8 (136)	217.4/232.7 (3314)
IHD	3	Sealed[7]	9/11/12	186.1/188.3, 198.5 (103)	210.5/237.9 (4310)[8]
IHD	3	Sealed[7]	9/11/12	187.6/188.9, 198.1 (100)	209.3/237.4 (4472)[9]
IHD	3	Sealed[7]	9/11/12	187.4/188.7, 198.6 (92)	213.8/239.8 (4306)[10]
IHD	4	Sealed[7]	10/1/12	187.5/190[4], 199.8 (123)	210.5/241.9 (4583)[11]
IHD	4	Sealed[7]	10/1/12	187.7/190[4], 199.7 (99)	211.8/240.5 (4203)[12]
IHD	4	Sealed[7]	10/1/12	187.8/190[4], 199.7 (92)	214.5/241.2 (4662)[13]

1. Onset of exothermic response reported to be obscured by endothermic response as indicated by software; 2.50 um laser drilled pinhole lid from TA Instruments; pinhole sample holder; 3. Sealed sample holder from TA Instruments; 4. Visually estimated from hard copy profile; 5. 75 um laser drilled pinhole lid from TA Instruments; 6. Pan break due to off gases; 7. Sealed, gold coated, high-pressure pans from SWISSI; 8. Additional peak on shoulder at 251.7°C; 9. Additional peak at 249.8°C; 10. Additional peak on shoulder at 250.3°C; 11. Additional peak on shoulder at 251.6°C; 12. Additional peak on shoulder at 252.5 °C; 13. Addotional peak at 251.8°C; 14. Data file labeled improperly: LANL, Set 3, Pinhole[5], 4,12,11, 188.6,198.8, 200.5,137 , 219.0,242.1,2148 [reference IDCA Data Report 033, April 13, 2011], changed after visual inspection of DSC profile; 15. In the calculation for the average values, the "~" sign wa dropped from the number and the number was taken at face value; 16. The exothermic onset overlaps with endothermic feature so onset values are approximate, no average for onset will be calculated, just a range; 17. Enthalpy data not used in average value calculations because of sample holder ruptured; 18. Value for exothermic enthalpy in original table incorrect: IHD, 3, Pinhole[5], 9/7/12, 187.4/188.7, 198.6 (92), 213.8/239.8 (4306);

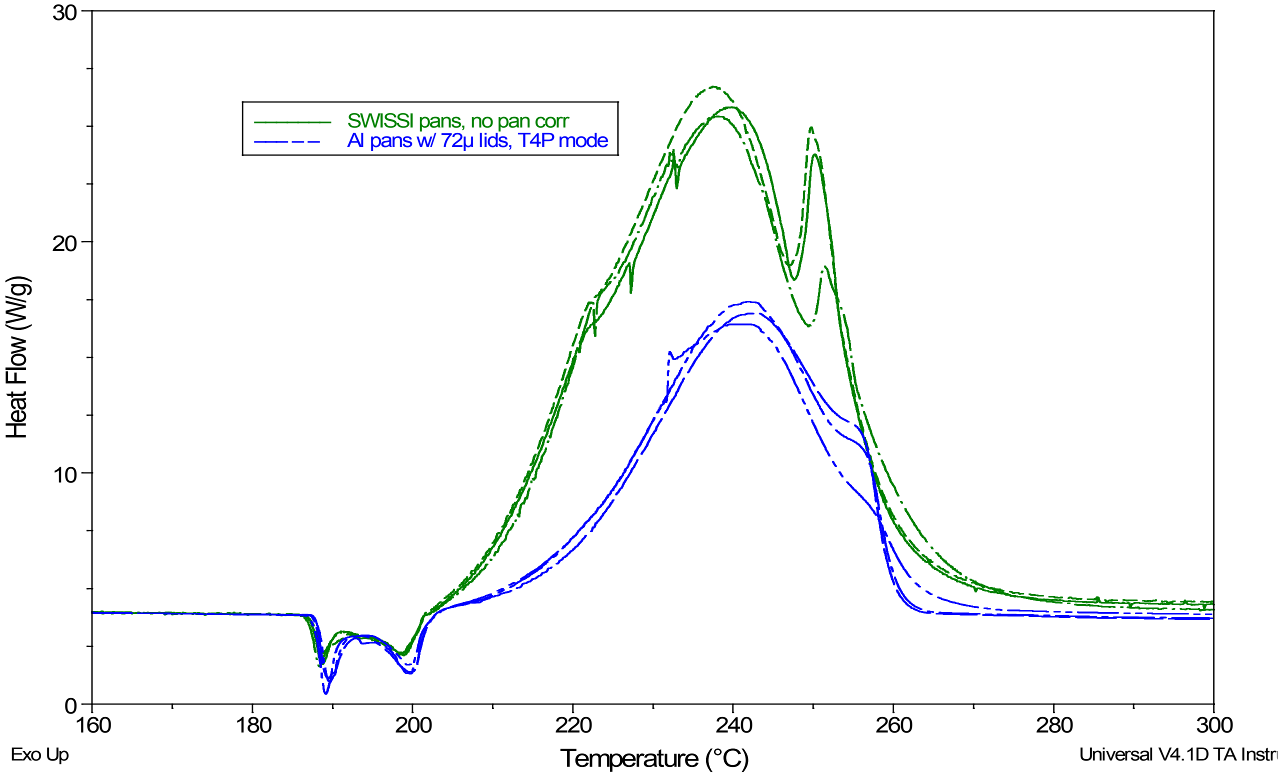

Figure 8. Example DSC scans of RDX in typical pinhole hermetic pans and one type of sealed pan.

Some insight into the results of Table 8 can be gained by grouping the data into categories and calculating simple averages and standard deviations. Table 9 summarizes these calculations. Table 8 was edited for these groupings in the following manner:

1. The enthalpy IHD data was not used from the measurements when pinhole sample holder ruptured,
2. Some temperatures were approximated directly from hard the hard copy of the profile and are designated with an "~" sign. For the calculations the number was taken at the face value,
3. Some data were grouped as LLNL pinhole data (old for older T Instruments equipment),
4. Some data were grouped as IHD and LANL pinhole data (new for newer TA Instruments equipment),
5. Some data were grouped as LLNL sealed sampled holder data (this is really a TA instrument sample without the pinhole—it is not really meant for high pressure work),
6. Some data were grouped as IHD sealed (which is the SWISSI pan which IS meant for high pressure work).
7. The transition between the endothermic and exothermic features is only listed with temperatures because the transition overlaps.

Examination of Table 9 shows that the pinhole and the sealed sample holders show differences in exothermic enthalpies—pinhole different than sealed; LLNL sealed different than the SWISSI sample holder. The sealed sample holders show higher enthalpies of decomposition in each case because they do not allow gas to escape. The pinhole sample holders allow gas to escape at a controlled rate and the escaping gas removes heat from the system, lowering the total observed enthalpy. Comparing the two sealed sample holders shows that they are distinctly different[32]. IHD used the SWISSI sample holder, the values for the enthalpy assigned to the exothermic feature are much higher than with the corresponding enthalpy for samples measured in the sealed sample holder that LLNL uses. This is probably due to the SWISSI sample holder is rated to hold 217 bar (3150 psi) at 400°C, while the LLNL sealed sample holder is not pressure rated. As a result, the LLNL sample holder probably allows some volatile gases escape, therefore cooling the sample.

The Endothermic enthalpy is less for the SWISSI sample holder than for the TA sample holder (maybe)—probably due to the more massive SWISSI sample holder. The SWISSI sample holder weighs more (~1 g)and the mass is distributed differently because of the sealed design than the TA sample holder (~0.1 g). As a result there is a shift of the response to slightly later times that effects the endothermic response (because of overlap with the much larger exothermic feature), so some of the response is lost in the exothermic transition. This has been seen before in the DSC profiles of AN[43].

The temperature of maximum exothermic enthalpy is lower for the sealed pans than for the open pans. Two mechanisms occur in the pinhole sample holders during this heating period that are counter to each other—evaporation which is endothermic, and decomposition which is exothermic. These mechanism compete causing the temperate of maximum exothermic enthalpy to be higher in the pinhole sample holder case.

Table 9. Ranges of DSC Parameters for RDX.

Parameter[1]	Pinhole Old	Pinhole New	Sealed LLNL	Sealed IHD
Endothermic Onset, ˚C Range (Average)	187.3-187.8 (187.7 ± 0.2)	187.4-188.6 (188.0 ± 0.2)	187.3-187.8 (187.6 ± 0.2)	186.1-187.8 (187.4 ± 0.6)
Endothermic Min., ˚C Range (Average)	188.3-189.2 (188.9 ± 0.3)	188.7-189.9 (189.4 ± 0.3)	188.3-189.1 (188.8 ± 0.3)	188.3-190.0 (189.3 ± 0.8)
Endothermic Min., °C Range (Average)	198.8-200.0 (199.2 ± 0.3)	198.6-200.8 (199.9 ± 0.5)	198.8-199.4 (199.0 ± 0.2)	198.1-199.8 (199.1 ± 0.8)
Endothermic Enthalpy, J/g Range (Average)	126-181 (142 ± 15)	92-146 (128 ± 14)	114-144 (133 ± 9)	92-123 (102 ± 11)
Exothermic Onset, ˚C Range[2]	203-219	201-225	203-220	209-215
Exothermic Max., ˚C Range (Average)	238.7-243.5 (241.6 ± 1.4)	239.8-244.2 (242.3 ± 1.0)	230.6-244.0 (235.3 ± 3.9)	237.4-241.9 (239.8 ± 1.8)
Exothermic Enthalpy, J/g Range (Average)	1890-2432 (2244 ± 177)	1947-2385[3] (2174 ± 120)	2003-3805 (3108 ± 495)	4203-4662 (4423 ± 179)

1. Onset is the beginning of the maximum or minimum as automatically identified by the equipment, endothermic min. is the minimum temperature of the endothermic feature, endothermic enthalpy is the overall enthalpy of the two overlapping endothermic features, exothermic max. is the maximum of the exothermic feature; 2. Range only because the transition between the endothermic and exothermic features overlap; 3. Two values from IHD Set 2 discarded due to sample holder rupturing during experiment.

4 DISCUSSION

The analyses above allows an assessment of the statistical differences among participants, average values, expected ranges, percent variability, dependence on method or environment, and possible causes for the differences that are observed. These are summarized in Table 10 for the various sensitivity and thermal tests. DH_{50} data from AFRL was excluded from this table since it showed up as a separate group in both Tukey and Fisher comparisons. No other data was excluded from this table.

Table 10. Results of Statistical Analyses of IDCA Small Scale Safety Testing of RDX.

	Equivalent results?, p-value from ANOVA	Average	Range	Percent Variability	Dependence on method or environment variables	Possible causes of differences
Impact DH_{50}	No p=0.000	21.5 cm	15-26.5 cm	27	Possibly grit	Operator, Detection method
BAM friction F_{50}	No p=0.001	21.0 kg	14.9-31.6 kg	40	No	Operator, Background noise
BAM friction TIL	No p=0.000	14.2 kg	9.6-21.6 kg	42	No	Operator, Background noise
BAM friction TIL+	No p=0.004	16.0 kg	12-24 kg	38	No	Operator, Background noise
ESD TIL	No p=N/A	0.051 J	0.025-0.095 J	N/A	Possibly RH	Detection method, Age of instrument
ESD TIL+	No p=N/A	0.099 J	0.0625-0.165 J	N/A	Possibly RH	Detection method, Age of instrument
DSC thermal	Yes p=N/A	See Table 9	See Table 9	N/A	No	Sample holder type,

The information in Table 10 shows that there is a statistically significant difference among the participants in all of the tests, whether evaluated by ANOVA or inferred by examination of the specific test results presented above. This is not pointing out deficiencies in the test methods but is highlighting the variability that can result from individual laboratories implementing detailed procedures within bounding facility conditions and within testing guidelines established from previous experience. This variability can be detected and quantified and may be used to determine whether a new laboratory is capable of making equivalent measurements if that becomes a goal for future directions.

For present IDCA purposes, the statistical difference also implies that it may be possible to ultimately determine the cause of variability in the results if enough details about the testing are tracked during future round-robin examinations. All of the test parameters, instrument details, sample characteristics and environment conditions will be important to track if this goal is undertaken.

The information in Table 10 also shows the ranges that can be expected for other materials with DH_{50}, F_{50} or TIL values near those for RDX. Translated to percent variability, these can suggest what might be expected when testing material at much higher or lower sensitivity values. This will be important to help understand whether differences observed with HME materials in future reports are truly significant.

The largest factors causing differences among participants appear to be the operator, method of detection, and testing environment. Sometimes these are inextricably linked, such as when the operator *is* the method of detection and their perception is limited by background noise in the laboratory. This is the case for BAM friction in which LANL vs. LLNL differences are due to operator and environment. For these tests there is no transducer, the LLNL friction machine has more shielding, and it is run with a vent fan during use. In other cases, such as with the DH_{50} differences between LLNL and LANL, only detection method and environment play a role since both participants use threshold sound levels to make Go vs. No-Go determinations. Differences in how these threshold levels are chosen may create an offset between DH_{50} values. The issues associated with operator-influenced results are being addressed informally at various testing laboratories through implementation of transducer-driven Go / No-Go discrimination and more formally by commercial entities such as Safety Management Services, which is developing full systems that integrate the test instrument, electronic detection methods, and result analysis.

5 CONCLUSIONS

The RDX results of this report validate the former assessment that HME materials evaluated by SSST testing are sensitive to the differences in the test methods and equipment employed by each laboratory. This further accentuates the expectations that differing evaluations of sensitivity are significant from a safety standpoint. Some of these differences can be eliminated by standardization, but others are inherent in the configurations and environments each laboratory has established to safely test energetic materials. Elimination of the differences will require further research, however. This work has shown that, even when a specific standard is carefully tested, variation in results occur and that it is important to be able to test materials under a variety of conditions because of the multiple types of insults possible to these materials. Exploring a range of variables provides the best chance of probing the particular set of test parameters that highlight the extent of sensitivity of the material. Sandpaper properties, striker mass, and the method of detecting the generated sound or reaction are all examples of important variables, and parameter variation is the topic of subsequent papers.

REFERENCES

1. Integrated Data Collection Analysis (IDCA) Program—Proficiency Study for Small Scale Safety Testing of Homemade Explosives, B. D. Olinger, M. M. Sandstrom, K. F. Warner, D. N. Sorensen, D. L. Remmers, J. S. Moran, T. J. Shelley, L. L. Whinnery, P. C. Hsu, R. E. Whipple, M. Kashgarian, and J. G. Reynolds, *IDCA Program Analysis Report* **001**, LLNL-TR-416101, December 3, 2009.
2. Integrated Data Collection Analysis (IDCA) Program—RDX Standard, Data Set 1, M. M. Sandstrom, G. W. Brown, D. N. Preston, C. J. Pollard, K. F. Warner, D. N. Sorensen, D. L. Remmers, J. S. Moran, T. J. Shelley, P. C. Hsu, R. E. Whipple, and J. G. Reynolds, *IDCA Program Analysis Report* **006**, LLNL-TR-479891, April 19, 2011.
3. Integrated Data Collection Analysis (IDCA) Program—Drying Procedures, B. D. Olinger, M. M. Sandstrom, G. W. Brown, K. F. Warner, D. N. Sorensen, D. L. Remmers, J. S. Moran, T. J. Shelley, L. L. Whinnery, P. C. Hsu, R. E. Whipple, and J. G. Reynolds, *IDCA Program Analysis Report* **004**, LLNL-TR-465872, April 27, 2010.

4. Integrated Data Collection Analysis (IDCA) Program—Mixing Procedures and Materials Compatibility, B. D. Olinger, M. M. Sandstrom, K. F. Warner, D. N. Sorensen, D. L. Remmers, J. S. Moran, T. J. Shelley, L. L. Whinnery, P. C. Hsu, R. E. Whipple, M. Kashgarian, and J. G. Reynolds, *IDCA Program Analysis Report* **002**, LLNL-TR-422028, December 27, 2009.
5. Holston Army Ammunition Plant, Kingsport, Tennessee.
6. RDX Chemical Analysis, D. N. Sorensen, K. B. Proctor, I. B. Choi, L. Tinsley, *IDCA Program Data Report* **009**, October 13, 2009.
7. RDX Particle Size, K. F. Warner, *IDCA Program Data Report* **056**, October 5, 2009.
8. Small Scale Safety Test Report for RDX (third in a series), P. C. Hsu and J. G. Reynolds, *IDCA Program Data Report* **068**, LLNL-TR-484412, May 16, 2011.
9. Small Scale Safety Test Report for RDX (4th in a series) revised to include 1-kg Data, P. C. Hsu and J. G. Reynolds, *IDCA Program Data Report* **080**, LLNL-TR-505971, January 23, 2012.
10. 50188 I RDX 3rd Time, M. M. Sandstrom and G. W. Brown, *IDCA Program Data Report* **033**, April 13, 2011.
11. 50188 V RDX 4th Time, M. M. Sandstrom and G. W. Brown, *IDCA Program Data Report* **065**, May 16, 2011.
12. RDX Report Run #3, D. L. Remmers, D. N. Sorensen, K. F. Warner, *IDCA Program Data Report* **117**, November 14, 2012.
13. Integrated Data Collection Analysis (IDCA) Program—SSST Testing Methods, B. D. Olinger, M. M. Sandstrom, G. W. Brown, K. F. Warner, D. N. Sorensen, D. L. Remmers, J. S. Moran, T. J. Shelley, L. L. Whinnery, P. C. Hsu, R. E. Whipple, and J. G. Reynolds, *IDCA Program Analysis Report* in preparation.
14. Reduced Sensitivity of RDX (RS-RDX) Part I: Literature Review and DSTO Evaluation, I. J. Lochert, M. D. Franson and B. L. Hamshere, DSTO-TR-1447, DSTO Systems Sciences Laboratory, July 2003.
15. Sigma-Aldrich Chemical Technical Library Particle Size Conversion Table, http://www.sigmaaldrich.com/chemistry/stockroom-reagents/learning-center/technical-library/particle-size-conversion.html.
16. Detailed Specification RDX (Cyclotrimethylenetrinitramine), Military Specification, MIL-DTL-398D, December 12, 1996.
17. Integrated Data Collection Analysis (IDCA) Program—RDX Set 2, M. M. Sandstrom, G. W. Brown, D. N. Preston, C. J. Pollard, K. F. Warner, D. N. Sorensen, D. L. Remmers, J. J. Phillips, T. J. Shelley, J. A. Reyes, P. C. Hsu, and J. G. Reynolds, *IDCA Program Analysis Report* **020**, LLNL-TR-619637 (725352), February 20, 2013.
18. A Method for Obtaining and Analyzing Sensitivity Data, W. J. Dixon and A.M. Mood, *J. Am. Stat. Assoc.,* 43, 109-126, 1948.
19. The Bruceton method also assumes that testing begins in the vicinity of the mean. Often this is not true and the initial testing to home in on the mean can skew the statistics. In practice, a "Modified" Bruceton method is used in which initial tests are used to bracket the mean before starting to count Go's and No-Go's. This is used by LANL in this work.
20. D-Optimality-Based Sensitivity Test, B. T. Neyer, *Technometrics*, **36**, 48-60, 1994.
21. Department of Defense Ammunition and Explosives Hazard Classification Procedures, TB 700-2 NAVSEAINST 8020.8B TO 11A-1-47 DLAR 8220.1, January 5, 1998.
22. Integrated Data Collection Analysis (IDCA) Program—PETN Class 4 Standard, M. M. Sandstrom, G. W. Brown, D. N. Preston, C. J. Pollard, K. F. Warner, D. N. Sorensen, D. L. Remmers, T. J. Shelley, P. C. Hsu, and J. G. Reynolds, *IDCA Program Analysis Report* **017**, LLNL-TR-568299 (634352), August 1, 2012.
23. Challenges of Small-Scale Safety and Thermal Testing of Home Made Explosives—Results from the Integrated Data Collection Analysis (IDCA) Program Proficiency Test, J. G. Reynolds, M. M. Sandstrom, G. W. Brown, K. F. Warner, T. J. Shelley, P. C. Hsu, *IDCA Program Presentation* **009**, LLNL-PRES-547780, May 2, 2012.
24. Integrated Data Collection Analysis (IDCA) Program—$KClO_3$/Icing Sugar (-100) mixture**,** M. M. Sandstrom, G. W. Brown, D. N. Preston, C. J. Pollard, K. F. Warner, D. N. Sorensen, D. L. Remmers, T. J. Shelley, P. C. Hsu, R. E. Whipple, and J. G. Reynolds, *IDCA Program Analysis Report* **007,** (2011), LLNL-TR-482149, May 10. 2011.
25. Integrated Data Collection Analysis (IDCA) Program—$KClO_3$ (as received)/Icing Sugar**,** M. M. Sandstrom, G. W. Brown, D. N. Preston, C. J. Pollard, K. F. Warner, D. N. Sorensen, D. L. Remmers, T. J. Shelley, P. C. Hsu and J. G. Reynolds, *IDCA Program Analysis Report,* **011**, LLNL-TR-484715, May 26, 2011.
26. Integrated Data Collection Analysis (IDCA) Program—$KClO_3$/dodecane Mixture, M. M. Sandstrom, G. W. Brown, D. N. Preston, C. J. Pollard, K. F. Warner, D. N. Sorensen, D. L. Remmers, J. S. Moran, T. J. Shelley, L. L. Whinnery, P. C. Hsu, and J. G. Reynolds, *IDCA Program Analysis Report* **012**, LLNL-TR-484788, June 21, 2011.
27. Integrated Data Collection Analysis (IDCA) Program—$KClO_4$/dodecane Mixture, M. M. Sandstrom, G. W. Brown, D. N. Preston, C. J. Pollard, K. F. Warner, D. N. Sorensen, D. L. Remmers, T. J. Shelley, L. L. Whinnery, P. C. Hsu, and J. G. Reynolds, *IDCA Program Analysis Report* **015**, LLNL-TR-522941, May 11, 2012.
28. Integrated Data Collection Analysis (IDCA) Program—$KClO_4$/Carbon, M. M. Sandstrom, G. W. Brown, D. N. Preston, C. J. Pollard, K. F. Warner, D. N. Sorensen, D. L. Remmers, T. J. Shelley, J. A. Reyes, P. C. Hsu, and J. G. Reynolds, *IDCA Program Analysis Report* **018**, LLNL-TR-614974 (717752), January 31, 2013.
29. Integrated Data Collection Analysis (IDCA) Program—$KClO_4$/Aluminum Mixture, M. M. Sandstrom, G. W. Brown, D. N. Preston, C. J. Pollard, K. F. Warner, D. N. Sorensen, D. L. Remmers, T. J. Shelley, P. C. Hsu, and J. G. Reynolds, *IDCA Program Analysis Report* **013**, LLNL-TR-518531, December 5, 2011.

30. Integrated Data Collection Analysis (IDCA) Program—$NaClO_3$/Icing Sugar, M. M. Sandstrom, G. W. Brown, D. N. Preston, C. J. Pollard, K. F. Warner, D. N. Sorensen, D. L. Remmers, T. J. Shelley, P. C. Hsu, R. E. Whipple, and J. G. Reynolds, *IDCA Program Analysis Report* **019**, LLNL-TR-617403 (721773), February 11, 2013..
31. Challenges of Small-Scale Safety and Thermal Testing of Home Made Explosives—Final Results from the Integrated Data Collection Analysis (IDCA) Program Proficiency Test, J. G. Reynolds, M. M. Sandstrom, G. W. Brown, K. F. Warner, J. J. Phillips, T. J. Shelley, J. A. Reyes, P. C. Hsu, *IDCA Program Presentation* **011**, LLNL-PRES-612072 (712792), January 17, 2013.
32. Methods report.
33. Small Scale Safety Test Report for RDX (second calibration), P. C. Hsu and J. G. Reynolds, *IDCA Program Data Report* **053**, LLNL-TR-480177, April 6, 2011.
34. 50188 I RDX Second Run, M. M. Sandstrom and G. W. Brown, *IDCA Program Data Report* **032**, April 6, 2011.
35. RDX Report Run #2, D. L. Remmers, D. N. Sorensen, and K. F. Warner, *IDCA Program Data Report* **103**, May 13, 2011.
36. SNL Small-Scale Sensitivity Report: RDX, J. J. Phillips, *IDCA Program Data Report* **104**, June 12, 2012.
37. IHD RDX #4
38. Small Scale Safety Test Report for IDCA—RDX, P. C. Hsu and J. G. Reynolds, *IDCA Program Data Report* **037**, LLNL-TR-426345, March 25, 2010
39. RDX 51088_rev 0, M. M. Sandstrom, D. N. Preston, *IDCA Program Data Report* **012** November 24, 2009.
40. RDX Report, D. L. Remmers, D. N. Sorensen, J. S. Moran, K. F. Warner, *IDCA Program Data Report* **001**, February 24, 2010.
41. RDX, Integrated Data Collection Analysis (IDCA) Program, Small Scale Safety Testing (SSST), J. A. Reyes, *IDCA Program Data Report* **036**, July 1, 2010.
42. Humidity and ESD Control, R. C. Allen, ESD Systems, Desco Industries, Inc. Report, April 1998, http://esdsystems.descoindustries.com/whitepapers/wp_humidity.html.
43. AN Report
44. http://www.sms-ink.com
45. [Neyer 1994] A D-Optimality Based Sensitivity Test, B. T. Neyer, *Technometrics*, 36, 61-70, 1994.
46. Devore 2012] Probability and Statistics for Engineering and the Sciences, 8th ed., J. L. Devore, Brooks/Cole pub., Boston, MA, 2012.
47. [Minitab] Minitab Inc., 1829 Pine Hall Rd, State College, PA, USA 16801. www.minitab.com
48. [Fisher 1935] Design of Experiments, R. A. Fisher, Oliver & Boyd, London, 1935.

ABREVIATIONS, ACRONYMS AND INITIALISMS

-100	Solid separated through a 100-mesh sieve
ABL	Allegany Ballistics Laboratory
AFRL	Air Force Research Laboratory, RXQL
Al	Aluminum
AR	As received (separated through a 40-mesh sieve)
ARA	Applied Research Associates
BAM	German Bundesanstalt für Materialprüfung Friction Apparatus
C	Chemical symbol for carbon
CAS	Chemical Abstract Services registry number for chemicals
cm	centimeters
DH_{50}	The height the weight is dropped in Drop Hammer that cause the sample to react 50% of the time, calculated by the Bruceton or Neyer methods
DHS	Department of Homeland Security
DSC	Differential Scanning Calorimetry
DTA	Differential Thermal Analysis
ESD	Electrostatic Discharge
F_{50}	The weight or pressure used in friction test that cause the sample to react 50% of the time, calculated by the Bruceton or Neyer methods
fps	feet per second
H	Chemical symbol for hydrogen
H_2O	Chemical formulation for water
HME	homemade explosives or improvised explosives
HMX	Her Majesty's Explosive, cyclotetramethylene-tetranitramine

IDCA	Integrated Data Collection Analysis
IHD	Indian Head Division, Naval Surface Warfare Center
j	joules
$KClO_3$	Potassium Chlorate
$KClO_4$	Potassium Perchlorate
kg	kilograms
LANL	Los Alamos National Laboratory
LLNL	Lawrence Livermore National Laboratory
MBOM	Modified Bureau of Mines
N	Chemical symbol for nitrogen
$NaClO_3$	Sodium Chlorate
NSWC	Naval Surface Warfare Center
O	Chemical symbol for oxygen
PETN	Pentaerythritol tetranitrate
psig	pounds per square inch, gauge reading
RDX	Research Department Explosive, 1,3,5-Trinitroperhydro-1,3,5-triazine
RH	Relative humidity
RT	Room Temperature
RXQL	The Laboratory branch of the Airbase Sciences Division of the Materials & Manufacturing Directorate of AFRL
s	Standard Deviation
SEM	Scanning Electron Micrograph
Si	silicon
SNL	Sandia National Laboratories
SSST	small-scale safety and thermal
TGA	Thermogravimetric Analysis
TIL	Threshold level—level before positive event

ACKNOWLEDGMENTS

The IDCA thanks Heidi Turner (LLNL) and Gary Hust (LLNL) for supporting data in this document.

This work was performed by the Integrated Data Collection Analysis (IDCA) Program, a five-lab effort supported by Los Alamos National Laboratory, Lawrence Livermore National Laboratory, Sandia National Laboratories, the Air Force Research Laboratory, and Indian Head Division, Naval Surface Warfare under sponsorship of the US Department of Homeland Security, Office of Science and Technology, Energetics Division. Los Alamos National Laboratory is operated by Los Alamos National Security, LLC, for the United States Department of Energy under Contract DE-AC52-06NA25396. Sandia is a multi-program laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under Contract DE-AC04-94AL85000. This work was performed under the auspices of the U. S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344. The Air Force Research Laboratory (AFRL/RXQF) and Indian Head Division, Naval Surface Warfare (NSWC IHD) also performed work in support of this effort. The work performed by AFRL/RXQL and NSWC IHD is under sponsorship of the US Department of Homeland Security, Office of Science and Technology, Energetics Division.

Questions:

Comments from author JGR

1. ANOVA analysis of data in appendix?
2. I have switched to the TIL and TIL+ convention in the other analysis reports, so I have changed that in the text and tables
3. Is there a reference for box plots? It seems to just be a nice representation of the median, mean and deviation, but does the width of the box have any significance?
4. In the experimental, there needs to be a description with references of the data analysis methods (ANOVA, Tukey, Fisher, etc.)
5. For my own clarification: in the ANOVA analyses of the BAM friction data, the F50 results for both Tukey and Fisher methods show that LLNL and IHD are the same, LLNL and LANL are different, and IHD and LANL are different. For the TIL, (leaving SNL out), LLNL is different than LANL, LLNL is different than IHD, LANL and IHD are the same. For TIL+ (leaving SNL out), shows the same as TIL. Combining all these analysis indicates that everyone is different. Is this correct?
6. There are no plots of BAM friction data vs. temperature and humidity. You refer to them in the text, but I think you left them out, or my computer screwed up the file.
7. In the figure of the DH_{50} values vs. striker weight, there are data points for 3 kg and 4 kg striker weights. I don't think we have striker weight data like that. Do we?
8. In the discussion of the DSC values, I checked most of the data, and it looks OK. I added the IHD Set 4 to the table. Here are the average values of the averages. Note All means no sealed sample holder data, and All H means only sealed sample holder data.

Participant[1,2]	T_{min} of En_1[3], °C	T_{min} of En_2[4], °C	ΔH of En_{1+2}[5], J/g	T_{max} of Ex_1[6], °C	ΔH of Ex_1[7], J/g
LLNL All	188.9 ± 0.3	199.1 ± 0.1	142 ± 12	239.6 ± 3.5	2469 ± 449
LLNL All H	188.9 ± 0.4	199.1 ± 0.2	133 ± 7	237.6 ± 4.6	2839 ± 596
LANL All	190.3 ± 2.7 (1.4)	200.2 ± 0.5 (0.2)	132 ± 7 (5)	242.6 ± 0.8 (0.3)	2150 ± 102 (5)
IHD All	189.3 ± 0.3	199.4 ± 0.4	124 ± 17	241.7 ± 0.7	2127 ± 721
IHD All H	189.3 ± 1.0	199.1 ± 0.9	99 ± 1	239.8 ± 2.0	4423 ± 85

I don't see any difference in the data except for the IHD sealed sample data. Both enthalpies are different than the rest of the pack. I think there is a hidden variable that I did not describe adequately in the table. The sealed sample holder than IHD uses the SWISSI pressure cell. This is a heavy-duty gold plated sample holder that has a lot of mass and is pressure rated. The sealed samples holder that LLNL uses for these measurements (not for the SETARAM) is the standard cell with a lid that does not have a laser-drilled hole. It really is not pressure rated, but does hold some pressure. The low temperature behavior of these two cells are vastly different. The extra mass of the SWISSI cell shows up as less endothermic enthalpy as well (we saw this in the AN report, not issued yet). The high temperature enthalpy is different because I don't believe the LLNL cell really hold pressure that well, so some gases escape. Yes these hold a little better, but not that much. You have to go to the official pressure cells to get the real number

Disclaimer

Lawrence Livermore National Laboratory is operated by Lawrence Livermore National Security, LLC, for the U.S. Department of Energy, National Nuclear