Validating Crack-Detection In-Line Inspections with API 1163

Tom Bubenik and Matt Ellinger
DNV USA

PPIM 2025
PIPELINE PIGGING & INTEGRITY MANAGEMENT
JANUARY 27-31 • HOUSTON

Organized by
CLARION
TECHNICAL CONFERENCES

https://doi.org/10.52202/078572-0090

https://doi.org/10.52202/078572-0090

Abstract

Pipeline operators use the verification and validation requirements of API Standard 1163, In-Line Inspection Systems Qualifications, to ensure high quality in-line inspections (ILIs). Verification confirms an ILI was conducted according to plan, procedures, and processes and that the inspection conditions are consistent with those used to establish the ILI Performance Specification. Validation evaluates the accuracy of the reported anomaly types and characteristics (depth, length, width, etc.) and provides assurance the ILI met its Performance Specification.

Many of the guidelines (i.e., “should” statements) and requirements (i.e., “shall” statements) in API 1163 require the pipeline operator to establish acceptance limits based on pipeline industry experience. Often, API 1163 does not provide guidance on how to establish the acceptance limits. To address this gap, PRCI developed a guidance document and spreadsheet for conducting the Level 2 and 3 validations and establishing acceptance limits in accordance with the third Edition of API 1163.

Both API 1163 and the PRCI documents are becoming widely used in industry to verify and validate metal-loss ILIs, but questions have been raised about the ability to use them for crack-detection inspections. These questions primarily concern the ability to use field non-destructive evaluation (NDE) techniques, which have inaccuracies similar to the stated ILI Performance Specifications, to validate the inspections. Simply put, how can an operator validate an inspection when field validation measurements are questionable?

This paper assesses the ability to use API 1163 for crack-detection ILIs including, but not limited to, the following:

- API 1163 and the PRCI guidance report/spreadsheet provide a method for incorporating field NDE tolerances into validation calculations. However, large field NDE tolerances increase the risk of accepting an ILI’s accuracy when, in fact, the ILI did not meet its Performance Specification.
- Analyses were used to evaluate the impact of typical field NDE uncertainties. The analyses demonstrate that using API 1163 and the PRCI spreadsheet for conducting Level 2 and Level 3 validations can be nonconservative if applied without correction factors.
- The analyses discussed in this paper provide a path forward for validating a crack-detection ILI based on paired ILI and field depth measurements. The implications of using such analyses are identified and discussed, and broad guidelines are given.

A goal of this work is to develop a methodology for validating ILI Performance Specifications from ILI and field NDE measurements, both of which contain errors. Two case studies are included to

https://doi.org/10.52202/078572-0090

demonstrate the impact of field NDE tolerances on API 1163 validations. Finally, areas for future work are identified.

Introduction

API Standard 1163, In-Line Inspection [ILI] Systems Qualifications, first published in 2005 is a tool for pipeline operators to ensure accurate and repeatable in-line inspections. The standard has evolved through two subsequent editions, one issued in April of 2013 and the other in September of 2021. The third (current) edition provides explicit requirements for Level 1, Level 2, and Level 3 ILI system validations not included in the previous editions.

To provide guidance, PRCI developed a report and spreadsheet for conducting validations. An expansion and revisions of the PRCI documents are underway and expected to be completed in 2025. The original PRCI documents are available as PRCI Project IM-1-06, "API-1163 Section 8 - Level 1, 2 & 3 Research" at the PRCI website (www.prci.org).

This paper discusses Level 2 and 3 validations as implemented in the PRCI guidance document and spreadsheet. Background information on API 1163 third edition validations is given in a sister publication.[*] A brief summary follows.

Level 2 validations involve analyses using *null hypothesis testing of a binomial event,* which was included in an optional appendix (annex) in API 1163, 2nd edition. The testing is based on the Agresti-Coull interval method[†‡] due to its accuracy and simplicity.

Level 3 validations estimate the actual ILI performance based on field measurements. Level 3 validations use tolerance interval calculations that estimate the actual tolerance associated with the certainty level given in the ILI performance specification. The methodology used is based on work by Hahn and Meeker[§], as implemented by Ludlow[**].

Level 2 and Level 3 validations serve different purposes. As a result, the safety factors to be used with Level 2 and Level 3 validations should be different. Level 2 validations allow the use of the ILI tool

---

[*] Bubenik, Tom and Ellinger, Matt, Lessons Learned Using API 1163 for Metal-Loss ILIs, Pipeline Pigging and Integrity Management Conference 2025, 2025.

[†] Agresti, A. and Coull, B., Approximate Is Better Than "Exact" for Interval Estimation of Binomial Proportions, The American Statistician, 52(2), 119–126, 1998.

[‡] Brown, L.D., Cai, T.T., and DasGupta, A., Interval Estimation for a Binomial Proportion. Statistical Science, 16(2), 101–133, 2001.

[§] Hahn, G.J. and Meeker, W.Q., Statistical Intervals—A Guide for Practitioners, New York: John Wiley & Sons, 1991

[**] Ludlow, J., Statistical Tolerance Intervals: A Better Approach to In-Line Inspection Performance Assessment, in NACE Corrosion 2012 Conference & Expo, 2012

https://doi.org/10.52202/078572-0090

tolerance specifications even when the inspection may not have been met, while Level 3 calculates an "actual" tolerance, which could be used in lieu of the ILI tool specification. The Level 3 tolerances can be higher than the ILI performance specification when the inspection just passes the Level 2 validations.

Level 2: Agresti-Coull null hypothesis testing of a binary event

The basis for Level 2 validations in API 1163 is *null hypothesis testing of a binary event*. *Null hypothesis testing* uses *statistical inference* to assess whether a null hypothesis can be rejected. The *null hypothesis* is that the ILI met its performance specification, i.e., that the actual depths of ILI-reported anomalies are within the performance specification's tolerance (e.g., ± 10% of the wall thickness) and certainty (80% of the time). The *alternative hypothesis*, which the method seeks to demonstrate, is that ILI anomalies are not within the performance specification's tolerance.

Statistical inference allows the user to make a judgment (whether the performance specification can be rejected) about a population (all ILI anomalies on a pipeline) based on a sample (a subset of the population). The *binary event* being evaluated is whether the actual depth of each ILI-reported anomaly is or is not within specification. The degree to which the actual depth is or is not within specification is not considered. The method, for example, does not distinguish between depths that are barely out of specification from those that are largely out of specification.

The API 1163 Level 2 third edition statistical calculations are based on a 90% two-tailed confidence level. The sample is the ILI-reported anomalies selected by the pipeline operator for field measurements. Typically, this sample is less, maybe much less, than the ILI population. The confidence level refers to how frequently a random sample of field measurements will meet the stated certainty level. So, for example, if the confidence level is 90% and the certainty level is 80%, about 90% of random samples of field measurements will have 80% or more within ± 10% of the actual depth.

In API 1163, the concept of null hypothesis significance testing is expanded to help the user make a judgment on whether the performance specification could or should be accepted based on a sample set of ILI-field-measurement data. As such, Level 2 validation does not prove that an ILI met its performance specification, but it provides confidence that the specification was reasonable for the ILI being evaluated.

In its simplest form, Level 2 validations estimate the range of certainties the ILI may have met at its stated tolerance. The estimated certainty range is given as an upper and lower limit or endpoint, which is a function of the confidence level used in the analysis. If the uncertainty range includes or exceeds the stated certainty, the inspection may have met its performance specification. In the PRCI

https://doi.org/10.52202/078572-0090

guidance document and spreadsheet, if the upper limit or endpoint exceeds the stated certainty and the lower limit or endpoint exceeds 0.6 (or 60%), the validation is considered successful and the performance specification is accepted.

Implementation of Agresti-Coull

Field measurements of pipeline anomaly depths have inherent uncertainty. That is, there is a potential error associated with the measured depth. The potential for error is recognized in API 1163 (all editions) and incorporated into the way the Agresti-Coull method is implemented. Specifically, the method considers whether the ILI-reported depth could be within specification when there is uncertainty in both the ILI-reported and field measured depths.

The Agresti-Coull method as described in API 1163 and the PRCI guidance document and as implemented in the PRCI spreadsheet is evaluated against the total inaccuracies or errors from both the ILI and field depth measurements:

$$\sigma_e = \sqrt{\sigma_{ILI}^2 + \sigma_{Field}^2}$$

where σ_e represents the standard deviation of the total error (i.e., the combined error from ILI-reported and field-measured values), σ_{ILI} represents the error attributed to the ILI, and σ_{Field} represents the field error. This approach assumes the ILI and field errors are independent and random. API 1163 states that ignoring field-measurement inaccuracies is generally conservative. While this may be true when the field inaccuracy is much less than the ILI inaccuracy, it is not true when the field and ILI inaccuracies are of similar magnitude.

The *binary* event being tested in the Agresti-Coull method as described in API 1163 and implemented in the PRCI spreadsheet evaluates whether the actual depths are within the combined tolerances of the ILI and field measurements. The method is agnostic as to where the error resides. So, for example, if the field depth error for a given case happens to be zero (i.e., the field measurement is perfectly accurate), all of the error can be attributed to the ILI and vice versa. As a result, the analyses in API 1163 and the PRCI spreadsheet do not determine where the combined error resides.

Typical inaccuracies in ILI and field NDE measurements of cracks

There are many sources of uncertainty in both ILI and field depth measurements. Many crack-like anomalies, such as hook cracks and lack of fusion in welds, are not ideal reflectors. Their surfaces can be curved and angled to the pipe surface and axis, and the crack tips can be irregular. Both affect the strength (magnitude) of ultrasonic echoes, which are often used to determine the ILI and field

https://doi.org/10.52202/078572-0090

depths. In addition, pipe geometry affects the inspection, especially around welds, where reinforcement and misalignment impact sound wave paths. Finally, inclusions and natural material artifacts (e.g., laminations) produce spurious signals that must be differentiated from signals from crack-like anomalies.

Typical inaccuracies of ILI crack detection, identification, and sizing systems are given in absolute terms as ±1 mm or ±40 mils. Alternatively, the inaccuracies may be defined as a percentage of the wall thickness, for example, ±10% wall thickness (WT) to ±25% WT. Typically, inaccuracies of field NDE depth measurements are between as ±0.5 to 1 mm or ± 20 to 40 mils. Table 1 shows equivalent tolerances in percent WT corresponding to a stated confidence of ±1 mm.

Table 1. Tolerances at ± 1 mm

Wall Thickness (inches)	Tolerance (%WT)
0.188	21%
0.250	16%
0.313	13%
0.375	10%
0.438	9%
0.500	8%

PRCI recommendations for Level 2 validations

A Level 2 validation estimates a two-tailed range of certainties associated with an ILI. At a given confidence level, the upper confidence limit or endpoint represents the “best” the ILI may have performed, and the lower confidence limit or endpoint represents the “worst” the ILI may have performed. Thus, for example, the range of uncertainties could be [27%,73%], meaning the actual certainty of the ILI-reported anomalies that have not been investigated is at least 27% but not more than 73% at the confidence level chosen.

Outcomes 1, 2, and 3 are related to where the tolerance stated in the ILI performance specification falls relative to the upper and lower confidence bounds or endpoints. Outcome 1 refers to the situation where the stated tolerance is above the upper and lower confidence bounds. Outcome 2 is when the stated tolerance is within the upper and lower confidence bounds. Outcome 3 is when the stated tolerance is less than the upper and lower confidence bounds.

https://doi.org/10.52202/078572-0090

API 1163 (third edition) states "Outcome 1 generally means the ILI did not meet its performance specification, Outcome 2 means the ILI may have met its performance specification, and Outcome 3 means the ILI met or exceeded its performance specification."

API 1163 states that a pipeline operator should establish an acceptable lower confidence bound in accordance with the operator's risk tolerance, but it does not offer further guidance. The PRCI guidance document recommends a lower confidence interval endpoint of 60%, which is implemented by default in the PRCI spreadsheet, with a two-sided Level 2 confidence level of 90%. A lower endpoint of 60% implies the certainty associated with an ILI stated tolerance (e.g., ±10% WT) may have been as low as 60%, versus a typical value of 80%.

Case study

The third edition of API 1163 and the PRCI spreadsheet were used to evaluate a crack detection ILI and its corresponding field depth measurements. No destructive (actual) depths were available. The ILI and field measurements were adjusted and scaled to maintain anonymity. Forty crack-like anomalies were reported by the ILI, which had an assumed tolerance of ±1 mm with 80% certainty. The tolerance of ±1 mm corresponds to 10.67% WT for the (adjusted) wall thickness of 0.375 inches. ILI-reported depths range from 13% WT to 47% WT, while field-measured depths are from 7% WT to 52% WT.

A unity plot of the ILI and field depths is shown in Figure 1. Twenty-six of the 40 ILI-reported anomalies were matched and within the ILI tolerance of the field measurements – there were no false positives – for a success rate of 65% versus the ILI certainty specification of 80%. Three of the calls that were out of specification were conservatively reported (the ILI-reported depth exceeded the field-measured depth by over 10.67% WT), and 11 were unconservatively reported (the ILI-reported depth was more than 10.67% less than the field-measured depth).

https://doi.org/10.52202/078572-0090

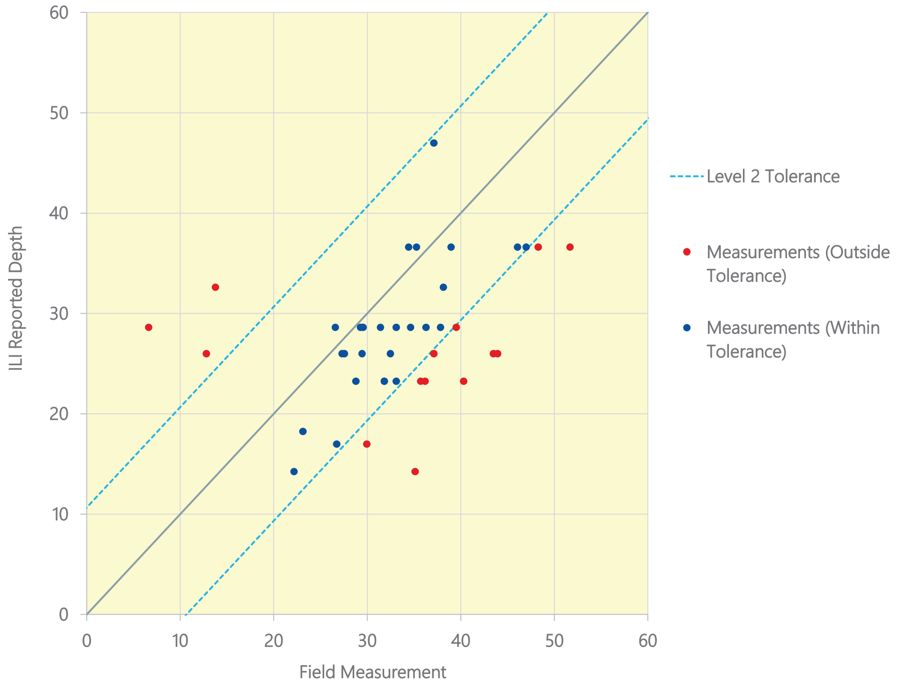


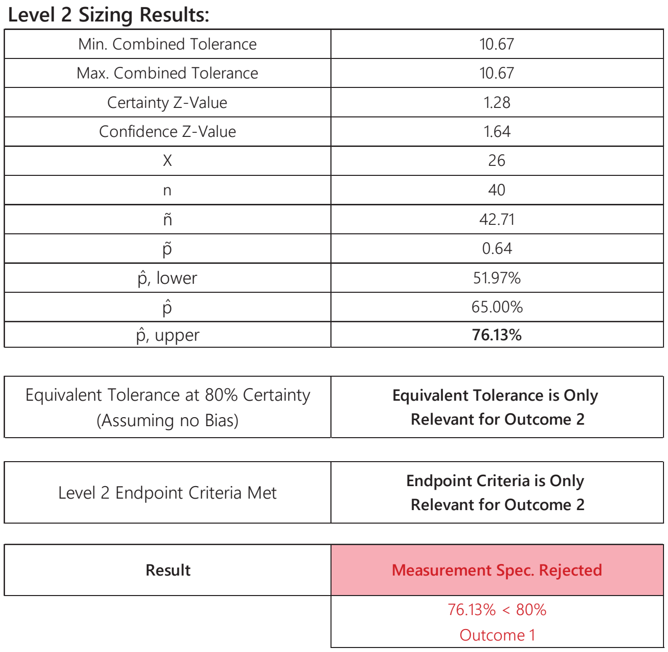
Figure 1. Level 2 Crack detection case study unity plot (field-measurement tolerance = 0)

Level 2 Validation

A Level 2 validation was conducted on the data shown in Figure 1 using the PRCI spreadsheet. The purpose was to determine the impact of various assumed field-measurement tolerances on the overall outcome of the validation exercise.

Figure 2 shows the validation results for a case where the field-measurement tolerance is zero. Here, "X" (in the fifth row of the table) represents the number of data points within specification (26), and n (sixth row of the table) is the total number of matched ILI and field depths (40). The lower confidence limit [($\tilde{p}$), lower] is 51.97% and the upper confidence limit [($\tilde{p}$), upper] is 76.13%. In this case, the lower confidence limit is below 60% and the upper confidence limit is below the certainty stated in the performance specification (80%). As a result, the validation concludes with Outcome 1 and the ILI performance specification can be rejected at the confidence level assumed in the analysis (90%).

https://doi.org/10.52202/078572-0090



Level 2 Sizing Results:

Min. Combined Tolerance	10.67
Max. Combined Tolerance	10.67
Certainty Z-Value	1.28
Confidence Z-Value	1.64
X	26
n	40
ñ	42.71
p̃	0.64
p̂, lower	51.97%
p̂	65.00%
p̂, upper	**76.13%**

Equivalent Tolerance at 80% Certainty (Assuming no Bias)	**Equivalent Tolerance is Only Relevant for Outcome 2**

Level 2 Endpoint Criteria Met	**Endpoint Criteria is Only Relevant for Outcome 2**

Result	Measurement Spec. Rejected
	76.13% < 80% Outcome 1

Figure 2. Crack detection case study Level 2 validation results (field-measurement tolerance = 0)

The above results are for the case where field-measurement uncertainty is zero. Adding a field uncertainty increases the combined tolerance evaluated in the Level 2 validation. Figure 3 shows the corresponding plot when the field tolerance equals the ILI tolerance (±1mm or 10.67% WT). The Level 2 tolerance (dashed) lines have moved apart because combined tolerance of the ILI and field measurements is now ±15.09% WT. Thirty-four of the 40 matched anomalies were within the combined tolerance for a success rate of 85% versus the ILI certainty specification of 80%.

https://doi.org/10.52202/078572-0090

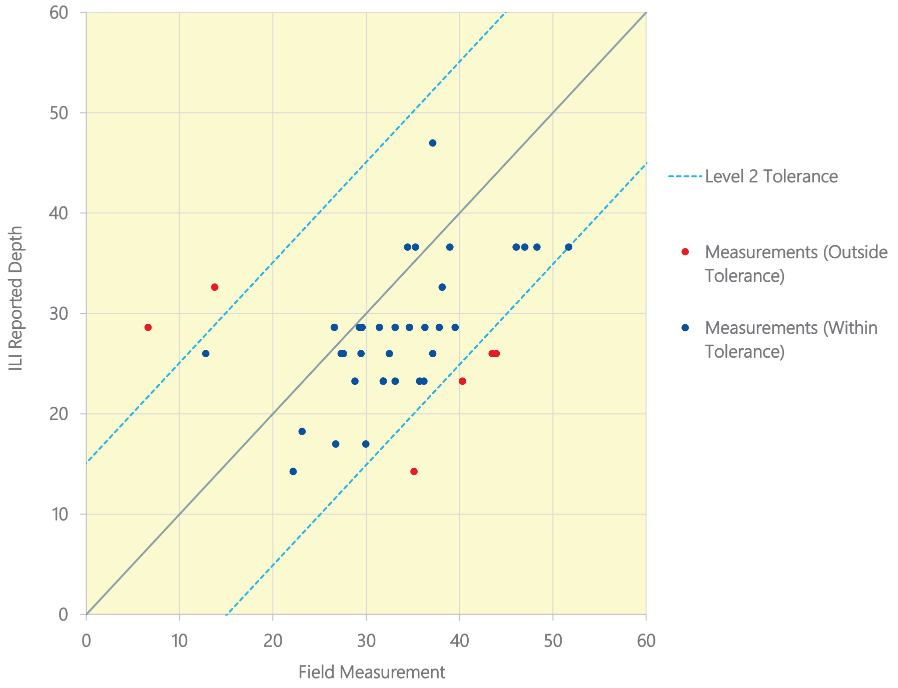
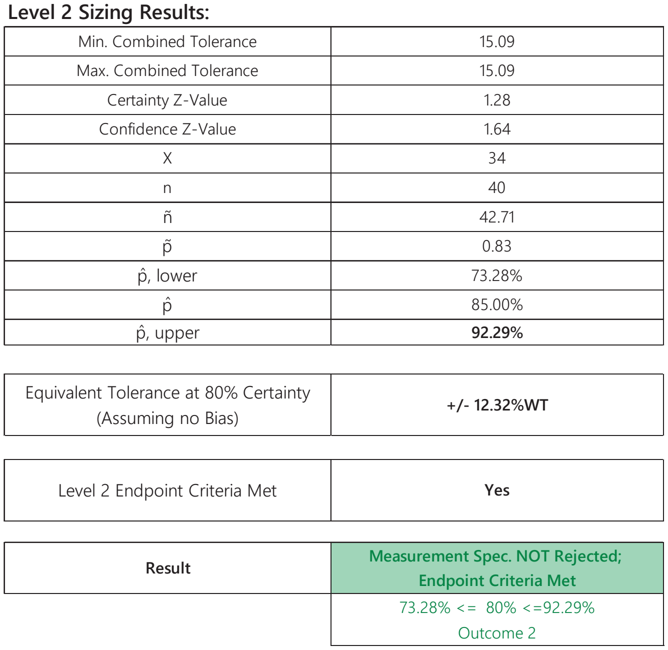


Figure 3. Level 2 crack detection case study unity plot (field-measurement tolerance = 10.67% WT)

Figure 4 shows the corresponding analysis results. The lower confidence limit is now 73.28%, and the upper confidence limit is 92.29%. In this case, the lower confidence limit is above 60% and the upper confidence limit is above the certainty stated in the performance specification (80%). As a result, the validation concludes with Outcome 2. Outcome 2 means the ILI performance specification may be accepted per API 1163. So, adding the field measurement tolerance changes the results from rejecting the ILI at a confidence level of 90% to accepting it.

https://doi.org/10.52202/078572-0090



Level 2 Sizing Results:

Min. Combined Tolerance	15.09
Max. Combined Tolerance	15.09
Certainty Z-Value	1.28
Confidence Z-Value	1.64
X	34
n	40
ñ	42.71
p̃	0.83
p̂, lower	73.28%
p̂	85.00%
p̂, upper	**92.29%**

Equivalent Tolerance at 80% Certainty (Assuming no Bias)	**+/- 12.32%WT**

Level 2 Endpoint Criteria Met	**Yes**

Result	**Measurement Spec. NOT Rejected; Endpoint Criteria Met**
	73.28% <= 80% <=92.29% Outcome 2

Figure 4. Crack detection case study Level 2 validation results (field-measurement tolerance = 10.67% WT)

Increasing the field measurement tolerance from 10.67% WT to ±1.5 mm or 16% WT further complicates the picture. Results for this case are shown in Figure 5 and Figure 6. At a combined tolerance of ±19.23% WT, the Level 2 validation results in Outcome 3, in which case, API 1163 says the ILI exceeded its performance specification.

https://doi.org/10.52202/078572-0090

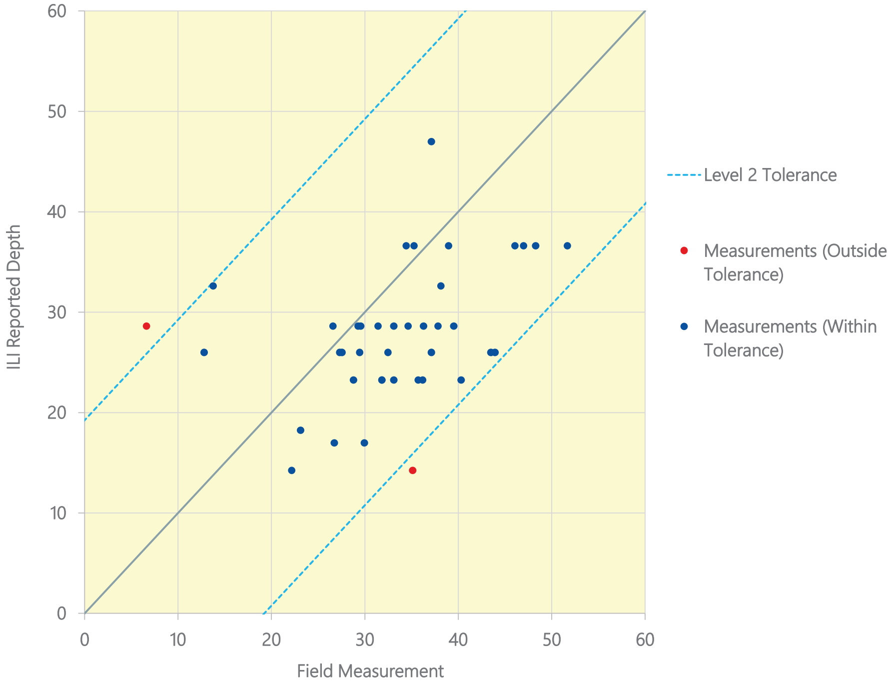
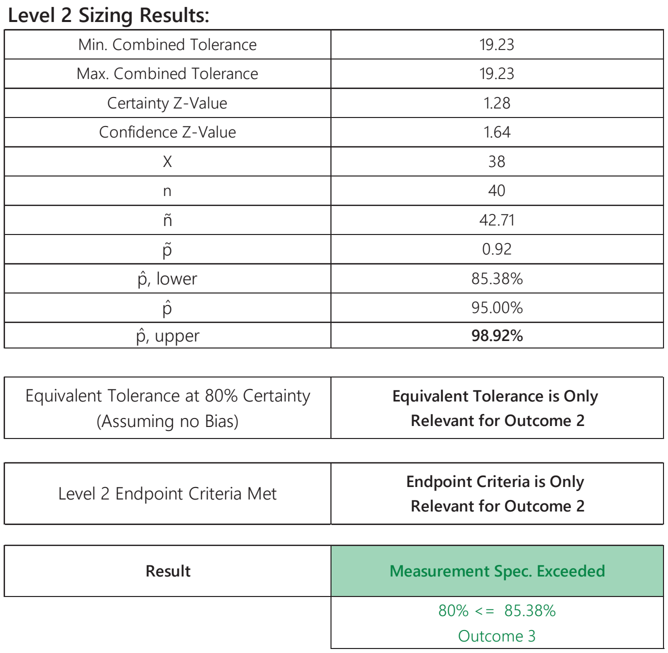


Figure 5. Level 2 crack detection case study unity plot (field-measurement tolerance = 16% WT)



Level 2 Sizing Results:

Min. Combined Tolerance	19.23
Max. Combined Tolerance	19.23
Certainty Z-Value	1.28
Confidence Z-Value	1.64
X	38
n	40
ñ	42.71
p̃	0.92
p̂, lower	85.38%
p̂	95.00%
p̂, upper	**98.92%**

Equivalent Tolerance at 80% Certainty (Assuming no Bias)	**Equivalent Tolerance is Only Relevant for Outcome 2**

Level 2 Endpoint Criteria Met	**Endpoint Criteria is Only Relevant for Outcome 2**

Result	**Measurement Spec. Exceeded**
	80% <= 85.38% Outcome 3

Figure 6. Crack detection case study Level 2 validation results (field-measurement tolerance = 16% WT)

The three cases demonstrate how assumed field measurement tolerances can impact the overall validation. By changing the assumed field tolerance, any of the three possible outcomes can be obtained. The difficulty posed by the three cases is due to the manner in which the Agresti-Coull method is implemented in API 1163, the PRCI guidance document, and the PRCI spreadsheet. Using Agresti-Coull, there is no simple way to decouple the ILI and field tolerances to determine whether the ILI errors are within the ILI tolerances. The Agresti-Coull evaluates the ILI performance relative to the combined ILI plus field tolerance, rather than the ILI tolerance alone.

Level 3 Validation

In API 1163, Level 3 validations are more sophisticated statistical methods that should provide more accurate estimates of the actual sizing bias and accuracy. The two example methodologies included in the API 1163 appendix (annex) call for the use of statistical tolerance intervals and Bayesian inference. The statistical tolerance interval method is easier to implement in Excel and is incorporated into the PRCI guidance document and spreadsheet. The method estimates the actual ILI performance from the available field measurements. The estimated performance is expressed as a bias and an estimated "actual" tolerance.

The tolerance interval method in API 1163, the PRCI guidance document, and the PRCI spreadsheet does not explicitly take into account field tolerances, but the calculated Level 3 tolerances are akin to the combined tolerances in the Level 2 validation. That is, they represent the total tolerance of the ILI and field measurements. So, they have the same limitations as the Level 2 results in that there is no way to determine if the actual tolerance is due to ILI inaccuracies, field NDE inaccuracies, or a combination of both.

For reference, Figure 7 and Figure 8 give the Level 3 unity plot and analysis results. The Level 3 validation returns a mean error or bias[††] of -5.46% WT (underprediction) and a tolerance of ±14.81% WT. This Level 3 tolerance is close to the Level 2 case where the ILI and field measurement tolerances were each assumed to be equal to ±1 mm or ±10.67% WT, giving a combined tolerance of 15.09% WT. The error bands in Figure 7 (blue dashed lines) depict the effects of incorporating the mean error and the tolerance (i.e., +9.35% WT to -20.27% WT).

[††] The effect of mean error or bias is not explicitly considered in the remainder of the paper.

https://doi.org/10.52202/078572-0090

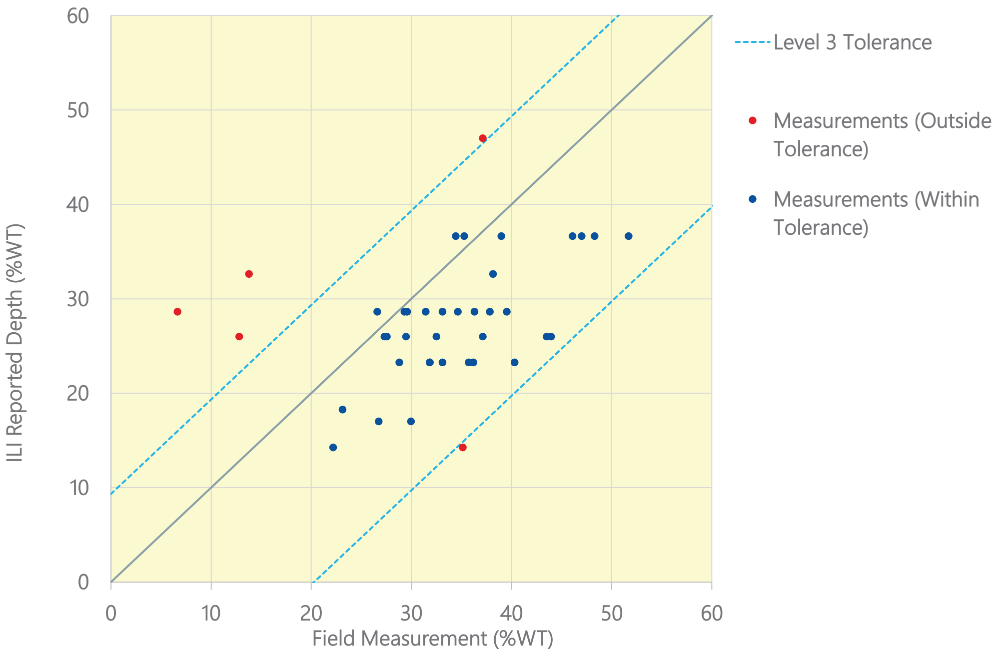
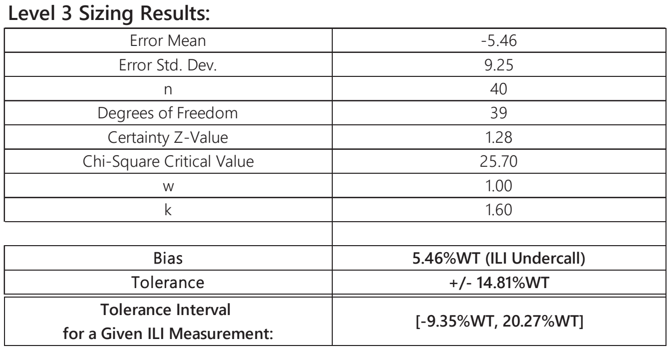


Figure 7. Level 3 crack detection case study unity plot

The Level 3 validation provides a combined ILI plus field tolerance of 14.81% WT, but it does not allow the ILI-only tolerance to be estimated. For this, more sophisticated statistical methods are required.


Level 3 Sizing Results:

Error Mean	-5.46
Error Std. Dev.	9.25
n	40
Degrees of Freedom	39
Certainty Z-Value	1.28
Chi-Square Critical Value	25.70
w	1.00
k	1.60
Bias	**5.46%WT (ILI Undercall)**
Tolerance	**+/- 14.81%WT**
Tolerance Interval for a Given ILI Measurement:	**[-9.35%WT, 20.27%WT]**

Figure 8. Crack detection case study Level 3 validation results

Combining Uncertain Data

When two sets of uncertain data are combined, the variances are additive, assuming the errors associated with each data set are independent:

$$Variance(ILI\ plus\ Field\ Error) = Variance(ILI\ Error) + Variance(Field\ Error) \quad (1)$$

The validations discussed previously represent the combined ILI plus field error. Solving for the ILI error alone requires knowledge of or an assumption about the actual field error. For example, if the

https://doi.org/10.52202/078572-0090

variance of the actual field measurement error is assumed to be equal the variance of the ILI error, solving Equation (1) for the ILI variance is trivial:

$$Variance(ILI\ Error) = Variance(ILI\ plus\ Field\ Error) \div 2 \tag{2}$$

However, this type of calculation is not valid if the ILI and field errors are correlated. If the errors are correlated, the combined variance of the ILI and field errors is as follows:

$$\begin{aligned} &Variance(ILI\ plus\ Field\ Error) \\ &\quad = Variance(ILI\ Error) + Variance(Field\ Error) - 2 * r * SD(ILI) * SD(Field) \end{aligned} \tag{3}$$

where *r* is the correlation coefficient between the ILI and field error and *SD(x)* represents the standard deviation (the square root of the variance) of *x*. Assuming the actual field measurement variance equals the ILI variance, solving for the ILI variance gives:

$$Variance(ILI) = \frac{Variance(ILI\ plus\ Field\ Error)}{2 - 2 * r} \tag{4}$$

Similarly, if the variance of the field error is some multiplier times the ILI error,

$$Variance(Field\ Error) = A * Variance(ILI\ Error) \tag{5}$$

where *A* is the multiplier. Solving for the ILI variance gives:

$$Variance(ILI) = \frac{Variance(ILI\ plus\ Field\ Error)}{1 + A\ -\ 2 * r * \sqrt{A}} \tag{6}$$

Thus, the ILI variance is a function of the multiplier, A, and the correlation coefficient, *r*, both of which are unknown.

To evaluate the impact of the multiplier and correlation coefficient, the authors evaluated a different data set, which contained ILI, field NDE, and destructive depth measurements. The data set consists of 60 data points with both ILI and field depth measurements, of which nine also had destructive depth measurements. Figure 9 shows results for cases where the field tolerance is assumed to be equal to the ILI tolerance (solid line) and for cases where the field tolerance is 60% and 80% of the actual ILI tolerance (dashed lines).

https://doi.org/10.52202/078572-0090

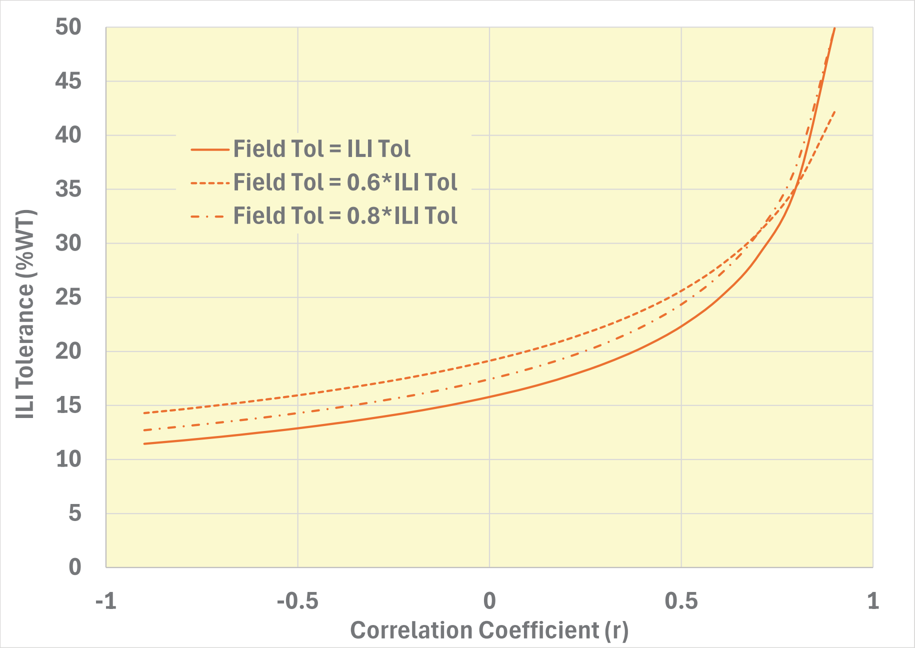


Figure 9. ILI Variance as a function of field multiplier, A, and correlation coefficient, r

It seems reasonable to assume that the field depth variances (or standard deviations or tolerances) are comparable to the ILI variances because many crack-detection ILI systems use ultrasonic shear beam technologies that are similar to the field NDE techniques. ILI tolerance is usually stated to be around ±1 mm (±40 mils), while field NDE tolerances are typically ±0.5 to ±1.0 mm (20 to 40 mils). Hence, the field multiplier, A, is likely in the range of 0.5 to 1.0 (the field tolerance is 50 to 100% of the ILI tolerance).

Errors associated with ILI and field depths versus actual depths might be correlated, depending on the ILI and field measurement technologies and data analysis. For the data set used in this analysis, the ILI and field depth measurements are weakly correlated (r=0.311). While the correlation coefficient between the ILI and field depth *measurements* is not necessarily the same as the correlation coefficient between the ILI and field depth *errors*, it seems plausible the ILI and field *error* correlation coefficient is similar. The following discussion assumes the correlation coefficient between ILI and field errors is between zero and 0.5.

Estimated Impact of Correlated ILI and Field NDE Errors

The calculated ILI tolerances as a function of the actual field accuracy (A=0.6, 0.8, 1.0) and ILI/field error correlation coefficient (r=0, 0.3, 0.5) are given in Table 2. Taking A=1.0 and r=0.3 as a base case, the calculated base ILI tolerance is ±18.9% WT. For the field accuracies and correlation coefficients considered, the calculated ILI tolerances range from ±15.8% WT (a decrease of 3.1%

https://doi.org/10.52202/078572-0090

WT from the base case tolerance) to ±25.6% WT (an increase of 6.7% WT). Over this range of field multipliers and correlation coefficients, the correlation coefficient is more impactful than the multiplier.

Table 2. Calculated ILI Tolerances in %WT

		Multiplier, A		
		0.6	**0.8**	**1**
Correlation Coefficient, r	**0**	19.4	17.4	15.8
	0.3	22.3	20.7	18.9
	0.5	25.6	24.3	22.3

The authors continue to analyze the destructive depth measurement data to determine the actual field depth tolerance, ILI/field error correlation coefficient, and ILI-truth error tolerance. Early analyses suggest:

- The actual field depth tolerance is smaller than the ILI depth tolerance (i.e., the field depths are more accurate than the ILI depths) – an actual multiplier less than 1.0 is justified.
- The ILI/field error correlation is of the order of 0.3, similar to the ILI/field depth correlation coefficient.
- The ILI/truth error tolerance is of the order of ±17.7, versus the values shown in Table 2.

The results of this ongoing work will be reported later. In addition, the authors are collecting more metallurgical data to allow better estimates of the field accuracy multiplier, the ILI/field error correlation coefficient, and the ILI/truth error.

Conclusions

This paper assesses the ability to use the third edition of API 1163 for crack-detection ILIs, where the uncertainty of the field-measured depth is similar to the uncertainty associated with the ILI reported depth. The paper also introduces a methodology for estimating the actual ILI tolerance (i.e., the difference between the ILI-reported depth and the true depth) from ILI and field NDE depths only (i.e., without "truth" data). The paper shows:

- The third edition of API 1163, and the corresponding PRCI guidance report and spreadsheet, provide a method for incorporating field NDE tolerances (uncertainties) into validation calculations. The approach used is largely applicable (and conservative) for cases where the field depth tolerance is small compared to the ILI depth tolerance – i.e., when the field depth measurements are more accurate than the ILI-reported depths.

- The third edition of API 1163 and the PRCI report and spreadsheet can provide nonconservative results when applied to cases where the field tolerance is large or assumed to be large (e.g., when the field and ILI tolerances are similar). That is, API 1163 and the PRCI report and spreadsheet can suggest the ILI met its Performance Specification when, in fact, it did not.
- Analyses were used to evaluate the impact of field NDE tolerances that are similar to ILI tolerances. The analyses show that the tolerance associated with the field depth measurements and the correlation coefficient between ILI errors and field NDE errors impact the calculated ILI tolerance used for validation.
- For a sample data set with ILI, field, and metallurgical depth measurements, the ILI tolerance calculated using only ILI and field depths is similar to the tolerance calculated directly from ILI and metallurgical ("truth") depths.

The analyses in this paper provide a path forward for validating a crack-detection ILI without access to "truth" data. Additional work is necessary to better define typical field depth measurement tolerances and the correlation coefficient between ILI and field depth errors.

https://doi.org/10.52202/078572-0090

https://doi.org/10.52202/078572-0090