

**NISTIR 8087**

**Feasibility Study on  
NIST Room Temperature  
Low-Energy and High-Energy  
Verification Specimens**

Enrico Lucon  
Chris N. McCowan  
Ray L. Santoyo

This publication is available free of charge from:  
<http://dx.doi.org/10.6028/NIST.IR.8087>

**NIST**  
**National Institute of  
Standards and Technology**  
U.S. Department of Commerce

**NISTIR 8087**

# **Feasibility Study on NIST Room Temperature Low-Energy and High-Energy Verification Specimens**

Enrico Lucon  
Chris N. McCowan  
Ray L. Santoyo  
*Applied Chemicals and Materials Division  
Material Measurement Laboratory*

This publication is available free of charge from:  
<http://dx.doi.org/10.6028/NIST.IR.8087>

October 2015



U.S. Department of Commerce  
*Penny Pritzker, Secretary*

National Institute of Standards and Technology  
*Willie May, Under Secretary of Commerce for Standards and Technology and Director*

Certain commercial entities, equipment, or materials may be identified in this document in order to describe an experimental procedure or concept adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the entities, materials, or equipment are necessarily the best available for the purpose.

**National Institute of Standards and Technology Internal Report 8087**  
**Natl. Inst. Stand. Technol. Int. Report 8087, 24 pages (October 2015)**  
**CODEN: NTNOEF**

**This publication is available free of charge from:**  
**<http://dx.doi.org/10.6028/NIST.IR.8087>**

## **Abstract**

The feasibility of certifying Charpy reference specimens for testing at room temperature ( $21\text{ }^{\circ}\text{C} \pm 1\text{ }^{\circ}\text{C}$ ) instead of  $-40\text{ }^{\circ}\text{C}$  was demonstrated at NIST by performing 130 room-temperature tests from five low-energy and four high-energy lots of steel on the three master machines located in Boulder, CO. The statistical analyses performed show that in most cases the variability of results (*i.e.*, the experimental scatter) is reduced when testing at room temperature. For eight out of the nine lots considered, both the coefficient of variation and the sample size were lower at  $21\text{ }^{\circ}\text{C}$  than at  $-40\text{ }^{\circ}\text{C}$ .

The results of this study will allow NIST to satisfy requests for room-temperature Charpy verification specimens that have been received from customers for several years: testing at  $21\text{ }^{\circ}\text{C}$  removes from the verification process the operator's skill in transferring the specimen in a timely fashion from the cooling bath to the impact position, and puts the focus back on the machine performance. For NIST, it also reduces the time and cost for certifying new verification lots.

For one of the low-energy lots tested with a C-shaped hammer, we experienced two specimens jamming, which yielded unusually high values of absorbed energy. For both specimens, the signs of jamming were clearly visible. Jamming is slightly more likely to occur at  $21\text{ }^{\circ}\text{C}$  than at  $-40\text{ }^{\circ}\text{C}$ , since at room temperature low-energy samples tend to remain in the test area after impact rather than exiting in the opposite direction of the pendulum swing. In the evaluation of a verification set, any jammed specimen should anyway be removed from the analyses.

## **Keywords**

Charpy master machines; Charpy reference specimens; coefficient of variation; high-energy specimens; indirect verification; low-energy specimens; room temperature; sample size; specimen jamming.

# Table of Contents

Abstract.....	iii
Keywords.....	iii
Table of Contents.....	iv
1. Introduction.....	1
2. Material, Test Equipment, and Test Matrix.....	3
3. Statistical Analyses.....	4
3.1 Test Result Statistics.....	4
3.2 Machine Statistics.....	5
3.3 Additional Statistics.....	5
3.4 Criteria for Assessing the Feasibility of RT Verification Specimens.....	6
4. Results.....	7
4.1 Lots Tested on the SI Machine.....	7
4.1.1 LL-139.....	7
4.1.2 HH-143.....	8
4.2 Lots Tested on the TO Machine.....	9
4.2.1 LL-138.....	9
4.2.2 HH-140.....	10
4.3 Lots Tested on the TK Machine.....	11
4.3.1 LL-119.....	11
4.3.2 LL-133.....	12
4.3.3 HH-136.....	14
4.4 Lots Tested on the Three Master Machines (SI, TK, TO).....	15
4.4.1 LL-140.....	15
4.4.2 HH-149.....	17
5. Discussion.....	19
5.1 Relationships between Results at $-40\text{ }^{\circ}\text{C}$ and $21\text{ }^{\circ}\text{C}$ .....	19
5.2 Specimen Jamming at the Low-Energy Level.....	20
6. Conclusions.....	23
References.....	24

## 1. Introduction

Charpy impact testing is frequently specified as an acceptance test for structural materials, and all companies performing acceptance tests are expected to periodically verify the performance of their Charpy impact machines. According to the ASTM E23-12c standard [1], the procedure for verifying the performance of Charpy machines consists of a physical part (*direct verification*) and an engineering part (*indirect verification*).

The direct verification corresponds to a detailed evaluation of the machine dimensions, alignment, etc., while the indirect verification of the machine performance is carried out by breaking sets of Charpy reference specimens with certified values of absorbed energy. The indirect verification procedure was added to ASTM E23 more than 50 years ago, when it was ascertained that direct verification alone could not explain certain unacceptable differences (as much as 100 %) among the results of the machines tested. Since some of the differences originated from interactions between the machine components and the specimens, only actual Charpy tests on reference specimens could resolve these effects [2].

Currently and for the last 26 years, NIST in Boulder has supplied impact reference specimens as a Standard Reference Material (SRM), which is used to indirectly verify the performance of Charpy machines in accordance with ASTM E23. Historically, the Charpy verification program was developed by the U.S. Army (Watertown Arsenal, AMMRC) that produced and distributed reference specimens for the verification of Charpy machines in the United States. The Army procedures were adopted by ASTM in their E23 standard in 1956 (ASTM E23-56T). As a result of the adoption of the E23 procedures and requirements, the differences between the Charpy machines of the Army contractors were reduced to 1 ft-lb (1.4 J) or 5 %, whichever was greater [2].

The Charpy verification program was taken over by NIST in 1989, and Army personnel helped to transfer the reference Charpy machines and their evaluation procedures to NIST. The three reference Charpy machines have been defined in ASTM E23 as the “master Charpy impact machines” for 25 years [3]. Each year, the NIST program evaluates the indirect verification test results of over 1,500 industrial machines. If the test results of an industrial machine agree with results of the NIST master machines within 1.4 J or 5 %, whichever is greater, the machine is certified for acceptance testing according to the requirements of ASTM Standard E23.

Currently, besides NIST, there are three other NMIs (National Metrology Institutes) in the world who certify and distribute reference Charpy specimens for the indirect verification of impact machines:

- *The Joint Research Center Institute for Reference Materials and Measurements (JRC-IRMM)* of the European Commission, located in Geel (Belgium). Their Charpy verification specimens cover four levels of absorbed energy, corresponding approximately to 25 J, 80 J, 120 J, and 150 J. All specimens must be tested at room temperature (RT, 20 °C), even though one batch of low-energy specimens also has certified values at 0 °C to avoid jamming [4]. Tests are performed and evaluated in accordance with ISO 148-2:2008 [5].
- *Laboratoire National de Métrologie et d’Essais (LNE)*, located in Trappes near Paris (France). Their reference specimens cover five absorbed energy levels, namely: low (approx. 25 J), medium (70 J to 80 J), high 1 (115 J to 125 J), high 2 (160 J to 175 J), and

super high (200 J to 220 J). All specimens have to be tested at room temperature (20 °C) in accordance with ISO 148-2:2008.

- *The National Metrology Institute of Japan* (NMIJ), located in Tsukuba (Japan). NMIJ (then called National Research Laboratory of Metrology, NRLM) used to certify and distribute Charpy reference specimens of steel corresponding to different absorbed energy levels [6]. However, a recent internet search has shown that at the time of writing, only reference materials for the Charpy impact strength of plastics (PVC and PMMA) are available from NMIJ [7].

To the authors' knowledge, similar Charpy verification programs are expected to be launched soon by other NMIs in the world, such as the Shanghai Research Institute of Materials (SRIM, China), the National Institute of Metrology, Quality and Technology (Inmetro, Brazil), and the National Physical Laboratory (CSIR-NPL, India).

Since the time the U.S. Charpy verification program was run by the Army, verification specimens have to be tested at  $-40\text{ }^{\circ}\text{C}$  ( $-40\text{ }^{\circ}\text{F}$ ) for the levels corresponding to low energy (14 J – 20 J at  $-40\text{ }^{\circ}\text{C}$ , SRM 2092) and high energy (88 J – 136 J at  $-40\text{ }^{\circ}\text{C}$ , SRM 2096). Initially, these two energy levels were the only ones available. With the development of new steels that have higher toughness and impact strength, a third absorbed energy level (super-high energy, approximately 200 J) was introduced in the mid-90s, following customers' demand. Super-high-energy verification specimens made from an 18 Ni, cobalt-strengthened maraging steel designated as T-200, were certified at NIST for testing at room temperature ( $21\text{ }^{\circ}\text{C} \pm 1\text{ }^{\circ}\text{C}$ ).

For many years, NIST has been approached by customers with requests and discussions on the possibility of certifying Charpy specimens for low- and high-energy verification at room temperature instead of  $-40\text{ }^{\circ}\text{C}$ . If the test is performed at room temperature, the operator's skill in transferring the specimen from the temperature bath in less than 5 seconds is removed from the verification test. The same applies for other ancillary experimental components, such as the accuracy and the calibration state of the temperature-measuring equipment. Therefore, it can be contended that the focus of the verification test is solely on the machine performance. Additional advantages of room-temperature SRMs are:

- for the customer, the need to invest in cooling-bath equipment is removed, if not needed for general testing;
- for NIST, the time and cost for the certification of a room-temperature lot is significantly reduced with respect to a  $-40\text{ }^{\circ}\text{C}$  lot.

The feasibility study described in this Internal Report is aimed at evaluating the possibility of providing our customers with the option of conducting their verification tests at room temperature ( $21\text{ }^{\circ}\text{C} \pm 1\text{ }^{\circ}\text{C}$ ) rather than at  $-40\text{ }^{\circ}\text{C}$ . As detailed above, this would put us in line with the remaining producers of Charpy verification specimens (NMIs), with the exception of the  $0\text{ }^{\circ}\text{C}$ , low-energy batch provided by IRMM.

To justify this change and satisfy our customers' request, we need to evaluate the influence that testing at RT would have of the variation (scatter) of our low- and high-energy SRMs. If this variation can be matched or even reduced, RT Charpy verification specimens could (and should) be produced and made available to our customer base.

## 2. Material, Test Equipment, and Test Matrix

Both the low- and high-energy SRMs are made from AISI 4340 steel bars from a single heat to minimize compositional and microstructural variations. The nominal composition of the 4340 steel is presented in Table 1.

*Table 1 – Chemical composition of 4340 steel, wt %.*

C	Si	Mn	P	S	Mo	Ni	Cr
0.4	0.28	0.66	0.004	0.001	0.28	1.77	0.83

The steel is produced by a double-vacuum-melting procedure (vacuum-induction-melt and vacuum-arc-remelt), in order to minimize elements such as P, S, Va, Nb, Ti, and Cu.

Ingots are forged, hot-rolled, and cold-finished to 12.7 mm square bars, and finally annealed. The maximum acceptable grain size is ASTM #8. The bars are then normalized at 950 °C and hardened to approximately 35 HRC (Rockwell Hardness C).

To produce different levels of Charpy absorbed energy, the steel is heat-treated by tempering for 1.5 h between 300 °C and 400 °C for low-energy specimens, and for 1.25 h at 593 °C for high-energy specimens.

Additional details on specimen production, sampling and machining are available in [3].

The Charpy machines used in this study are the three master machines located at NIST in Boulder, CO. Their principal characteristics are listed in Table 2.

*Table 2 – Characteristics of NIST master Charpy machines.*

Machine ID	Hammer weight (N)	Hammer length (mm)	Fall angle (°)	Capacity (J)	Impact speed (m/s)	Hammer type
SI	296.6	800.4	136.3	409.05	5.20	U
TK	295.3	899.1	110.7	359.52	4.89	C
TO	267.7	900.7	119.2	358.63	5.12	U

For this study, we tested specimens from nine lots of verification specimens: five at the low-energy level (LL) and four at the high-energy level (HH). Of these nine lots, one (HH-149) was a “failed” lot, *i.e.*, rejected for use as verification specimens at –40 °C, based on a sample size greater than 5<sup>1</sup>.

Two of the lots were tested on all three master machines (typically, 25 tests per machine); each of the remaining seven were tested on one machine only (again, typically 25 tests). All tests were performed at room temperature (21 °C ± 1 °C).

The complete test matrix is presented in Table 3.

<sup>1</sup> For the explanation of sample size, refer to the following section on Statistical Analyses.



Table 3 – Test matrix for the feasibility study on RT SRMs.  
(LL = low-energy lot; HH = high-energy lot).

Specimen lot	Number of tests on			Total tests
	SI	TK	TO	
LL-119		25		25
LL-133		30		30
LL-138			25	25
LL-139	25			25
LL-140	25	25	25	75
HH-136		25		25
HH-140			25	25
HH-143	24			24
HH-149 <sup>2</sup>	25	25	25	75

### 3. Statistical Analyses

The statistics listed below are returned when Charpy test results from a single machine or multiple machines are analyzed by means of the NIST statistical analysis software *summary\_stats.r*, developed by Jolene Splett (jolene.splett@nist.gov) in the freely available programming language *R*.

#### 3.1 Test Result Statistics

- Number of tests performed ( $N$ ).
- Mean value of absorbed energy ( $\overline{KV}$ ):

$$\overline{KV} = \frac{1}{N} \sum_{i=1}^N KV_i \quad (1)$$

where  $KV_i$  is the value of absorbed energy (in J) obtained from the  $i$ -th test, with  $i = 1, \dots, N$ .

- Standard deviation ( $\sigma_{KV}$ ):

$$\sigma_{KV} = \sqrt{\frac{1}{N} \sum_{i=1}^N (KV_i - \overline{KV})^2} \quad (2)$$

- Variance ( $\sigma_{KV}^2$ ):

$$\sigma_{KV}^2 = \frac{1}{N} \sum_{i=1}^N (KV_i - \overline{KV})^2 \quad (3)$$

- Degrees of Freedom ( $\nu$ ):

$$\nu = N - 1 \quad (4)$$

---

<sup>2</sup> “Failed” lot.

- Standard Error of the mean ( $SE_{KV}$ ):

$$SE_{KV} = \frac{\sigma_{KV}}{\sqrt{N}} \quad (5)$$

- Smallest ( $KV_{min}$ ) and largest ( $KV_{max}$ ) value of absorbed energy.
- Range of absorbed energy values ( $KV_{max} - KV_{min}$ ).
- Coefficient of variation ( $CV$ ):

$$CV = \frac{\sigma_{KV}}{KV} \quad (6)$$

### 3.2 Machine Statistics

The same statistics listed under Section 3.1 are individually outputted for each of the impact machines used.

### 3.3 Additional Statistics

- Equality of variances: the hypothesis that the machine variances are equal is verified by means of Levene's test [8]. The output of the test is a *p-value*. If this is lower than the significance value  $\alpha = 0.05$ , the assumption of equal variances is rejected and the observed differences in sample variances are unlikely to have occurred based on random sampling from a population with equal variances.
- Pooled standard deviation: in statistics, pooled variance is a method for estimating the variance of several different populations when the mean of each population may be different, but one may assume that the variance of each population is the same [9]. The square root of a pooled variance is known as a pooled standard deviation ( $s_p$ ). It accounts for possibly different sample sizes for each machine.
- ASTM Pass/Fail: firstly, the deviation between the mean of each machine and the grand mean (mean of the means for each machine) is calculated. If the deviation is less than 1.4 J or 5 % of the grand mean (whichever is larger), the machine passes the ASTM E23 criterion. Additionally, the *k-ratio* is calculated for each machine, by dividing the machine's standard deviation by the pooled standard deviation. The *k-ratio* should be less than 1.25, based on 3 machines and 25 measurements per machine [3,10]. If any of the *k-ratio* values is greater than 1.25, the variability in energy values attributable to that machine is questionable and appropriate actions should be taken (direct verification, repairs, testing of additional specimens, etc.).
- Sample size: this represents the minimum number of specimens from a given lot that should be tested in a verification test for the outcome to be statistically significant. It is a very important statistical metric for assessing the quality of a reference specimen lot. It is defined as:

$$n_{SS} = \left( \frac{3s_p}{E} \right)^2, \quad (7)$$

where  $E$  is 1.4 J or 5 % of the grand mean, whichever is greater. The sample size is one of the statistics used to determine the acceptability of a lot<sup>3</sup> and the performance of the machines.

(e) Maximum  $s_p$ : for low-energy specimens, it is given by:

$$\text{Max } s_p = 1.4 \frac{\sqrt{5}}{3} = 1.043 \text{ J} \quad ; \quad (8)$$

for high-energy specimens, it is given by:

$$\text{Max } s_p = 0.037 \cdot \overline{KV_{gm}}, \quad (9)$$

where  $\overline{KV_{gm}}$  is the grand mean of the test results.

Obviously, the statistics listed in (a-c) above are meaningful only when tests are performed on more than one machine.

### 3.4 Criteria for Assessing the Feasibility of RT Verification Specimens

In this study, two statistical parameters will be primarily used to characterize the variability (scatter) of Charpy results, and hence to assess the feasibility of producing NIST verification specimens to be tested at room temperature:

- the coefficient of variation  $CV$ , eq. 6, and
- the sample size  $n_{SS}$ , eq. 7.

If both  $CV$  and  $n_{SS}$  calculated from room temperature tests are lower than or equivalent<sup>4</sup> to the values obtained at  $-40$  °C under the same experimental conditions (same machine(s) and approximately the same number of tests), the feasibility is demonstrated for a particular specimen lot.

---

<sup>3</sup> The NIST verification program routinely rejects lots with  $n_{SS} > 5$  [3].

<sup>4</sup> In this study, we arbitrarily assumed that RT values ( $CV$ ,  $n_{SS}$ ) can be considered equivalent to  $-40$  °C values if less than 20 % greater (based on engineering judgement).

## 4. Results

### 4.1 Lots Tested on the SI Machine

#### 4.1.1 LL-139

Twenty-five Charpy specimens from lot LL-139 were tested at 21 °C on the SI machine. The resulting statistics are given in Table 4. Detailed test results are provided in Appendix 1.

*Table 4 - Statistics resulting from LL-139 specimens tested at RT on the SI machine.*

Statistic	Value
$N$	25
$\overline{KV}$ (J)	17.906
$\sigma_{KV}$ (J)	0.559
$\sigma^2_{KV}$ (J <sup>2</sup> )	0.313
$SE_{\overline{KV}}$	0.112
$KV_{min}$ (J)	16.873
$KV_{max}$ (J)	18.952
Range (J)	2.079
$CV$	0.031
$n_{SS}$	1.437

The comparison between the results obtained at –40 °C (pilot lot and production lot)<sup>5</sup> and RT is provided in Table 5.

*Table 5 – Results for LL-139 tested at –40 °C and RT on the SI machine.*

Statistic	Test temperature (°C)	Value
$\overline{KV}$ (J)	–40 (pilot lot)	16.422
	–40 (production lot)	16.306
	21	17.906
$\sigma_{KV}$ (J)	–40 (pilot lot)	0.187
	–40 (production lot)	0.667
	21	0.559
Range (J)	–40 (pilot lot)	4.053
	–40 (production lot)	2.486
	21	2.079
$CV$	–40 (pilot lot)	0.057
	–40 (production lot)	0.041
	21	0.031
$n_{SS}$	–40 (pilot lot)	4.004
	–40 (production lot)	2.043
	21	1.437

<sup>5</sup> The first batch of 100 specimens that a supplier ships to NIST for preliminary qualification is called “pilot” lot. If the results of the pilot lot are acceptable (*i.e.*,  $n_{SS} \leq 5.0$ ), another batch of 100 specimens (“production” lot) is shipped for the final qualification of the Charpy verification specimen lot and the establishment of the certified reference value.

Both the coefficient of variation and the sample size at 21 °C are lower than those at -40 °C.

#### 4.1.2 HH-143

Twenty-four Charpy specimens from lot HH-143 were tested at 21 °C on the SI machine. The resulting statistics are given in Table 6. Detailed test results are provided in Appendix 2.

*Table 6 - Statistics resulting from HH-143 specimens tested at RT on the SI machine.*

Statistic	Value
$N$	24
$\overline{KV} \text{ (J)}$	106.799
$\sigma_{KV} \text{ (J)}$	3.165
$\sigma^2_{KV} \text{ (J}^2\text{)}$	10.017
$SE_{\overline{KV}}$	0.646
$KV_{min} \text{ (J)}$	101.800
$KV_{max} \text{ (J)}$	114.230
<b>Range (J)</b>	12.430
$CV$	0.030
$n_{SS}$	3.162

The comparison between the results obtained at -40 °C (pilot lot and production lot) and RT is provided in Table 7.

*Table 7 – Results for HH-143 tested at -40 °C and RT on the SI machine.*

Statistic	Test temperature (°C)	Value
$\overline{KV} \text{ (J)}$	-40 (pilot lot)	97.794
	-40 (production lot)	99.554
	21	106.799
$\sigma_{KV} \text{ (J)}$	-40 (pilot lot)	2.567
	-40 (production lot)	3.283
	21	3.165
<b>Range (J)</b>	-40 (pilot lot)	9.065
	-40 (production lot)	12.497
	21	12.430
$CV$	-40 (pilot lot)	0.026
	-40 (production lot)	0.033
	21	0.030
$n_{SS}$	-40 (pilot lot)	2.480
	-40 (production lot)	3.916
	21	3.162

The coefficient of variation and the sample size at room temperature are higher than those obtained for the pilot lot at -40 °C, but lower than those obtained for the production lot at -40 °C.

## 4.2 Lots Tested on the TO Machine

### 4.2.1 LL-138

Twenty-five Charpy specimens from lot LL-138 were tested at 21 °C on the TO machine. The resulting statistics are given in Table 8. Detailed test results are provided in Appendix 3.

*Table 8 - Statistics resulting from LL-138 specimens tested at RT on the TO machine.*

Statistic	Value
$N$	25
$\overline{KV}$ (J)	18.753
$\sigma_{KV}$ (J)	0.406
$\sigma^2_{KV}$ (J <sup>2</sup> )	0.164
$SE_{\overline{KV}}$	0.081
$KV_{min}$ (J)	17.721
$KV_{max}$ (J)	19.627
Range (J)	1.906
$CV$	0.022
$n_{ss}$	0.753

The comparison between the results obtained at -40 °C (pilot lot and production lot) and RT is provided in Table 9.

*Table 9 – Results for LL-138 tested at -40 °C and RT on the TO machine.*

Statistic	Test temperature (°C)	Value
$\overline{KV}$ (J)	-40 (pilot lot)	16.285
	-40 (production lot)	16.592
	21	18.753
$\sigma_{KV}$ (J)	-40 (pilot lot)	0.761
	-40 (production lot)	0.468
	21	0.406
Range (J)	-40 (pilot lot)	3.441
	-40 (production lot)	1.639
	21	1.906
$CV$	-40 (pilot lot)	0.047
	-40 (production lot)	0.028
	21	0.022
$n_{ss}$	-40 (pilot lot)	2.659
	-40 (production lot)	1.006
	21	0.753

Both the coefficient of variation and the sample size at 21 °C are lower than at -40 °C.

#### 4.2.2 HH-140

Twenty-five Charpy specimens from lot HH-140 were tested at 21 °C on the TO machine. The resulting statistics are given in Table 10. Detailed test results are provided in Appendix 4.

*Table 10 - Statistics resulting from HH-140 specimens tested at RT on the TO machine.*

Statistic	Value
$N$	25
$\overline{KV}$ (J)	103.307
$\sigma_{KV}$ (J)	2.934
$\sigma^2_{KV}$ (J <sup>2</sup> )	8.611
$SE_{\overline{KV}}$	0.587
$KV_{min}$ (J)	98.235
$KV_{max}$ (J)	109.020
Range (J)	10.785
$CV$	0.028
$nss$	2.905

The comparison between the results obtained at -40 °C (pilot lot and production lot) and RT is provided in Table 11.

*Table 11 – Results for HH-140 tested at -40 °C and RT on the TO machine.*

Statistic	Test temperature (°C)	Value
$\overline{KV}$ (J)	-40 (pilot lot)	94.335
	-40 (production lot)	97.498
	21	103.307
$\sigma_{KV}$ (J)	-40 (pilot lot)	2.485
	-40 (production lot)	3.506
	21	2.934
Range (J)	-40 (pilot lot)	9.149
	-40 (production lot)	12.294
	21	10.785
$CV$	-40 (pilot lot)	0.026
	-40 (production lot)	0.036
	21	0.028
$nss$	-40 (pilot lot)	2.498
	-40 (production lot)	4.656
	21	2.905

The coefficient of variation and the sample size at room temperature are equivalent (within 20 %) to those obtained for the pilot lot at -40 °C, and lower than those obtained for the production lot at -40 °C.

### 4.3 Lots Tested on the TK Machine

#### 4.3.1 LL-119

Twenty-five Charpy specimens from lot LL-119 were tested at 21 °C on the TK machine. The resulting statistics are given in Table 12. Detailed test results are provided in Appendix 5.

*Table 12 - Statistics resulting from LL-119 specimens tested at RT on the TK machine.*

Statistic	Value
$N$	25
$\overline{KV}$ (J)	17.104
$\sigma_{KV}$ (J)	0.604
$\sigma^2_{KV}$ (J <sup>2</sup> )	0.364
$SE_{\overline{KV}}$	0.121
$KV_{min}$ (J)	15.565
$KV_{max}$ (J)	18.272
Range (J)	2.707
CV	0.035
$n_{SS}$	1.671

The comparison between the results obtained at -40 °C (pilot lot and production lot) and RT is provided in Table 13.

*Table 13 – Results for LL-119 tested at -40 °C and RT on the TK machine.*

Statistic	Test temperature (°C)	Value
$\overline{KV}$ (J)	-40 (pilot lot)	14.241
	-40 (production lot)	14.053
	21	17.104
$\sigma_{KV}$ (J)	-40 (pilot lot)	0.802
	-40 (production lot)	0.678
	21	0.604
Range (J)	-40 (pilot lot)	3.097
	-40 (production lot)	2.798
	21	2.707
CV	-40 (pilot lot)	0.056
	-40 (production lot)	0.048
	21	0.035
$n_{SS}$	-40 (pilot lot)	2.953
	-40 (production lot)	2.112
	21	1.671

For LL-119, both the coefficient of variation and the sample size for LL-119 at 21 °C are lower than those at -40 °C.



#### 4.3.2 LL-133

Thirty<sup>6</sup> Charpy specimens from lot LL-133 were tested at 21 °C on the TK machine. The resulting statistics are given in Table 14. Detailed test results are provided in Appendix 6.

*Table 14 - Statistics resulting from LL-133 specimens tested at RT on the TK machine.*

Statistic	Value
<i>N</i>	30
$\overline{KV}$ (J)	17.057
$\sigma_{KV}$ (J)	1.687
$\sigma^2_{KV}$ (J <sup>2</sup> )	2.848
$SE_{\overline{KV}}$	0.308
$KV_{min}$ (J)	15.666
$KV_{max}$ (J)	23.207
Range (J)	7.541
<i>CV</i>	0.099
<i>n<sub>SS</sub></i>	13.078

The comparison between the results obtained at -40 °C (pilot lot and production lot) and RT is provided in Table 15.

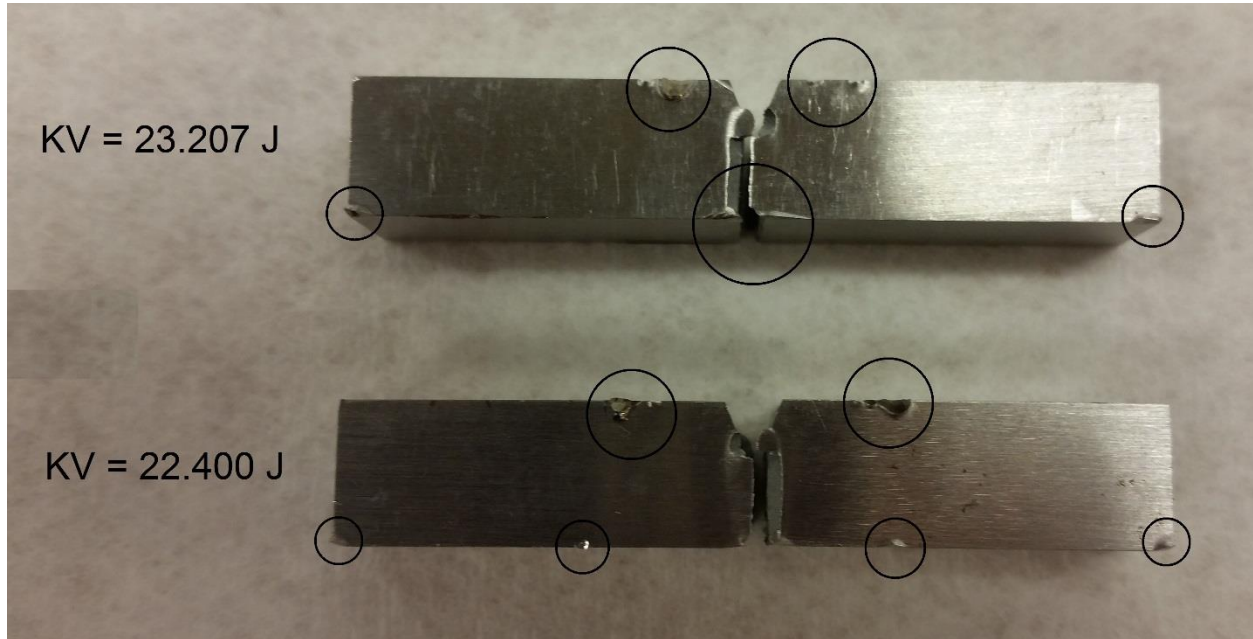
*Table 15 – Results for LL-133 tested at -40 °C and RT on the TK machine.*

Statistic	Test temperature (°C)	Value
$\overline{KV}$ (J)	-40 (pilot lot)	14.054
	-40 (production lot)	13.914
	21	17.057
$\sigma_{KV}$ (J)	-40 (pilot lot)	0.874
	-40 (production lot)	0.803
	21	1.687
Range (J)	-40 (pilot lot)	3.698
	-40 (production lot)	3.097
	21	7.541
<i>CV</i>	-40 (pilot lot)	0.062
	-40 (production lot)	0.058
	21	0.099
<i>n<sub>SS</sub></i>	-40 (pilot lot)	3.508
	-40 (production lot)	2.962
	21	13.078

For LL-133, the results obtained at RT were unsatisfactory, with very high values of *CV* and *n<sub>SS</sub>*. The cause was identified in two outlier tests, which yielded comparatively high values of absorbed energy: 23.207 J and 22.4 J (compared to a general mean of 17.057 J).

<sup>6</sup> The number of tests, higher than usual (25), was justified by the occurrence of jamming for two of the first 25 specimens tested.

On both specimens, clear indications of jamming were visible (see Figure 1, where the signs of jamming are circled in the figure).



*Figure 1 - LL-133 outlier specimens, showing clear signs of jamming (circled).*

Jamming between specimen and machine parts (anvils, supports, shrouds, or striker) is an obvious reason for rejecting the test result. If both outlier  $KV$  values are excluded from the analyses, the statistics for lot LL-133 tested at 21 °C become perfectly acceptable (Table 16), and both  $CV$  and  $n_{SS}$  are lower than those corresponding to the pilot and production lots tested at -40 °C (Table 17).

More details about the outlier identification procedure and a follow-up discussion on the possibility of jamming at room temperature are given in the Discussion section below.

*Table 16 - Statistics resulting from LL-133 specimens tested at RT on the TK machine, after the exclusion of the two jammed specimens.*

<b>Statistic</b>	<b>Value</b>
$N$	28
$\overline{KV}$ (J)	16.647
$\sigma_{KV}$ (J)	0.652
$\sigma^2_{KV}$ (J <sup>2</sup> )	0.426
$SE_{\overline{KV}}$	0.123
$KV_{min}$ (J)	15.666
$KV_{max}$ (J)	18.073
<b>Range (J)</b>	2.407
$CV$	0.039
$n_{SS}$	1.956

Table 17 – Comparison between the results for LL-133 tested at –40 °C and RT on the TK machine, after exclusion of the two jammed specimens.

Statistic	Test temperature (°C)	Value
$\overline{KV}$ (J)	–40 (pilot lot)	14.054
	–40 (production lot)	13.914
	21	16.647
$\sigma_{KV}$ (J)	–40 (pilot lot)	0.874
	–40 (production lot)	0.803
	21	0.652
Range (J)	–40 (pilot lot)	3.698
	–40 (production lot)	3.097
	21	2.407
CV	–40 (pilot lot)	0.062
	–40 (production lot)	0.058
	21	0.039
nss	–40 (pilot lot)	3.508
	–40 (production lot)	2.962
	21	1.956

#### 4.3.3 HH-136

Twenty-five Charpy specimens from lot HH-136 were tested at 21 °C on the TK machine. The resulting statistics are given in Table 18. Detailed test results are provided in Appendix 7.

Table 18 - Statistics resulting from HH-136 specimens tested at RT on the TK machine.

Statistic	Value
N	25
$\overline{KV}$ (J)	84.401
$\sigma_{KV}$ (J)	1.994
$\sigma^2_{KV}$ (J <sup>2</sup> )	3.975
$SE_{\overline{KV}}$	0.399
$KV_{min}$ (J)	81.523
$KV_{max}$ (J)	88.302
Range (J)	6.779
CV	0.024
nss	2.009

The comparison between the results obtained at –40 °C (pilot lot and production lot) and RT is provided in Table 19.

Table 19 – Results for HH-136 tested at –40 °C and RT on the TK machine.

Statistic	Test temperature (°C)	Value
$\overline{KV}$ (J)	–40 (pilot lot)	80.047
	–40 (production lot)	77.176
	21	84.401
$\sigma_{KV}$ (J)	–40 (pilot lot)	2.415
	–40 (production lot)	1.587
	21	1.994
Range (J)	–40 (pilot lot)	8.860
	–40 (production lot)	5.523
	21	6.779
CV	–40 (pilot lot)	0.030
	–40 (production lot)	0.021
	21	0.024
nss	–40 (pilot lot)	3.277
	–40 (production lot)	1.521
	21	2.009

For HH-136, the coefficient of variation and the sample size at room temperature are lower than those for the pilot lot at –40 °C, and higher than those for the production lot at –40 °C.

#### 4.4 Lots Tested on the Three Master Machines (SI, TK, TO)

##### 4.4.1 LL-140

Seventy-five Charpy specimens from lot LL-140 were tested at 21 °C, 25 on each of the three master machines (SI, TK, TO). The resulting statistics are given in Table 20 (all machines and individual machines). Detailed test results are provided in Appendix 8.

Table 20 - Statistics resulting from LL-140 specimens tested at RT on the three master machines.

Statistic	Impact machines			
	All	SI	TK	TO
$N$	75	25	25	25
$\overline{KV}$ (J)	18.917	19.242	17.880	19.629
$\sigma_{KV}$ (J)	0.938	0.649	0.603	0.413
$\sigma^2_{KV}$ (J <sup>2</sup> )	0.880	0.421	0.363	0.171
$SE_{\overline{KV}}$	0.108	0.130	0.121	0.083
$KV_{min}$ (J)	16.380	17.870	16.380	19.112
$KV_{max}$ (J)	20.936	20.240	19.190	20.936
Range (J)	4.556	2.370	2.810	1.824
CV	0.050	0.034	0.034	0.021
nss	1.462	1.933	1.667	0.785

The variances of the three machines were found to be equal, based on Levene’s test, with a calculated  $p$ -value of 0.072. The pooled standard deviation is  $s_p = 0.564$  J, which is lower than the maximum allowable value (1.043 J).

The ASTM E23 pass/fail criterion and the corresponding  $k$ -ratio of the three master machines, based on LL-140 test results, are documented in Table 21. The results are satisfactory for all machines (deviations  $\leq 1.4$  J;  $k$ -ratio  $\leq 1.25$ ). For more details, see section 3.3, item (c).

*Table 21 - Results of additional machine statistics based on LL-140 tested at 21 °C.*

Machine	Deviation (J)	ASTM E23 criterion	$k$ -ratio
SI	0.325	PASS	1.150
TK	-1.037	PASS	1.068
TO	0.712	PASS	0.732

The comparison between LL-140 Charpy results obtained at RT and  $-40$  °C (pilot and production lot) is documented in Table 22 for all machines.

*Table 22 – Results for LL-140 tested at  $-40$  °C and RT on the three master machines.*

Statistic	Test temperature (°C)	Value
$\overline{KV}$ (J)	$-40$ (pilot lot)	15.988
	$-40$ (production lot)	15.966
	21	18.917
$\sigma_{KV}$ (J)	$-40$ (pilot lot)	1.284
	$-40$ (production lot)	1.238
	21	0.938
Range (J)	$-40$ (pilot lot)	5.199
	$-40$ (production lot)	5.047
	21	4.556
CV	$-40$ (pilot lot)	0.080
	$-40$ (production lot)	0.078
	21	0.050
$n_{SS}$	$-40$ (pilot lot)	3.432
	$-40$ (production lot)	2.452
	21	1.462

Both the coefficient of variation and the sample size for LL-140 are at 21 °C lower than those at  $-40$  °C.

Since the LL-140 tests at room temperature were performed on all three master machines, we also established the certified reference value for absorbed energy at  $21$  °C  $\pm 1$  °C, as well as the expanded uncertainty, based on the standard procedures adopted by the NIST Charpy Verification Program and documented in [10,11]. The calculations were performed with *summary\_uncertainty.r*, software developed by Jolene Splett (jolene.splett@nist.gov) in R.

The certified reference value is defined as the grand average of the 75 specimens tested:  $KV_{ref} = 18.917$  J.

The combined standard uncertainty ( $u_c$ ) of the reference value is obtained by combining the within-machine standard uncertainty, the standard uncertainty due to machine bias, and the standard uncertainty of specimen homogeneity. For the RT tests on lot LL-140, we obtained:

$$u_c = 0.065 \text{ J} ,$$

which corresponds to 65 effective degrees of freedom (calculated with the Welch-Satterthwaite formula [12]).

The expanded uncertainty ( $U_c$ ), which corresponds to a 95 % confidence interval on the true reference value, is obtained by multiplying  $u_c$  by a coverage factor  $k$  which depends on the effective degrees of freedom. For LL-140 at room temperature,  $k = 1.9971$  and:

$$U_c = 0.13 \text{ J} ;$$

therefore, the lower and upper 95 % confidence limits on the reference value are 18.787 J and 19.047 J, respectively.

#### 4.4.2 HH-149

Seventy-five Charpy specimens from lot HH-149 were tested at 21 °C, 25 on each of the three master machines (SI, TK, TO). The resulting statistics are given in Table 23 (all machines and individual machines). Detailed test results are provided in Appendix 9.

HH-149 is a “failed” lot, for which a sample size of 6.711 was determined when certified at -40 °C.

*Table 23 - Statistics resulting from HH-149 specimens tested at RT on the three master machines.*

Statistic	Impact machines			
	All	SI	TK	TO
$N$	75	25	25	25
$\overline{KV}$ (J)	139.286	140.824	137.225	139.809
$\sigma_{KV}$ (J)	6.848	6.597	6.724	6.978
$\sigma^2_{KV}$ (J <sup>2</sup> )	46.892	43.516	45.208	48.687
$SE_{\overline{KV}}$	0.791	1.319	1.345	1.396
$KV_{min}$ (J)	125.63	132.25	125.63	125.79
$KV_{max}$ (J)	154.22	154.22	147.30	153.03
Range (J)	28.590	21.97	21.67	27.24
$CV$	0.049	0.047	0.049	0.050
$n_{SS}$	8.499	7.900	8.643	8.967

The variances of the three machines were found to be equal, based on Levene’s test, with a calculated  $p$ -value of 0.919. The pooled standard deviation is  $s_p = 6.768$  J, which is higher than the maximum allowable value (5.154 J, calculated by means of eq. 9).

The ASTM E23 pass/fail criterion and the corresponding *k-ratio* of the three master machines, based on the HH-149 test results, are documented in Table 24. The results are satisfactory for all machines (deviations  $\leq 5\%$  of grand mean, or 6.96 J; *k-ratio*  $\leq 1.25$ ). For more details, see section 3.3, item (c).

*Table 24 - Results of additional machine statistics based on HH-149 tested at 21 °C.*

Machine	Deviation (J)	ASTM E23 criterion	<i>k-ratio</i>
SI	-0.839	PASS	1.025
TK	2.366	PASS	0.927
TO	-1.527	PASS	1.045

The comparison between HH-149 Charpy results obtained at RT and  $-40\text{ °C}$  (pilot and production lot) is documented in Table 25 for all machines.

*Table 25 – Results for HH-149 tested at  $-40\text{ °C}$  and RT on the three master machines.*

Statistic	Test temperature (°C)	Value
$\overline{KV}$ (J)	$-40$ (pilot lot)	123.196
	$-40$ (production lot)	124.303
	21	139.286
$\sigma_{KV}$ (J)	$-40$ (pilot lot)	5.686
	$-40$ (production lot)	5.562
	21	6.848
Range (J)	$-40$ (pilot lot)	27.680
	$-40$ (production lot)	25.660
	21	28.590
CV	$-40$ (pilot lot)	0.046
	$-40$ (production lot)	0.045
	21	0.049
$n_{SS}$	$-40$ (pilot lot)	7.369
	$-40$ (production lot)	6.711
	21	8.499

As can be seen from Table 24 and Table 25, this high-energy lot also failed when tested at room temperature ( $n_{SS} > 5$ ).

If HH-149 specimens were to be sold for the verification of the Charpy machines at room temperature, the minimum number of specimens in a verification set would be 9.

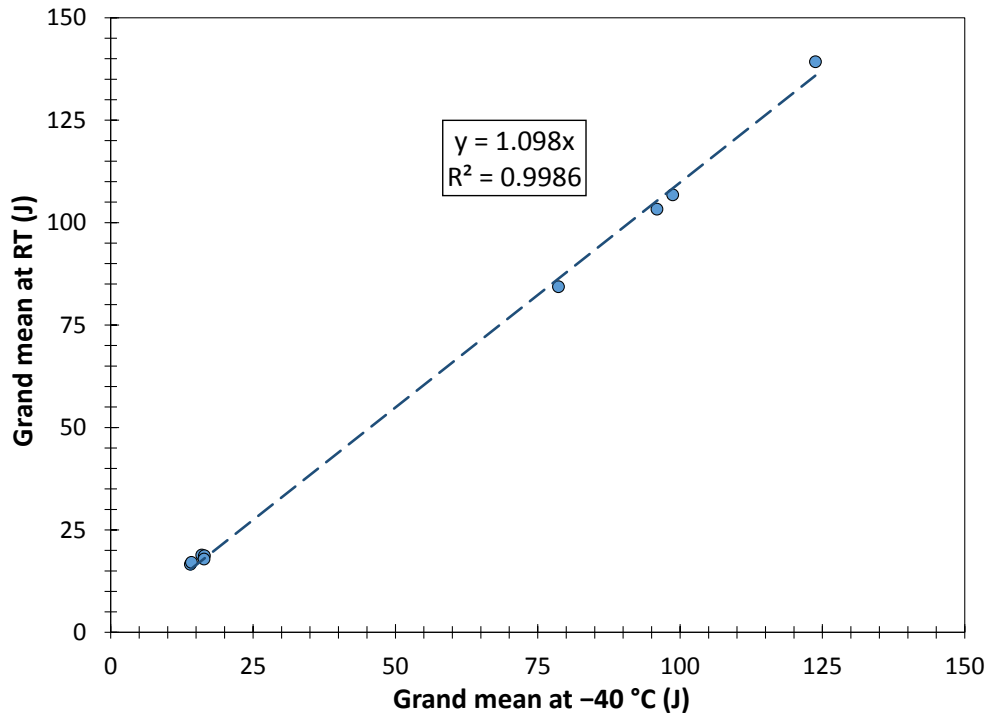
The certified reference value, corresponding to the grand average of the 75 specimens tested, is  $KV_{ref} = 139.29$  J.

The combined standard uncertainty is  $u_c = 0.782$  J, with 72 effective degrees of freedom. The expanded uncertainty, with a coverage factor  $k = 1.9935$ , is  $U_c = 1.99$  J. The lower and upper 95 % confidence limits on the reference value are 137.728 J and 140.844 J respectively.

## 5. Discussion

### 5.1 Relationships between Results at $-40\text{ }^{\circ}\text{C}$ and $21\text{ }^{\circ}\text{C}$

The grand means obtained at  $21\text{ }^{\circ}\text{C}$  and  $-40\text{ }^{\circ}\text{C}$  (average of pilot and production lot tests) are compared in Figure 2. Based on a linear fit, the absorbed energy values at room temperature are  $10\% \pm 4\%$  (95% confidence) higher than at  $-40\text{ }^{\circ}\text{C}$ .



*Figure 2 - Relationship between absorbed energy (grand means) at room temperature and  $-40\text{ }^{\circ}\text{C}$ .*

A similar comparison is shown in Figure 3 for the sample size  $n_{SS}$ . In the figure, the upper left half corresponds to an increased variability at RT with respect to  $-40\text{ }^{\circ}\text{C}$ , the lower right half to a reduced variability. Sample sizes at  $-40\text{ }^{\circ}\text{C}$  were obtained by averaging the values calculated for the pilot lot and the production lot.

All lots examined in this study, with the exception of the “failed” lot HH-149, show lower variability (lower sample size) at  $21\text{ }^{\circ}\text{C}$  than at  $-40\text{ }^{\circ}\text{C}$ . It’s interesting to note that the scatter reduction is more significant for low-energy specimens than for high-energy specimens.

A summary of the major statistical metrics (coefficient of variation and sample size) is given in Table 26.



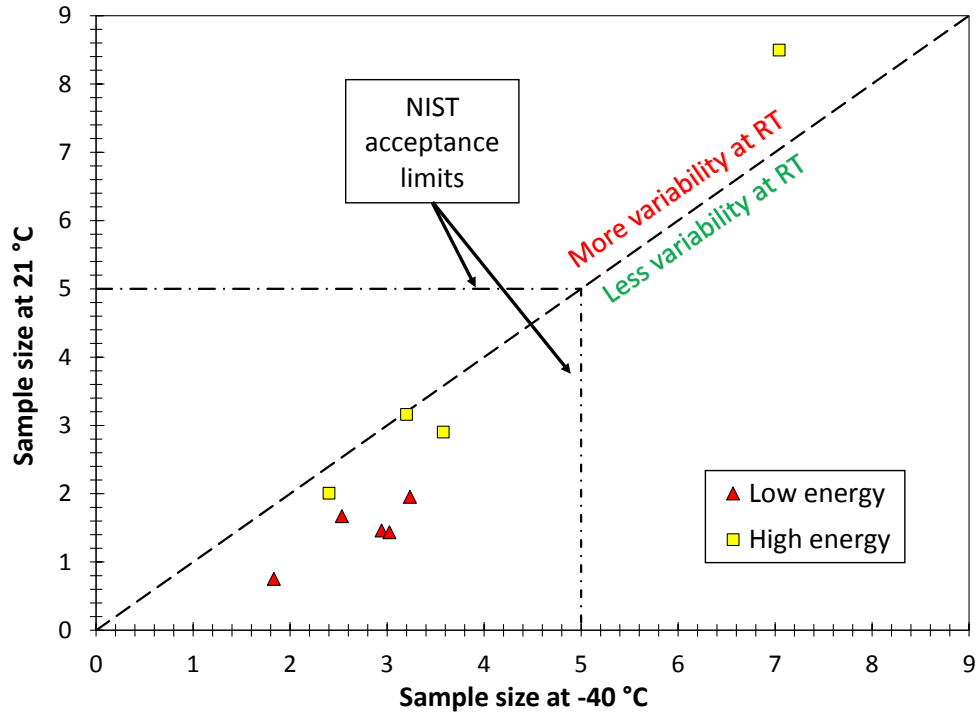


Figure 3 - Relationship between sample size at room temperature and  $-40\text{ }^{\circ}\text{C}$ .

Table 26 - Coefficients of variation and sample sizes obtained at  $-40\text{ }^{\circ}\text{C}$  and room temperature.

Energy level	Lot	Machines	CV		$n_{SS}$	
			$-40\text{ }^{\circ}\text{C}$	RT	$-40\text{ }^{\circ}\text{C}$	RT
Low	LL-140	All	0.079	0.050	2.94	1.46
	LL-133	TK	0.060	0.039	3.24	1.96
	LL-119 <sup>7</sup>	TK	0.052	0.035	2.53	1.67
	LL-138	TO	0.037	0.022	1.83	0.75
	LL-139	SI	0.049	0.031	3.02	1.44
High	HH-149	All	0.045	0.049	7.04	8.50
	HH-136	TK	0.025	0.024	2.40	2.01
	HH-140	TO	0.031	0.028	3.58	2.91
	HH-143	SI	0.030	0.030	3.20	3.16

## 5.2 Specimen Jamming at the Low-Energy Level

When low-energy verification specimens are tested at  $-40\text{ }^{\circ}\text{C}$ , in most cases the broken specimens exit the machine in a direction opposite to the pendulum swing. This minimizes the chances of post-test secondary interactions between specimen halves and the swinging pendulum, or other parts of the machine (anvils, supports, shrouds if present).

However, when testing low-energy specimens at room temperature, we noticed that oftentimes the broken specimens are not ejected from the machine, but remain close to the test area because of the slightly higher impact toughness (around 10 % according to our results, see Figure 2). As a consequence, secondary impacts with the swinging hammer become more frequent

<sup>7</sup> After excluding the two outliers.

and the likelihood of one or both specimen halves jamming and dissipating pendulum energy increases.

Out of 130 low-energy specimens tested at room temperature, only two specimens (1.5 %) showed clear evidence of jamming, as indicated by both their significantly high *KV* values and the marks visible on the broken halves (Figure 1). Both tests were performed on the same master machine (TK) and on the same low-energy lot (LL-133).

To confirm that the results of these tests are outliers and ought to be excluded, we used a common statistical test for outlier detection: Grubbs' test, also known as the maximum normed residual test or extreme studentized deviate test [13]. Both tests were identified as outliers:

- (a) The highest *KV* value (23.207 J, compared to an average of 16.845 J for the remaining 29 tests) corresponded to a *Z-value* of 3.64447, which was higher than the critical value of *Z* (2.90847) at a significance level of 0.05.
- (b) The second highest *KV* value (22.400 J, compared to an average of 16.647 J for the remaining 28 tests) corresponded to a *Z-value* of 4.45913, which was higher than the critical value of *Z* (2.89270) at a significance level of 0.05.

Grubbs' test performed on the remaining 28 test results did not detect any residual outliers. The coefficient of variation dropped from 0.099 to 0.074 (first outlier removed) to 0.039 (second outlier removed); the sample size decreases from 13.078 to 7.127 (first outlier removed) to 1.956 (second outlier removed).

The observation that both outliers were tested on the same machine suggests an effect of the machine design on the occurrence of specimen jamming. The TK is the only master machine that has a C-shape pendulum, which might be more susceptible to significant losses of energy due to specimen/machine interactions.

It is also interesting to note that jamming occurred only for LL-133, but not for the other two low-energy lots tested on the TK machine (LL-119 and LL-140). We therefore decided to compare the three low-energy lots in terms of full energy vs. temperature transition curves, obtained by performing tests between -180 °C and 300 °C. The comparison of the transition curves in Figure 4 shows that LL-133 is the toughest of the three low-energy lots, but the differences in absorbed energy are negligible both at -40 °C and 21 °C.

It is questionable, therefore, whether a modification of the heat treatment for the low-energy material, such as lowering the tempering temperature below 400 °C or modifying the duration of the heat treatment, could effectively decrease the likelihood of jamming. Furthermore, the trend of absorbed energy as a function of tempering temperature for 4340 shown in Figure 5 [3] indicates that *KV* is not very sensitive to tempering temperatures below 400 °C.

All things considered, the slightly higher likelihood of a low-energy specimen jamming at room temperature does not seem a serious hurdle for developing room temperature SRMs. Even for specimens tested at -40 °C, our current procedure calls for removing from the analyses any specimen showing evidence of jamming or other test-related issues (such as a specimen struck off-center or badly positioned, etc.). When a customer sends back a sample that has clearly jammed and whose absorbed energy is significantly higher than the rest of the verification set, its result will be ignored and the machine verification will be based on the *KV* values from the remaining specimens.

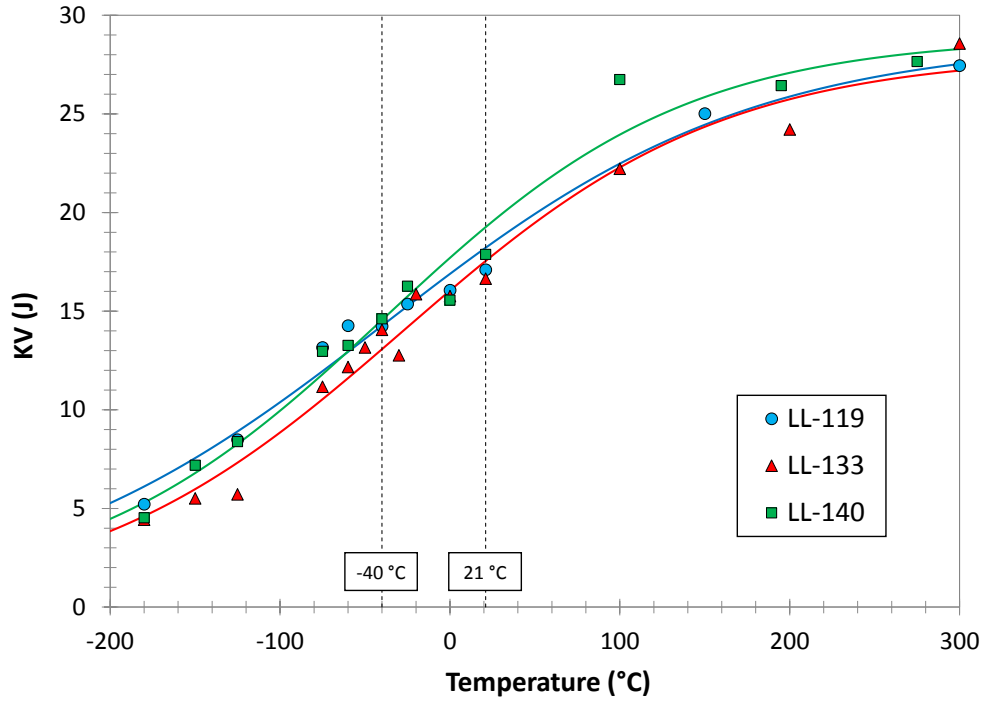


Figure 4 - KV transition curves for LL-119, LL-133, and LL-140.

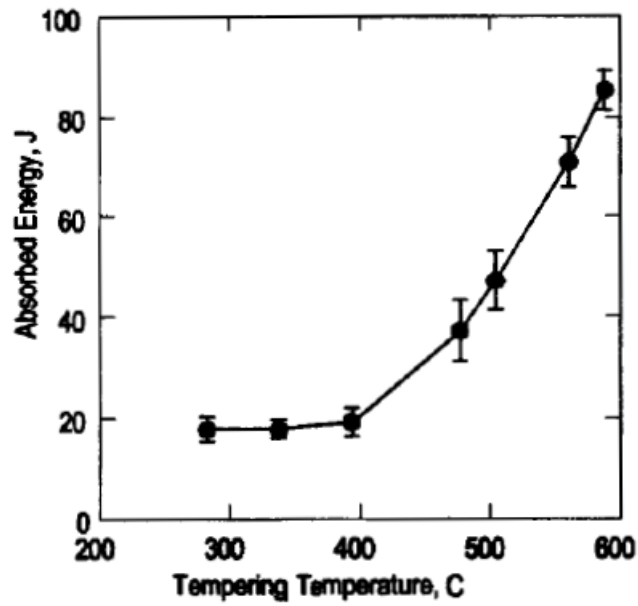


Figure 5 - Effect of tempering temperature on RT absorbed energy for 4340 steel [3].

## 6. Conclusions

This study has clearly demonstrated the feasibility (and the benefit) of certifying our low-energy and high-energy Charpy verification specimens at room temperature ( $21\text{ }^{\circ}\text{C} \pm 1\text{ }^{\circ}\text{C}$ ) instead of  $-40\text{ }^{\circ}\text{C}$ .

The room-temperature tests that we conducted on 5 low-energy lots and 4 high-energy lots, tested on the three master machines located in Boulder, indicated that the variability in absorbed energy values decreased in 8 out of 9 cases, as demonstrated by lower coefficients of variation and lower sample sizes. The only lot for which both statistical metrics were higher at room temperature than at  $-40\text{ }^{\circ}\text{C}$  was a “failed” high-energy lot, which had already proven inadequate (sample size  $> 5.0$ ) during the original certification of the pilot and production lots.

For one of the low-energy lots tested on the TK machine (the only machine with a C-shaped hammer), two specimens jammed and yielded unusually high absorbed energy values. Signs of jamming were clearly visible on the broken samples, and their values of absorbed energy were classified as statistical outliers according to Grubbs’ test. Although the likelihood of jamming at RT appears larger than at  $-40\text{ }^{\circ}\text{C}$ , given that most specimens tend to remain close to the anvil/support area instead of being ejected backward, it seems unlikely that this type of behavior could be changed by modifying the heat treatment of the low-energy 4340 steel. Jamming can be clearly recognized however, and the results from a jammed specimen can be easily removed from the evaluation of a set of verification specimens.

## References

- [1] ASTM E23-12c, “*Standard Test Methods for Notched Bar Impact Testing of Metallic Materials*,” ASTM Book of Standards, Vol. 03.01, ASTM International, West Conshohocken, PA, 2015.
- [2] D. E. Driscoll, “*Reproducibility of Charpy Impact Test*,” ASTM, Philadelphia, PA, 1955.
- [3] C. N. McCowan, T. A. Siewert, and D. P. Vigliotti, “*The NIST Charpy V-notch Verification Program: Overview and Operating Procedures*,” in **Charpy Verification Program: Reports Covering 1989-2002**, NIST Technical Note 1500-9, Materials Reliability Series, September 2003, pp. 3-42.
- [4] IRMM, Standards for Innovation and Sustainable Development Unit, “*Certified Reference Materials – 2015*,” available on line at [https://ec.europa.eu/jrc/sites/default/files/rm\\_catalogue.pdf](https://ec.europa.eu/jrc/sites/default/files/rm_catalogue.pdf) (retrieved 6/18/2015).
- [5] ISO 148:2-2008, “*Metallic materials -- Charpy pendulum impact test -- Part 2: Verification of testing machines*,” International Standards Organization, Geneva, Switzerland.
- [6] C. N. McCowan, J. Pauwels, G. Revise, and H. Nakano, “*International Comparison of Impact Verification Programs*,” in **Pendulum Impact Testing: A Century of Progress**, ASTM STP 1380, T. A. Siewert and M. P. Manahan, Eds., American Society for Testing and Materials, West Conshohocken, PA, 1999, pp. 210-219.
- [7] “*NMIJ CRM Catalog 2014-2015*,” available on line at [https://www.nmij.jp/english/service/C/CRM\\_Catalog\\_20140828.pdf](https://www.nmij.jp/english/service/C/CRM_Catalog_20140828.pdf) (retrieved 6/18/2015).
- [8] H. Levene, “*Robust tests for equality of variances*,” in Ingram Olkin, Harold Hotelling, *et alia*, Stanford University Press, 1960, pp. 278-292.
- [9] P. R. Killeen, “*An alternative to null-hypothesis significance tests*,” *Psychological Science*, Vol. 16, No. 5, May 2005, pp. 345-353.
- [10] C. M. Wang and J. D. Splett, “*Consensus Values and Reference Values Illustrated by the Charpy Machine Certification Program*,” *Journal of Testing and Evaluation*, JTEVA, Vol. 25, No. 3, May 1997, pp. 308-314.
- [11] J. D. Splett, C. N. McCowan, H. K. Iyer, and C.-M. Wang, “*NIST Recommended Practice Guide: Computing Uncertainty for Charpy Impact Machine Test Results*,” NIST Special Publication 960-18, Sep 2007.
- [12] F. E. Satterthwaite, “*An Approximate Distribution of Estimates of Variance Components*,” *Biometrics Bulletin* 2, 1946, pp. 110-114.
- [13] F. E. Grubbs, “*Sample criteria for testing outlying observations*,” *The Annals of Mathematical Statistics*, 21(1), Mar 1950, pp. 27-58.

# Appendix 1

Test results for LL-139 lot (SI machine)

KV (J)
17.874
17.444
17.874
17.731
18.520
17.731
17.946
18.880
17.016
17.158
17.659
17.516
18.952
17.587
17.731
18.017
18.089
18.952
17.731
17.946
18.736
16.873
17.802
17.659
18.233

# Appendix 2

Test results for HH-143 lot (SI machine)



KV (J)
109.89
109.43
105.11
102.43
108.17
104.75
106.46
102.87
105.38
109.43
105.47
102.96
104.84
106.64
109.52
101.80
109.71
105.56
104.84
105.38
109.43
114.23
106.55
112.33

# Appendix 3

Test results for LL-138 lot (TO machine)

KV (J)
18.759
18.673
18.933
18.586
19.019
19.627
18.586
18.586
19.366
19.019
18.413
18.500
18.153
17.721
18.759
18.759
19.019
18.586
18.586
18.586
19.193
18.846
19.193
19.106
18.240

# Appendix 4

Test results for HH-140 lot (TO machine)

KV (J)
102.87
104.00
102.77
102.21
102.87
99.75
98.52
109.02
102.11
104.19
101.07
103.62
106.37
102.77
102.96
105.52
102.02
100.50
101.54
100.50
105.80
107.32
98.24
107.98

# Appendix 5

Test results for LL-119 lot (TK machine)

KV (J)
17.068
16.967
18.272
17.068
17.469
16.667
17.268
17.469
17.971
17.469
17.268
17.068
17.068
17.770
16.867
16.967
16.967
18.272
15.565
17.068
17.168
16.366
16.466
16.366

# Appendix 6

Test results for LL-133 lot (TK machine)



KV (J)	Note
16.969	
15.666	
16.067	
15.967	
17.069	
18.073	
16.267	
16.167	
17.471	
17.772	
16.167	
15.867	
16.067	
16.468	
17.170	
16.668	
16.167	
16.568	
16.668	
<b>23.207</b>	<b>Outlier #1</b>
<b>22.400</b>	<b>Outlier #2</b>
17.872	
16.267	
17.370	
16.769	
16.065	
16.566	
16.667	
17.268	

# Appendix 7

Test results for HH-136 lot (TK machine)

KV (J)
84.860
84.652
87.051
81.523
82.774
84.339
88.302
83.191
87.572
82.879
83.191
83.504
83.713
82.149
87.468
84.130
84.547
86.320
83.087
82.253
82.461
85.173
85.903
87.155

# Appendix 8

Test results for LL-140 lot  
(all master machines)

Machine	KV (J)	Machine	KV (J)	Machine	KV (J)
SI	19.66	TO	19.63	TK	17.88
SI	18.30	TO	19.72	TK	17.78
SI	19.74	TO	20.24	TK	17.58
SI	18.58	TO	19.55	TK	18.08
SI	20.10	TO	19.11	TK	17.98
SI	18.87	TO	19.29	TK	17.68
SI	18.37	TO	19.29	TK	19.19
SI	19.09	TO	19.63	TK	18.18
SI	20.24	TO	20.94	TK	19.09
SI	18.73	TO	19.55	TK	17.98
SI	19.74	TO	19.29	TK	17.48
SI	20.03	TO	19.11	TK	17.08
SI	17.87	TO	19.37	TK	18.68
SI	19.38	TO	19.46	TK	17.78
SI	18.30	TO	19.37	TK	17.88
SI	19.30	TO	19.89	TK	17.76
SI	20.03	TO	19.81	TK	18.08
SI	19.59	TO	19.98	TK	17.78
SI	19.02	TO	19.55	TK	17.38
SI	19.66	TO	20.15	TK	17.78
SI	18.80	TO	19.11	TK	17.38
SI	18.94	TO	19.37	TK	18.78
SI	19.88	TO	19.89	TK	17.88
SI	19.23	TO	19.81	TK	16.38

# Appendix 9

Test results for HH-149 lot  
(all master machines)

Machine	KV (J)	Machine	KV (J)	Machine	KV (J)
SI	154.22	TO	143.92	TK	129.87
SI	146.59	TO	153.03	TK	140.35
SI	140.38	TO	149.09	TK	143.93
SI	136.86	TO	144.68	TK	143.53
SI	136.68	TO	151.53	TK	125.63
SI	133.82	TO	140.15	TK	128.11
SI	144.55	TO	135.62	TK	139.53
SI	146.04	TO	144.96	TK	130.18
SI	140.75	TO	132.98	TK	140.25
SI	142.33	TO	142.98	TK	141.58
SI	132.99	TO	147.12	TK	131.52
SI	135.11	TO	143.36	TK	137.89
SI	135.48	TO	135.72	TK	142.71
SI	143.25	TO	134.21	TK	143.73
SI	148.27	TO	131.47	TK	133.47
SI	147.24	TO	139.21	TK	134.19
SI	138.62	TO	141.29	TK	128.42
SI	136.03	TO	142.89	TK	145.47
SI	132.25	TO	145.71	TK	143.93
SI	153.01	TO	134.49	TK	147.30
SI	152.36	TO	132.88	TK	131.83
SI	136.31	TO	125.79	TK	126.87
SI	134.56	TO	130.43	TK	134.71
SI	136.12	TO	137.61	TK	140.56