Materials Performance and Characterization

doi:10.1520/MPC20240077 / Vol. 14 / No. 1 / 2025 / available online at go.astm.org/pubs

Enrico Lucon[1]

On the Possibility of Adding a Repeatability Requirement for the Indirect Verification of Charpy Machines in ASTM E23

Manuscript received October 3, 2024; accepted for publication January 6, 2025; published online March 27, 2025. Issue published March 27, 2025.

[1] Applied Chemicals and Materials Division, National Institute of Standards and Technology, 325 Broadway St., Boulder, CO 80305, USA (Corresponding author), e-mail: enrico.lucon@nist.gov, https://orcid.org/0000-0002-3021-4785

Reference

E. Lucon, “On the Possibility of Adding a Repeatability Requirement for the Indirect Verification of Charpy Machines in ASTM E23,” *Materials Performance and Characterization* 14, no. 1 (2025): 35–48. https://doi.org/10.1520/MPC20240077

ABSTRACT

Since 1989, the National Institute of Standards and Technology (NIST) Charpy Machine Verification Program in Boulder, Colorado, has been indirectly verifying Charpy machines around the world according to ASTM E23, *Standard Test Methods for Notched Bar Impact Testing of Metallic Materials*, and ISO 148-2, *Metallic Materials—Charpy Pendulum Impact Test—Part 2: Verification of Testing Machines*. Most of the indirect verifications results have been logged in a database, which at the end of May 2024 included 77,105 records, each one corresponding to 4–5 absorbed energy values at one of three energy levels (low, high, and super-high). Based on the analysis of this database, the acceptability criteria of ASTM E23, exclusively based on bias, have been compared with those of ISO 148-2, which also include repeatability. The objective of this study is to support the proposal of revising ASTM E23 by introducing repeatability requirements in addition to the existing bias criteria. Four proposals have been considered and assessed, leading to the selection of a set of repeatability limits (larger of 4 J and 15 % of certified absorbed energy) that only moderately increases the severity of the standard but, most importantly, prevents deeming machines acceptable that satisfy ASTM E23 bias but fail ISO 148-2 repeatability (approximately one sixth of all the machines verified by the NIST since 1990).

Keywords

ASTM E23, bias, indirect verification, ISO 148-2, National Institute of Standards and Technology Charpy database, National Institute of Standards and Technology Charpy Program, repeatability

Introduction

The Charpy impact test is one of the oldest[1,2] and still most popular mechanical tests. Numerous industry codes and construction specifications, as well as standards and procedures for safety analysis, prescribe Charpy testing as one of the tools for establishing the toughness of metallic materials or their resistance to the propagation of a pre-existing defect under applied external loads. The most important normative documents covering the performance and interpretation of Charpy tests are ASTM E23, *Standard Test Methods for Notched Bar Impact Testing of Metallic* Materials,[3] whose first edition dates to 1933, and ISO 148-1:2016, *Metallic Materials — Charpy Pendulum Impact Test — Part 1: Test Method*,[4] first issued in 1983 as simply ISO 148, *Steel — Charpy Impact Test (V-notch)*.*

The concept of periodically conducting Charpy tests on certified reference specimens to verify the performance and maintenance state of impact machines (what came to be known as “indirect verification”) was introduced in the 1950s, when several researchers started to question the reliability and soundness of the test itself due to the significant scatter in Charpy data generated by different laboratories on specimens extracted from the same materials. The Chief of the Mechanical Metallurgy Branch of the Watertown Arsenal in Massachusetts, David E. Driscoll, countered this opinion by demonstrating that much tighter agreement between testing laboratories could be achieved, provided Charpy machines were correctly maintained and periodically checked with specimens accurately machined to strict tolerances, using materials with very low scatter of mechanical properties.[5,6] His published research significantly contributed to the reevaluation of the Charpy test. Soon after, the US Army set up a program at the Watertown Arsenal aimed at producing and supplying reference specimens for the indirect verification of Charpy machines. The program, which started in the late 1960s, was eventually transferred to the National Bureau of Standards (which later became the National Institute for Standards and Technology [NIST]) in Boulder, Colorado, and is active to the present day.

Around the same time (late 1960s/early 1970s), ASTM E23 was updated to include the requirement to indirectly verify Charpy machines at least once a year by testing certified reference specimens, whose scatter needed to be kept to a minimum level to allow statistically meaningful comparisons between individual laboratories and the institution certifying the absorbed energy values of the reference specimens. The indirect verification requirements eventually evolved in different directions for ASTM E23 and ISO 148-2, *Metallic Materials—Charpy Pendulum Impact Test—Part 2: Verification of Testing Machines*† and presently correspond to considerably different degrees of severity:

- ASTM E23[3] prescribes that the difference between the average absorbed energy from five Charpy tests on certified reference specimens ($\overline{KV}$) and the certified value (KV_R) be within the larger of 1.4 J and 5 % of KV_R for a specific energy level. According to A2.4.1.2 of the standard, the certified value must have been established on the three reference machines located at NIST in Boulder, Colorado.
- ISO 148-2[7] requires that $|\overline{KV} - KV_R|$ be within the larger of 4 J and 10 % KV_R (bias), and the difference between maximum and minimum KV, KV_{max}–KV_{min} (repeatability), be within the larger of 6 J and 15 % KV_R. Historically, ISO 148-2 decided to adopt much looser requirements than ASTM E23 under pressure from several countries, which were concerned that the ASTM level of severity would have prevented the use of many of their older machines.

In terms of bias, ASTM E23 is at least twice stricter than ISO 148-2. However, ASTM E23 does not have a reproducibility criterion, and, therefore, no limit is set to the scatter of absorbed energy values obtained from a series of indirect verification tests. The obvious implication of this is that any machine failing ASTM E23 would automatically fail ISO 148-2, whereas the reverse is not true, because of the absence of a repeatability requirement in ASTM E23.

---

* ISO 148 (1983) did not include provisions for the verification, direct or indirect, of impact machines.

† It may be worthwhile to specify here that the ASTM and ISO requirements for indirect verification, although different, have not changed since they were first introduced in ASTM E23 and ISO 148 many years ago.

The objective of this study is to evaluate the possibility and advantages of adding a suitable repeatability criterion to a future edition of ASTM E23.

Analysis of the NIST Charpy Database

The Charpy reference specimens certified by the NIST correspond to three distinct energy levels: low energy ($KV_R \approx 17$ J to 20 J), high energy ($KV_R \approx 90$ J to 120 J), and super-high energy ($KV_R \approx 180$ J to 220 J). Furthermore, they can be classified as "self-verification" or "NIST-verification." When purchasing the former type of specimens, customers evaluate the conformity of their machine to ASTM E23, ISO 148-2, or both, based on their own test results, and there is no post-test service provided by NIST. Conversely, in the case of "NIST-verification" specimens, customers are required to send test results and pictures of the tested specimens to NIST, who evaluate the results, certify their compliance with ASTM E23, ISO 148-2, or both (depending on the customer's preference), and provide an official verification letter and a sticker to the testing laboratory. Should the indirect verification be unsuccessful, NIST also provides recommendations for troubleshooting the situation.

Since 1990 and more consistently starting in 1993, absorbed energy values, KV, reported by customers have been stored in the NIST Charpy database, which has become an invaluable tool for investigating general trends in worldwide Charpy testing, as well as studying the effects of various machine features on test results, such as the shape of the impact hammer.[8]

The NIST Charpy Machine Verification Program was originally tasked with supporting the American industry, but, through the years, more and more companies all over the world used NIST specimens to verify their machines, and the program became truly international. Consequently, in the last 10 years, NIST has started verifying machines not only according to ASTM E23, but also according to ISO 148-2. Nevertheless, every customer record in the database is currently also evaluated in accordance with the ISO 148-2 requirements for both bias and reproducibility.

The analyses performed within this study were conducted on the NIST database as of May 31, 2024. By that date, 77,105 records were included in the database, each corresponding to (nominally) five tests[‡] performed on reference specimens of a given energy level. Database records are split between low-energy tests (44.4 %), high-energy tests (44.1 %), and super-high–energy tests (11.4 %).

Although the oldest records in the database are from October 8, 1990, customer results only started to be regularly input in January 1993. Verified machines and companies served over the last 35 years are illustrated in **figure 1**.

The number of indirect verifications processed by NIST are listed in **Table 1**, subdivided by energy level and time period.

The whole NIST database cannot, for obvious reasons, be made public because it contains confidential information about companies and machines. However, the Microsoft Excel spreadsheets used for this study containing all indirect verification test results from 1990–2024, have been completely anonymized by removing company/machine information, and are available to anyone interested by contacting the corresponding author.

RESULTS OF INDIRECT VERIFICATIONS ACCORDING TO ASTM E23 AND ISO 148-2

The percentage of indirect verifications that failed the requirements of ASTM E23 (bias greater than the larger of 1.4 J and 5 % K_R) is shown in **figure 2**. The general trend at all energy levels is for the fail rate[§] to decrease over time, which indicates an overall improvement of the quality/maintenance state of Charpy machines around the world. It is interesting to note a spike for failed verifications at the super-high–energy level around 2017, two years before those reference specimens became available again, after being out of stock since 2010.

[‡] In most cases, all five test results are used for evaluation. Sporadically, one energy value must be removed from the analysis because of experimental issues (jamming, incorrect positioning on the machine anvils, etc.).

[§] In this paper, "fail rate" is defined as the ratio between the number of unsuccessful indirect verifications and the total number of indirect verifications, expressed in percent. "Pass rate" is defined similarly, using the number of successful verifications.

FIG. 1

Machines verified and companies served by the NIST Charpy Machine Verification Program between 1990 and 2024.

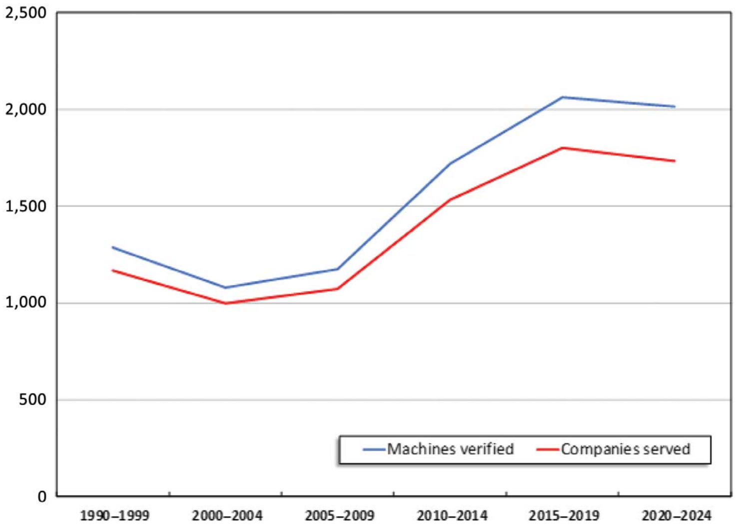


TABLE 1

Indirect verifications processed by NIST between 1990 and 2024

	Energy Level			
Period	Low	High	Super-High	Total
1990–1999	5,606	5,601	1,728	12,935
2000–2004	4,226	4,153	1,842	10,221
2005–2009	3,908	3,886	1,923	9,717
2010–2014	6,168	6,103	200	12,471
2015–2019	7,731	7,627	39	15,397
2020–2024	6,620	6,665	3,079	16,364

Note: Super-high energy reference specimens were out of stock between 2010 and 2019.

Historical ASTM E23 fail rates are between 6 % and 13 %, except for the super-high energy spike mentioned earlier. The fail rate for low-energy specimens is consistently higher than that for high-energy specimens, demonstrating that tests at the lowest energy level are the most challenging for the average Charpy machine. When the material behavior is markedly brittle, the vibrational modes of some machines can be significantly activated, with the consequence that vibrational losses can artificially increase the absorbed energy[9] and therefore cause the machine to fail at the low-energy level. Another significant contribution is given by the compliance of the machine striker, which depends on its design, and can in turn excite specific vibrational modes of the swinging hammer.[10]

The same information is shown by **figure 3** for failed verifications in accordance with ISO 148-2 (bias greater than the larger of 4 J and 10 % K_R; repeatability greater than the larger of 6 J and 15 % K_R).

Clear differences between ASTM E23 and ISO 148-2 can be appreciated by comparing **figures 2** and **3**.

- Overall, fail rates are significantly lower for ISO 148-2 (between 1 % and 7.5 %, excluding the super-high energy spike) than for ASTM E23.
- Contrary to ASTM E23, fail rates for low-energy specimens are much lower than at the high-energy or super-high–energy level.

FIG. 2

Percentage of failed ASTM E23 verifications between 1990 and 2024.

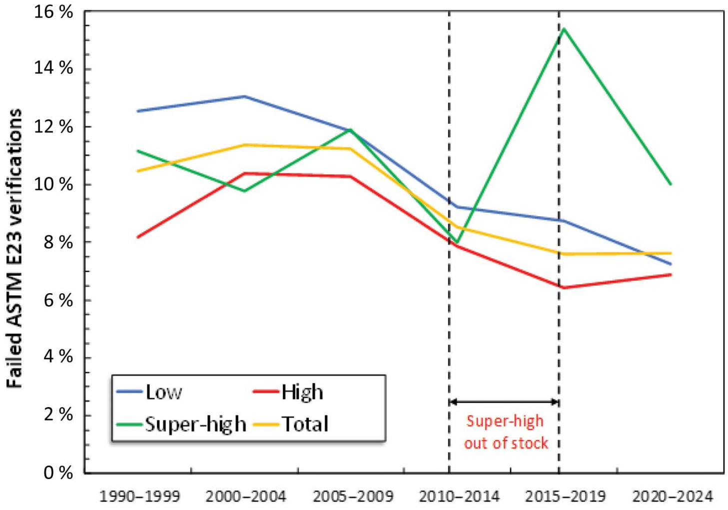
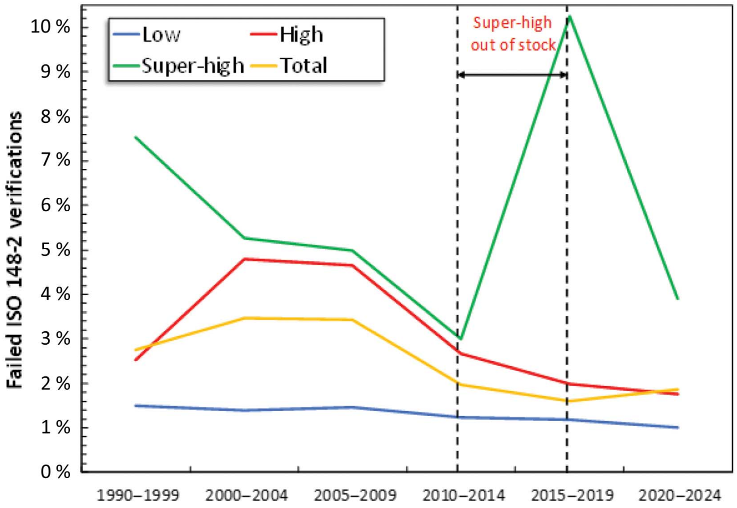


FIG. 3

Percentage of failed ISO 148-2 verifications between 1990 and 2024.



The former remark confirms the well-known fact that the ISO 148-2 indirect verification requirements are way more lenient than those of ASTM E23, as **figure 4** also clearly shows across all energy levels. A comparison between **figures 2** and **3** shows that this is particularly true at the low-energy level, where the ISO 148-2 limit (4 J) is almost three times larger than that of ASTM E23 (1.4 J). At the other energy levels, the ISO 148-2 limit is twice as large as the ASTM E23 limit (10 % versus 5 %).

A CLOSER LOOK AT FAILED ISO 148-2 INDIRECT VERIFICATIONS

The percentages of indirect machine verifications from 1990–2024 that failed ISO 148-2 because of the bias requirements, the repeatability requirements, or both, are shown in **Table 2** and **figure 5** for each energy level and for all levels combined.

We can observe that more than half of the failures (50.8 %) are caused by failing just the repeatability criteria. The importance of the bias requirements decreases with the energy of the reference specimens, whereas

FIG. 4

Overall fail rates for ASTM E23 and ISO 148-2.

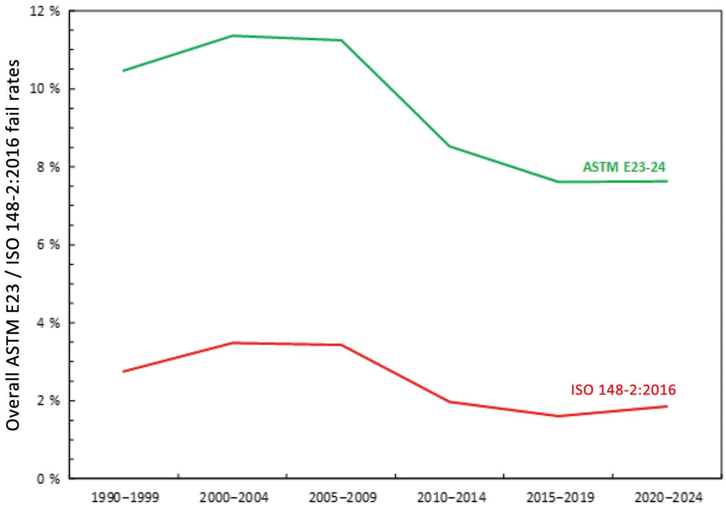


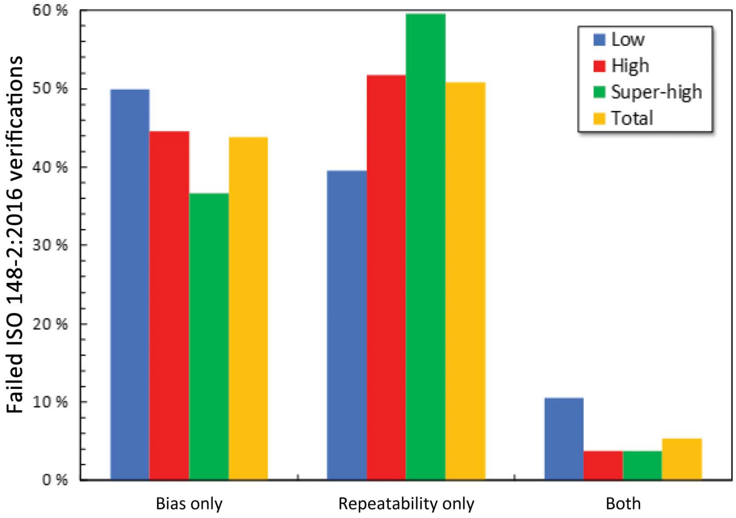
TABLE 2

Charpy machine indirect verifications, conducted with NIST reference specimens, that failed the ISO 148-2 requirements (1990–2024)

Requirements	Low	High	Super-High	Total
Bias only	49.9 %	44.5 %	36.6 %	43.9 %
Repeatability only	39.5 %	51.7 %	59.6 %	50.8 %
Both bias and repeatability	10.6 %	3.8 %	3.8 %	5.4 %

FIG. 5

Failed ISO 148-2 indirect verifications (1990–2024).



repeatability becomes progressively more dominant. Machines failing both bias and repeatability remain a minority; however, their share at the low-energy level (10.6 %) is more than double than at the high- or super-high–energy level (3.8 %).

Considering the global percentages of noncompliant machines that failed the bias requirements, whether alone or in conjunction with repeatability, the figures are 60.5 % (low energy), 48.3 % (high energy), 40.4 % (super-high energy), and 49.3 % (all energy levels). The corresponding figures for repeatability (alone or in conjunction with bias) are 50.1 % (low), 55.5 % (high), 63.4 % (super high), and 56.2 % (all levels). We can therefore state that the weight of the repeatability requirements increases with the energy of the reference specimens.

"EXTREME" SITUATIONS (ACCEPTABLE PER E23, BUT UNACCEPTABLE PER ISO 148-2 BECAUSE OF REPEATABILITY)

As already mentioned, ASTM E23 only requires that the average value of absorbed energy from five certified specimens, $\overline{K_V}$, matches the reference value, K_R, within the larger of 1.4 J and 5 % of K_R. However, certified Charpy specimens are produced by NIST according to strict metallurgical and machining specifications, which are specifically targeted at providing low variability (spread) of absorbed energies when tested on an acceptable machine.[11] Therefore, if we exclude specific and blatant experimental issues (jamming, incorrect specimen positioning, wrong test temperature, etc.), there should be a limit for the maximum difference between the highest and lowest absorbed energy values yielded by an acceptable machine. If such a limit is exceeded, the machine should not be deemed acceptable, even if $\overline{K_V}$ is relatively close to K_R.

Several indirect verifications included in the NIST database can be considered "extreme" situations, in that the corresponding test results satisfy the bias requirements of ASTM E23 (and automatically the bias requirements of ISO 148-2) but also fail the fairly generous repeatability requirements of ISO 148-2. According to what we previously stated, machines exhibiting this type of behavior should not be considered acceptable, despite passing the bias criteria. "Failing" machines or "failed" verifications, in this context, refer to passing ASTM E23 bias but failing ISO 148-2 repeatability.

Between 1990 and 2024, 735 verification sets (<1 % of all verifications) fell into this category. The corresponding number of machines is 565, about one sixth (15.3 %) of all the machines included in the NIST database. Almost one quarter (129) failed more than once over time. Although most of these failed twice, a small number of machines failed four (10) or five (14) times. Just eight of these machines failed at two different energy levels during the same indirect verification, and none failed at all three levels.

Of the 129 machines that failed more than once, about two thirds (62.8 %) experienced all failures over <5 years, 22.1 % between 6–10 years, 9.7 % between 11–15 years, and 4.8 % between 16–20 years. One machine had 26 years elapse between the oldest (1994) and the most recent (2020) failed verifications.

The information provided earlier shows that, although the overall number of failed verifications is small (only 1 % of all recorded verifications), a significant number of machines were affected, with a considerable share failing more than once, sometimes many years after the first failure. In other words, machine issues not addressed on first occurrence may recur multiple times over the years. The introduction of repeatability criteria for ASTM E23 would eliminate this problem and only marginally alter the overall ASTM E23 pass/fail ratio. The observation that repeatability issues seem to occur sporadically for many machines, rather than consistently, does not change the fact that a machine exhibiting this type of behavior should not be allowed to be successfully verified according to ASTM E23, until the causes have been investigated and suitably addressed. Moreover, as already mentioned, the steels used by NIST to produce reference specimens are very well-characterized and extremely homogeneous, as demonstrated by the very rigorous selection procedures and acceptability requirements implemented by NIST,[11] which are aimed at minimizing the variability of Charpy test results. Therefore, excessive scatter in Charpy absorbed energy values will be primarily caused by the machine performance, and any contribution of materials' variability can be considered negligible.

A detailed examination of these 595 machines revealed that more than half (310 [55 %]) subsequently failed the ASTM E23 bias requirements (88 machines failed three times or more). This may indicate that poor repeatability is often a symptom of machine issues that eventually cause ASTM E23 failures. It is also interesting that almost half of these 595 machines (263 [47 %]) have not been indirectly verified with NIST specimens in the last five years (2020–2024), which could mean they were either put out of service by the owners or simply deemed unsuitable to fulfill the ASTM E23 requirements. Obviously, because we do not have access to specific information about any of the machines in question, this latter remark should be considered purely speculative.

IMPLEMENTING A REPEATABILITY CRITERION FOR ASTM E23

As previously mentioned, ASTM E23 only requires test results from an indirect verification to be within a specified range with respect to the certified value of absorbed energy, K_R, established on the three NIST reference machines in Boulder, Colorado. The limit set by ASTM E23 is the larger between 1.4 J and 5 % of K_R. In practical terms, the limit corresponds to 1.4 J for low-energy specimens (typical K_R range: 13 J - 20 J) and 5 % for high and super-high–energy specimens (90 J - 140 J and 175 J - 240 J). The cutoff energy value between the absolute limit and the relative limit is 28 J (i.e., 1.4 J = 5 % of 28 J).

In this study, we attempted to answer the following questions:

(1) How does adding a repeatability criterion affect the ASTM E23 indirect verification pass/fail rates, with respect to the current bias-only version and the corresponding requirements of ISO 148-2?
(2) What repeatability limits could be proposed for a future ASTM E23 revision if the scientific community decides that the ASTM standard should impose a limit to the range of absorbed energies obtained during an indirect verification?

We reanalyzed the 77,105 customer records contained in the NIST Charpy database as of May 31, 2024, to assess the fail rates caused by the application of four different sets of repeatability requirements. The first (strictest) set of limits was established as follows:

- the ratio between low-energy bias limit for ASTM E23 (1.4 J) and ISO 148-2 (4 J) is 0.35;
- if the ISO 148-2 low-energy repeatability limit (6 J) is multiplied by 0.35, the result is 2.1 J; and
- 2.1 J corresponds to 7.5 % of 28 J (ASTM E23 cutoff energy between absolute and relative limits, as previously mentioned).

The first set of potential repeatability limits for ASTM E23 (indicated herein as “Rep1”) was therefore chosen as the larger between 2.1 J and 7.5 % of K_R. Three additional, more lenient, repeatability limits were then arbitrarily established, corresponding to 3 J and 11 % K_R (“Rep2”), 4 J and 14 % K_R (“Rep3”), and 5 J and 18 % K_R (“Rep4”). Note that, in the latter case, the absolute limit is below the ISO 148-2 value (6 J), whereas the corresponding relative limit is above the ISO 148-2 value (15 %).

ASTM E23 REP1 (PROPOSED REPEATABILITY LIMIT: LARGER BETWEEN 2.1 J AND 7.5 % OF K_R)

ASTM E23 fail rates, yielded by a combination of the current bias limits (1.4 J and 5 % K_R) and the first set of proposed repeatability limits (2.1 J and 7.5 % K_R), are provided in **Table 3**.

The number of failed indirect verifications for ASTM E23 jumps from 4,593 (6.0 %) for the current version to 26,491 (34.3 %, an almost six-fold increase) across all energy levels. For comparison, ISO 148-2 corresponds to 1,886 failed verifications (2.4 %). According to **Table 3**, more than one quarter of all failed verifications are caused by the repeatability criteria. **Figure 6** clearly shows that the first set of repeatability limits is overly severe.

TABLE 3

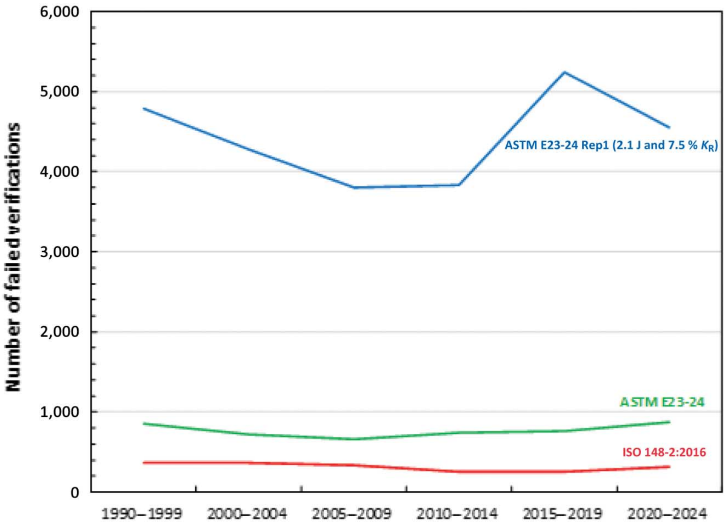
1990–2024 indirect verification fail rates for ASTM E23 Rep1, after adding the 2.1 J and 7.5 % K_R repeatability requirements

Requirements	Low	High	Super-High	Total
Bias only	7.0 %	4.8 %	6.0 %	6.0 %
Repeatability only	20.6 %	27.5 %	31.3 %	24.9 %
Both	3.6 %	3.2 %	4.6 %	3.5 %

Note: Percentages in the table are calculated with respect to the total number of indirect verifications at the specific energy level.

FIG. 6

Number of failed verifications from 1990–2024 according to ASTM E23, ISO 148-2, and ASTM E23 Rep1 (repeatability limits: larger of 2.1 J and 7.5 % of K_R).



ASTM E23 REP2 (PROPOSED REPEATABILITY LIMIT: LARGER BETWEEN 3 J AND 11 % OF K_R)

ASTM E23 fail rates produced by a combination of the current bias limits (1.4 J and 5 % K_R) and the second set of proposed repeatability limits (3 J/11 % K_R) are provided in **Table 4**.

The number of failed indirect verifications for ASTM E23 almost halves, dropping from 26,941 (34.3 %) for Rep1 to 13,796 (17.9 %) for Rep2, across all energy levels. **Table 4** reveals that bias and repeatability almost equally account for the failed verifications. Very few verifications (1.6 %) fail both bias and repeatability. Failed indirect verifications for ASTM E23, ISO 148-2, and ASTM E23 Rep2 are compared in **figure 7**, where the same Y-axis scale is used as in **figure 6** to facilitate comparison.

ASTM E23 REP3 (PROPOSED REPEATABILITY LIMIT: LARGER BETWEEN 4 J AND 14 % OF K_R)

ASTM E23 fail rates produced by a combination of the current bias limits (1.4 J and 5 % K_R) and the third set of proposed repeatability limits (4 J and 14 % K_R) are provided in **Table 5**.

The number of failed indirect verifications for ASTM E23 Rev3 drops further to 9,451 (12.2 %) with respect to 26,941 for Rep1 and 13,796 for Rep2 (34.3 % and 17.9 %, respectively) across all energy levels. In this case, bias is responsible for three times more failed verifications than repeatability. The number of verifications failing both bias and repeatability is extremely small (597 [0.8 %]). Failed indirect verifications for ASTM E23, ISO 148-2, and ASTM E23 Rep3 are compared in **figure 8**, where the same Y-axis scale used in **figures 6** and **7** is used to facilitate comparison.

TABLE 4

Indirect verification fail rates from 1990–2024 for ASTM E23 Rep2, after adding the 3 J and 11 % K_R repeatability requirements

Requirements	Low	High	Super-High	Total
Bias only	8.6 %	6.8 %	8.6 %	7.8 %
Repeatability only	9.4 %	7.0 %	10.5 %	8.4 %
Both	2.0 %	1.2 %	2.0 %	1.6 %

Note: Percentages in the table are calculated with respect to the total number of indirect verifications at the specific energy level.

FIG. 7

Number of failed verifications from 1990–2024 according to ASTM E23, ISO 148-2, and ASTM E23 Rep2 (repeatability limits: larger of 3 J and 11 % of K_R). Note: for easier comparison with **figure 6**, the same scale has been used for the Y-axis.

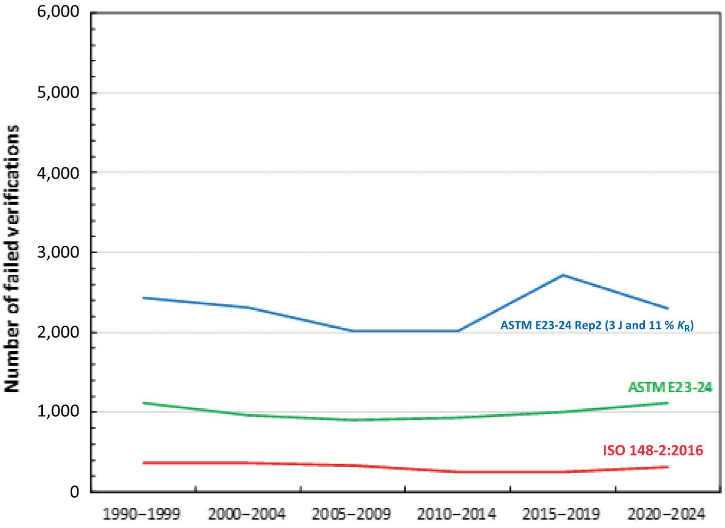


TABLE 5

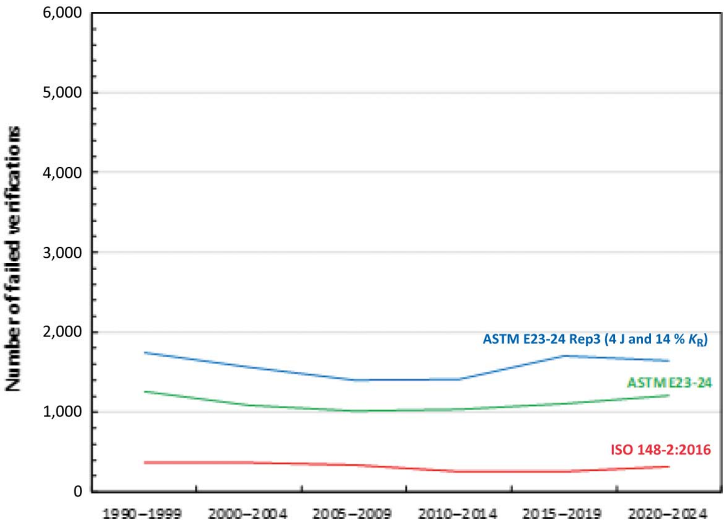
Indirect verification fail rates from 1990–2024 for ASTM E23 Rep3, after adding the 4 J and 14 % K_R repeatability requirements

Requirements	Low	High	Super-High	Total
Bias only	9.6 %	7.5 %	9.8 %	8.7 %
Repeatability only	3.4 %	1.9 %	3.9 %	2.8 %
Both	1.0 %	0.5 %	0.8 %	0.8 %

Note: Percentages in the table are calculated with respect to the total number of indirect verifications at the specific energy level.

FIG. 8

Number of failed verifications from 1990–2024 according to ASTM E23, ISO 148-2, and ASTM E23 Rep3 (repeatability limits: larger of 4 J and 14 % of K_R). Note: for easier comparison with **figures 6** and **7**, the same scale has been used for the Y-axis.



ASTM E23 REP4 (PROPOSED REPEATABILITY LIMIT: LARGER BETWEEN 5 J AND 18 % OF K_R)

ASTM E23 fail rates produced by a combination of the current bias limits (1.4 J and 5 % K_R) and the fourth set of proposed repeatability limits (5 J and 18 % K_R) are provided in **Table 6**.

For E23 Rep4, the number of failed indirect verifications (7,876 [10.2 %]) across all energy levels is only slightly higher than for the current E23-24 (7,014 [9.1 %]), indicating that these repeatability requirements might be too lenient because they are only responsible for <1 % of all failed verifications, as shown in **Table 6** and **figure 9**.

Discussion and Proposal for ASTM E23 Repeatability

Absolute and relative failed indirect verifications across all energy levels are summarized in **Table 7** for the current standards (ASTM E23 and ISO 148-2) and for the revised ASTM E23 using four different potential sets of repeatability limits (Rep1, Rep2, Rep3, and Rep4).

The number of failed indirect verifications are also illustrated in **figure 10**.

Another way to look at the information contained in the NIST database, with the aim of identifying reasonable repeatability criteria for ASTM E23, is by defining an "ASTM/ISO severity ratio," corresponding to the ratio between ASTM E23 and ISO 148-2 fail rates. The meaning of this parameter is intuitive: for example, a ratio of four between ASTM E23 and ISO 148-2 means that ASTM E23 is four times more severe than ISO 148-2. This ratio was calculated for verifications failing bias, failing repeatability, and overall failures (**Table 8**).

TABLE 6

Indirect verification fail rates from 1990–2024 for ASTM E23 Rep4, after adding the 5 J and 18 % K_R repeatability requirements

Requirements	Low	High	Super-High	Total
Bias only	10.0 %	7.8 %	10.3 %	9.1 %
Repeatability only	1.1 %	0.3 %	1.0 %	0.8 %
Both	0.5 %	0.2 %	0.4 %	0.4 %

Note: Percentages in the table are calculated with respect to the total number of indirect verifications at the specific energy level.

FIG. 9

Number of failed verifications from 1990–2024 according to ASTM E23, ISO 148-2, and ASTM E23 Rep4 (repeatability limits: larger of 5 J and 18 % of K_R). Note: for easier comparison with **figures 6–8**, the same scale has been used for the Y-axis.

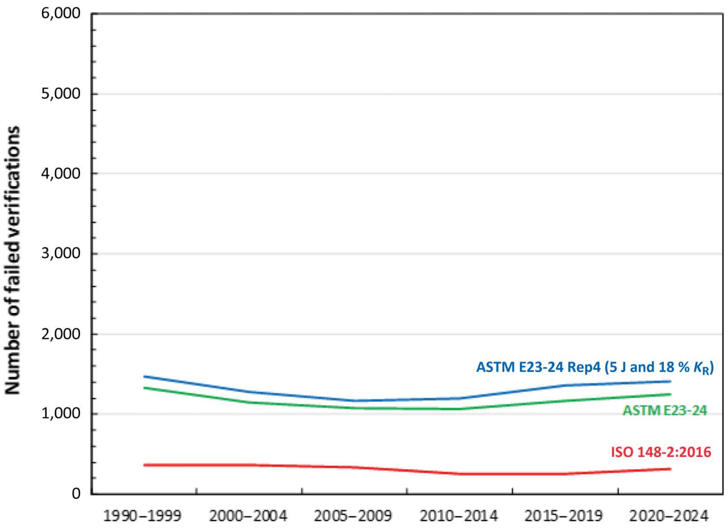


TABLE 7

Absolute and relative failed verifications for ASTM E23, ISO 148-2, ASTM E23 Rep1, ASTM E23 Rep2, ASTM E23 Rep3, and ASTM E23 Rep4

Failed Verifications	ASTM E23	ISO 148-2	ASTM E23 Rep1	ASTM E23 Rep2	ASTM E23 Rep3	ASTM E23 Rep4
Absolute	7,294	1,886	26,491	13,796	9,451	7,876
Relative	9.5 %	2.4 %	34.3 %	17.9 %	12.2 %	10.2 %

FIG. 10

Failed indirect verifications between 1990–2024 for ASTM E23, ISO 148-2, and the four potential revisions of ASTM E23.

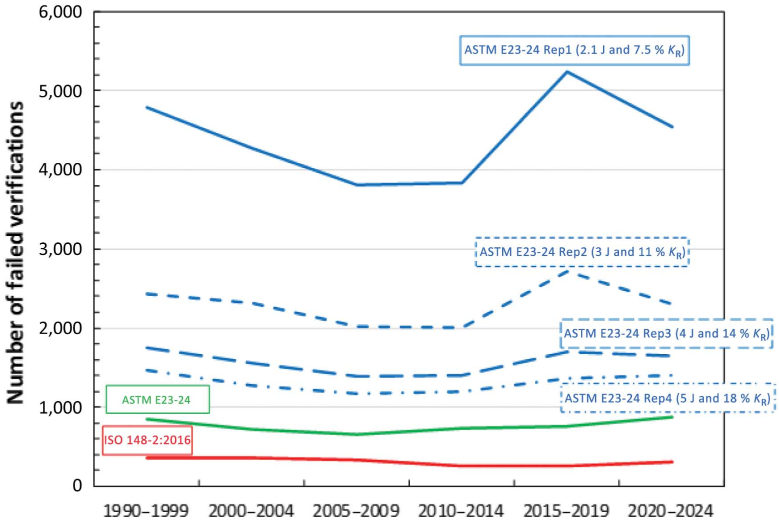


TABLE 8

ASTM/ISO severity ratios for the current and proposed versions of ASTM E23

	Ratio between ASTM and ISO Fail Rates (“Severity Ratio”)				
Indirect Verifications	ASTM E23	ASTM E23 Rep1	ASTM E23 Rep2	ASTM E23 Rep3	ASTM E23 Rep4
Failing bias	7.9	7.9	7.9	7.9	7.9
Failing repeatability		35.9	7.3	2.6	0.8
Overall failures	3.9	14.0	7.3	5.0	4.2

For the current version of ASTM E23, which does not include requirements on repeatability, the overall severity ratio with respect to ISO 148-2 is close to four. However, if only bias requirements are considered, the ratio almost doubles and approaches 8.

As progressively more lenient requirements on repeatability are considered for ASTM E23, the overall severity ratio decreases from 14 to just over 4.

Examination of the data displayed in **Tables 7** and **8** generates the following observations.

- The strictest set of requirements (Rep1) would make the ASTM E23 repeatability limits almost 36 times more severe than the ISO 148-2 limits and cause more than one third of all the verifications recorded by NIST since 1990 to become unacceptable (**Table 7**).
- The most lenient set of requirements (Rep4) provides a larger repeatability limit (18 %) for certified energies above 28 J than ISO 148-2 (15 %), which would make ASTM E23 unusually more generous than ISO 148-2.

The overall severity ratio (4.2) is marginally higher than that for ASTM E23 (3.9), and failed verifications only increase by 582 or an almost negligible 0.75 % of all indirect verifications between 1990–2024.

- The remaining two proposals (Rev2 and Rep3) obviously correspond to less extreme situations. The former (Rev2) yields a repeatability severity ratio (7.3) similar to the bias severity ratio (7.9), and the overall severity ratio increases by 87 % from 3.9 to 7.3. The overall number of failed indirect verifications would almost double, from 7,294 to 13,796. Conversely, the repeatability limits of Rev3, although remaining more severe than those of ISO 148-2, cause only a moderate increase of severity ratio from about 4 to 5 and an overall increase of 2,157 failures [2.8 %].

As far as the "extreme" situations previously described are concerned (verifications passing ASTM E23 but failing the ISO 148-2 repeatability criteria), they would be completely eliminated by the proposed Rep1, Rep2, and Rep3, whereas Rep4, which has a more lenient relative repeatability requirement (18 % versus 15 % for ISO 148-2) would end up retaining 431 verifications [0.6 % of the entire NIST database].

Based on the observations listed earlier, it is clear that Rep1 is overly restrictive and Rep4 is overly lenient, whereas both Rep2 and Rep3 represent reasonable options for a future revision of ASTM E23 that would introduce requirements on the maximum spread of indirect verification test results.

The recommendation of this author is to implement Rep3 (larger between 4 J and 14 % of K_R), which moderately increases the overall fail rate (from 9.5 % to 12.2 %), invalidates all verifications corresponding to "extreme" situations as defined in this study, and still allows ASTM E23 to remain more restrictive than ISO 148-2.

Conclusions

An in-depth examination of the NIST Charpy database, which includes test results reported by customers since 1990, showed that the current ASTM E23 requirements for the indirect verification of a Charpy machine, which are solely based on bias (the difference between the average and reference absorbed energies), are approximately four times more severe than those of ISO 148-2, which instead concern both bias and repeatability (the difference between the largest and smallest values of absorbed energy).

The absence of repeatability requirements in ASTM E23 caused 1 % of all recorded indirect verifications to be considered acceptable, although they would violate the rather generous repeatability criteria of ISO 148-2. Albeit this percentage is marginal, the corresponding share of customer machines falling within this category is about one sixth of all the machines verified by NIST since 1990. A quarter of these machines failed ISO repeatability more than once through the years (a few even four or five times) at different energy levels. This should not be considered acceptable, because NIST reference specimens are expected to provide low scatter among replicate tests. It is the personal but firm opinion of the author that machines displaying this behavior should not be considered in good working order, and as a minimum, a simplified direct verification and a repeated indirect verification would be warranted. We therefore advocate that ASTM E23 be revised to address this circumstance.

Four potential sets of repeatability criteria for ASTM E23 have been investigated in this study and compared with ISO 148-2. The one recommended herein for a possible revision of ASTM E23 would require the repeatability of indirect verification results to be within the larger of 4 J and 14 % of the certified absorbed energy, K_R. When applied to the 77,105 indirect verification sets included in the NIST Charpy database, these requirements would cause a moderate increase of the overall fail rate (from 9.5 % to 12.2 %), but also fail machines that currently pass ASTM E23 bias and fail ISO 148-2 repeatability. The proposed limits are offered to the consideration of ASTM E23 and the Charpy community as a compromise between avoiding that a machine with poor repeatability be considered acceptable by ASTM E23, and an excessive increase of the current fail rate.

This proposal, although causing the ASTM E23 overall acceptability criteria to become only marginally tighter, would result in better machine verification and would de facto constitute an improvement of the standard. The occurrence of a structural failure due to the use of a machine that passed ASTM E23 but would have failed

ISO 148-2 would become unlikely, and liability issues, which could destroy a company, would be avoided. The industry would generally be helped and better protected.

References

1. S. B. Russell, "Experiments with a New Machine for Testing Materials by Impact (Reprint from 1898)," in *Pendulum Impact Testing: A Century of Progress*, ed. T. A. Siewert and M. P. Manahan, Sr. (West Conshohocken, PA: ASTM International, 2000), 237–250, https://doi.org/10.1520/stp14385s
2. G. Charpy, "Essay on the Metals Impact Bend Test of Notched Bars (Reprint from 1901)," in *Pendulum Impact Testing: A Century of Progress*, ed. T. A. Siewert and M. P. Manahan, Sr. (West Conshohocken, PA: ASTM International, 2000), 848–877, https://doi.org/10.1520/STP14386S
3. *Standard Test Methods for Notched Bar Impact Testing of Metallic Materials*, ASTM E23 (West Conshohocken, PA: ASTM International, approved April 1, 2024), https://doi.org/10.1520/E0023-24
4. *Metallic Materials — Charpy Pendulum Impact Test — Part 1: Test Method*, ISO 148-1:2016 (Geneva, Switzerland: International Standards Organization, 2016).
5. D. E. Driscoll, "The Charpy Impact Machine and Procedure for Inspection and Testing Charpy 'V' Notch Impact Specimens," *ASTM Bulletin* 191 (1953): 60–64.
6. D. E. Driscoll, "Reproducibility of Charpy Impact Test," in *Symposium on Impact Testing*, ed. F. G. Tatnall (West Conshohocken, PA: ASTM International, 1956): 70–75, https://doi.org/10.1520/STP47578S
7. *Metallic Materials — Charpy Pendulum Impact Test — Part 2: Verification of Testing Machines*, ISO 148-2 (Geneva, Switzerland: International Standards Organization, 2016).
8. E. Lucon and R. L. Santoyo, "Analyzing the NIST Charpy Program Database: Influence of Impact Hammer Type (C Versus U) on Test Results," *Journal of Testing and Evaluation* 50, no. 5 (September 2022): 2465–2481, https://doi.org/10.1520/JTE20220198
9. M. P. Manahan, Sr. and R. B. Stonesifer, "The Difference between Total Absorbed Energy Measured Using an Instrumented Striker and That Obtained Using an Optical Encoder," in *Pendulum Impact Testing: A Century of Progress*, ed. T. A. Siewert and M. P. Manahan, Sr. (West Conshohocken, PA: ASTM International, 2000): 181–197, https://doi.org/10.1520/STP14394S
10. J. Schuurmans, "Influence of Striker Design on the Absorbed Energy of Low-Energy Reference Specimens Used for the Indirect Verification of Charpy Impact Testers," *Journal of Testing and Evaluation* 51, no. 5 (September 2023): 3574–3583, https://doi.org/10.1520/JTE20220515
11. 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* (Boulder, CO: National Institute of Standards and Technology, 2003), 3–42, https://doi.org/10.6028/NIST.TN.1500-9