

Decision Rules & Statements of Conformity

A Simplified Guidance Document for Calibration Laboratories

Addressing Common Questions from the Field

Based on ISO/IEC 17025:2017 [1] | UKAS LAB 48 [2] | ILAC-G8:09/2019 [3] | JCGM 106:2012 [4]

Primarily written by Henry Zumbrun

A special thank you to Greg Cenker and Dilip Shah for being on some of these email chains and co-authoring [Decision Rules Guidance](#), which was heavily drawn from for this document.

Introduction

This guidance document addresses practical questions that calibration laboratories frequently encounter regarding decision rules, statements of conformity, and the handling of measurement uncertainty when reporting pass/fail determinations. It draws on ISO/IEC 17025:2017 [1], UKAS LAB 48 [2], ILAC-G8:09/2019 [3], and JCGM 106:2012 [4].

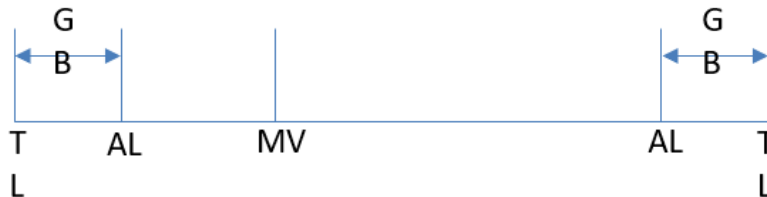
The questions and scenarios addressed here arose from real-world correspondence between calibration professionals and represent common points of confusion in the industry. This document systematically addresses those questions to help calibration laboratories and quality managers understand their obligations and best practices when issuing statements of conformity. AI-assisted tools were used in the preparation of this document, specifically to generate technical illustrations, apply and verify citations against source standards, and organize the original correspondence into a concise, structured guidance format. All technical content has been reviewed for accuracy against the referenced standards.

Who is this document for?

Laboratory managers, calibration technicians, quality managers, accreditation assessors, NUPIC, and anyone responsible for issuing purchase orders and/or issuing and interpreting calibration certificates that include pass/fail (conformity assessment) or in-tolerance/out-of-tolerance statements.

1. Key concepts and definitions

Before addressing specific scenarios, it is essential that all parties share a common understanding of the foundational terms in this area.



Term	Definition
Decision Rule	A rule that describes how measurement uncertainty is accounted for when stating conformity with a specified requirement. (ISO/IEC 17025:2017 [1], Clause 3.7)
Guard band (w)	An interval between a tolerance limit and a corresponding acceptance limit, used as a buffer to account for measurement uncertainty. Guard band = $ TL - AL $.
Tolerance Limit (TL)	The specified upper or lower bound of permissible values of a property (i.e., the specification limit).
Acceptance Limit (AL)	The specified upper or lower bound of permissible measured quantity values after applying a guard band.
Measured Quantity Value (MV)	The quantity value represents a measured result.

PFA – Probability of False Accept	The probability that a measurement result is declared conforming when the true value is actually non-conforming (consumer's risk).
PFR – Probability of False Reject	The probability that a measurement result is declared non-conforming when the true value is actually conforming (producer's risk).
TUR – Test Uncertainty Ratio	The ratio of the tolerance (TL) divided by the 95 % expanded measurement uncertainty of the measurement process. $TUR = TL / U$.
Simple Acceptance	A decision rule where a measurement is passed if the reported value falls within the tolerance limits, with no guard band applied. This carries significant PFA risk near the limits. ILAC-G8:09/2019 [3] (definition 1.9) defines simple acceptance as $AL = TL$. The broader guidance in the same document indicates that simple acceptance without a precondition constraint on measurement uncertainty (such as $TUR > N:1$ or an upper limit on allowable U) may not satisfy the Clause 3.7 requirement to account for measurement uncertainty in the decision rule.
Specific Risk	A decision rule that applies conditional probability at each measurement point to limit the risk of a false accept to a specified level.

Type I - Type II Error			
Statement of Conformity		Consequence	
		Actually in Conformance (ACCEPT)	Actually out of Conformance (REJECT)
Decision Made: (Based on Decision Rule)	Called In Conformance: ACCEPT	(1- α) Calibration Lab's Confidence (Probability of Correct Accept - PCA)	β : Type II Error (Probability of False Accept - PFA) CONSUMER'S RISK
	Called Out of Conformance: REJECT	α : Type I Error (Probability of False Reject - PFR) PRODUCER'S RISK	(1- β) End User's Confidence (Probability of Correct Reject - PCR)

Figure 1 Courtesy of Dilip Shah

Understanding that every conformity decision carries both Type I and Type II risk raises an obvious question: what does the standard actually require laboratories to do about it? The answer is less prescriptive than many expect. ISO/IEC 17025:2017 [1] does not mandate a specific guard band, a particular risk threshold, or any single mathematical method. What it does require is that the laboratory define, communicate, and document the decision rule it uses. The following is a plain-language summary of the four key clauses that govern this obligation.

2. What ISO/IEC 17025:2017 [1] requires

A frequent source of confusion is precisely what the standard mandates. The following is a plain-language summary of the four key clauses from ISO/IEC 17025:2017 [1] that govern decision rules.

2.1 Clause 3.7 – Definition

ISO/IEC 17025:2017 [1] defines a decision rule as: "a rule that describes how measurement uncertainty is accounted for when stating conformity with a specified requirement." Any conformity statement must be accompanied by a defined decision rule.

Note: This is the foundational requirement. If there is no measurement uncertainty, how can the measurement be metrologically traceable?

2.2 Clause 7.1.3 – Customer agreement

"When the customer requests a statement of conformity to a specification or standard for the test or calibration (e.g. pass/fail, in-tolerance/out-of-tolerance), the specification or standard and the decision rule shall be clearly defined. Unless inherent in the requested specification or standard, the decision rule selected shall be communicated to, and agreed with, the customer."

Note: This means the laboratory must proactively communicate the decision rule to the customer.

Additionally, Clause 6.2.6 requires that the laboratory authorize personnel to perform "analysis of results, including statements of conformity or opinions and interpretations." This means personnel making pass/fail determinations must be specifically authorized and competent to apply the decision rule in use. ILAC-G8:09/2019 [3] reinforces this requirement as part of its guidance on laboratory resources and competence.

2.3 Clause 7.8.6.1 – Documentation and risk

"When a statement of conformity to a specification or standard is provided, the laboratory shall document the decision rule employed, taking into account the level of risk (such as false accept and false reject and statistical assumptions) associated with the decision rule employed, and apply the decision rule."

2.4 Clause 7.8.6.2 – Reporting requirements

The calibration report must clearly identify: (a) to which results the statement of conformity applies; (b) which specifications, standards or parts thereof are met or not met; and (c) the decision rule applied (unless it is inherent in the requested specification or standard).

Important Clarification

ISO/IEC 17025:2017 [1] does NOT require guard bands, probability calculations, or any specific mathematical risk model. It requires that the decision rule be defined, applied, and (when applicable) agreed with the customer. Simple acceptance (zero guard band) may be used, though the laboratory must understand and document the associated risk, which can be very high near specification limits. Also, if the calibration certificate does not report measurement uncertainty, the measurement is not metrologically traceable.

3. Why simple acceptance without considering uncertainty is problematic

UKAS LAB 48 [2] illustrates this clearly with an example that has become an industry benchmark. Consider a situation where:

- Tolerance: ± 1 unit
- Decision Rule: PASS if measured value is within ± 1 unit (simple acceptance, no guard band)
- Measured Value: 0.5 units

Under simple acceptance, this result is a PASS. But now consider what happens when the measurement uncertainty is large, say, ± 0.8 units ($k = 2$). The true value could plausibly range from approximately -0.3 to $+1.3$ units. There is a meaningful probability that the true value lies outside the tolerance of ± 1 unit, even though the reported measurement was 0.5.

The UKAS LAB 48 [2] example **makes the critical point: ALL of the following measurement scenarios would also result in a PASS under simple acceptance**, regardless of the measurement uncertainty associated with each result:

Measured Value	Measurement Uncertainty ($k = 2$)	Simple Acceptance Result
0.5	± 0.1	PASS (PFA $\approx 0\%$ — low risk, appropriate)
0.5	± 0.8	PASS (PFA $\approx 11\%$ — questionable)
0.9	± 0.8	PASS (PFA $\approx 40\%$ — likely non-conforming)
0.99	± 0.8	PASS (PFA $\approx 49\%$ — probably non-conforming)

This demonstrates why simply ignoring measurement uncertainty in a decision rule is not merely a technical oversight; it is a failure to properly protect the end-user and can have serious consequences in safety-critical applications.

MEASUREMENT DISTRIBUTIONS VS. SPECIFICATION LIMITS

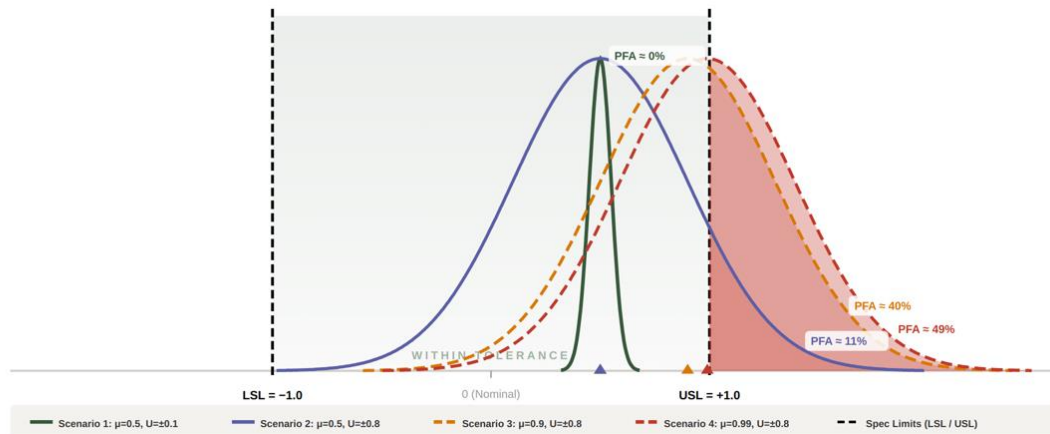


Figure 2 Specific Risk Curves Showing the Actual Risk

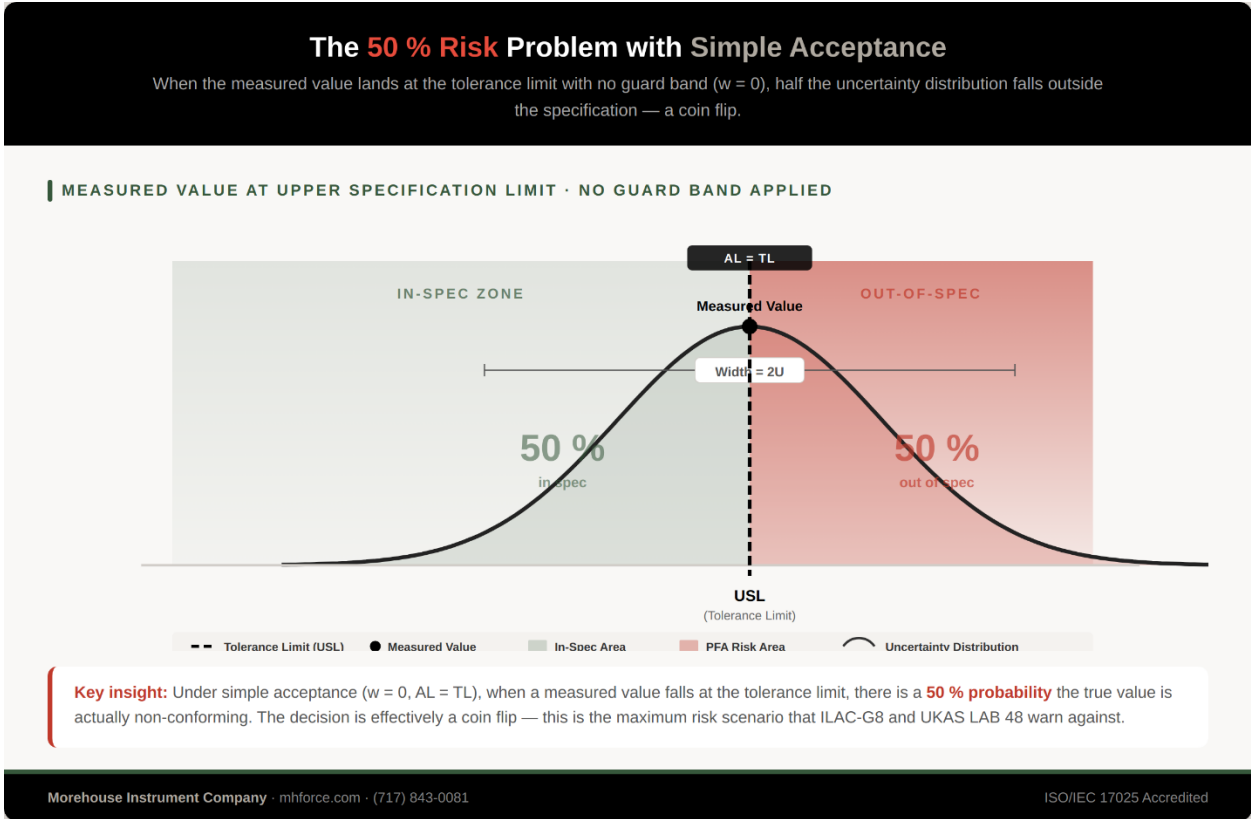


Figure 3 The 50 % Risk Problem

Under simple acceptance using a single probability distribution (specific risk, bench level), when a measured value falls exactly at the tolerance limit, the conditional probability that the true value is non-conforming can approach 50 %, effectively the same as a coin flip. This is a specific risk statement; it describes the risk associated with one individual conformity decision, given where the observed value landed relative to the tolerance boundary.

It is important to understand that specifying a Test Uncertainty Ratio (TUR) does not change this specific risk. A stated TUR is a global risk control, intended to limit overall risk across a population of measurements by ensuring adequate measurement capability. At any single measurement point near the tolerance limit, the conditional PFA under simple acceptance remains as high as the probability distribution allows, regardless of the TUR requirement written into the purchase order. The impact of TUR ratios using specific risk only changes the acceptance limits; high TUR equals more room to pass, or larger guard banded acceptance limits.

Therefore, specific risk governs the correctness of an individual pass/fail decision, while global risk governs long-term performance and overall exposure across a population. Both are legitimate frameworks; they answer fundamentally different questions. A laboratory that reports 2 % global PFA may still be issuing individual conformity statements with specific PFA approaching 50 % near the tolerance boundary.

Note: Global/Average (Unconditional) Risk and Specific (Conditional) Risk, also known as bench-level risk, are discussed in greater detail later in this document.

From Design to Verification: How Uncertainty Shrinks Your Usable Spec

At the design phase, the full specification range is available. During verification, measurement uncertainty creates zones of ambiguity at every tolerance boundary.

SPECIFICATION LIMITS VS. MEASUREMENT UNCERTAINTY ZONES

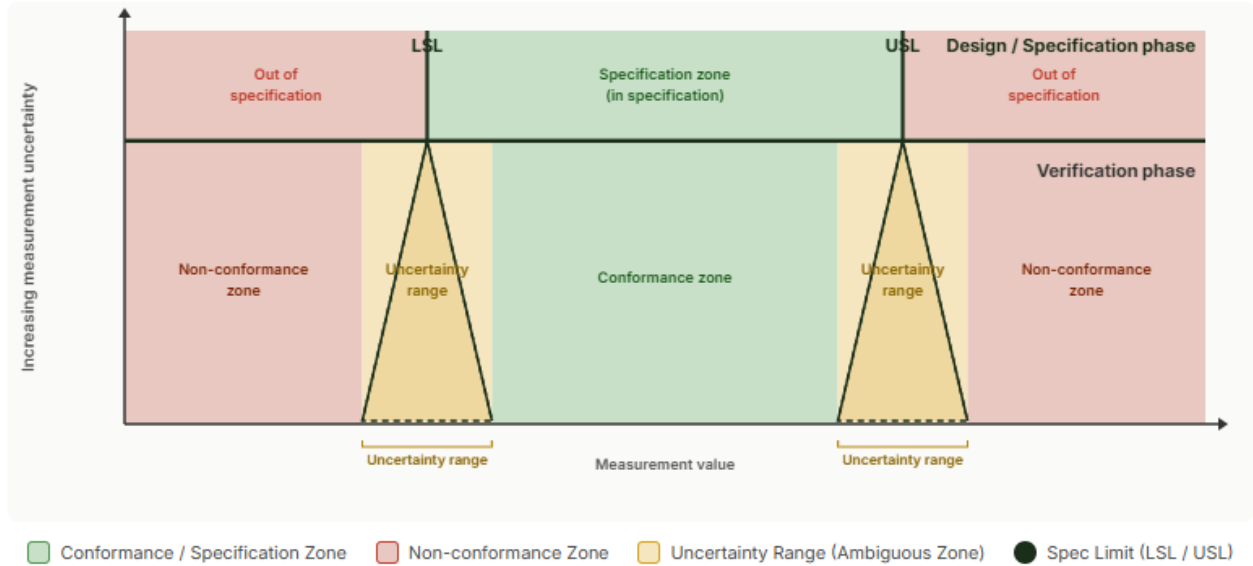


Figure 4 Setting Specifications

This graphic illustrates a fundamental problem in metrology: **what you specify on paper and what you can reliably verify in practice are not the same thing.**

The Design / Specification Phase (top band)

When an engineer designs a product or process, they define a tolerance, a Lower Specification Limit (LSL), and an Upper Specification Limit (USL). At this stage, the boundary is assumed to be a sharp, clean line. Something is either in spec or it isn't. There is no ambiguity. The world is simple, in theory as many might try to ask for something they do not need, or pad the tolerance from what they really need. Maybe this is from experience, and maybe they do not trust the measurement chain. Though we will assume it's simple.

The Verification Phase (bottom band)

The moment you pick up a measurement instrument to check whether something actually meets that specification, reality gets complicated. Your measurement tool has uncertainty; it doesn't tell you the exact true value, it gives you an estimate with a range of possible errors around it. That uncertainty is represented by the triangles: a single sharp spec limit at the top fans out into a fuzzy zone at the bottom.

The Uncertainty Range (amber zones)

This is the critical insight the graphic is communicating. Near each spec limit, there is a band of measurement values where you genuinely cannot tell, based on the measurement alone, whether the item is truly in spec or truly out of spec. If your measurement falls within this zone, the true value could lie on

either side of the limit. The wider your measurement uncertainty, the wider these amber zones become, and the more of your usable specification range gets consumed by ambiguity.

What does this mean practically?

You face three distinct regions in the verification phase, not two:

- **Non-conformance zone** — far enough outside the limit that even accounting for uncertainty, the item is almost certainly bad. Reject with confidence.
- **Uncertainty range** — close enough to the limit that you cannot reliably determine conformance. This is where false accepts (PFA) and false rejects (PFR) happen.
- **Conformance zone** — far enough inside the spec that even accounting for uncertainty, the item is almost certainly good. Accept with confidence.

The practical consequence

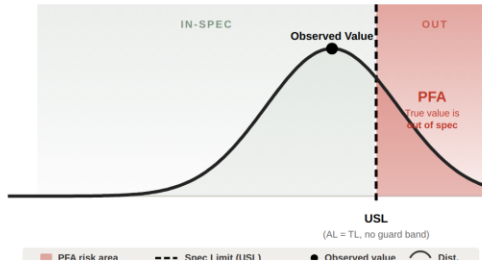
If your measurement uncertainty is large relative to your tolerance, the amber zones eat up a significant portion of your specification range. Your "safe" conformance zone shrinks. You're forced to either accept more risk (by ignoring the uncertainty) or tighten your effective acceptance limits, which means rejecting more borderline-good parts and increasing cost.

In most practical calibration programs, the most direct way to recover a usable specification range without increasing decision risk is to reduce measurement uncertainty at the source. That is precisely what a lower uncertainty calibration lab like Morehouse delivers, keeping those amber zones narrow so your full specification range remains usable, and your pass/fail decisions actually mean something.

Understanding PFA and PFR Risk Zones

Two distinct risks arise when measurement uncertainty overlaps a specification boundary. Each is best understood in its own scenario.

PFA — Probability of False Accept CONSUMER'S RISK



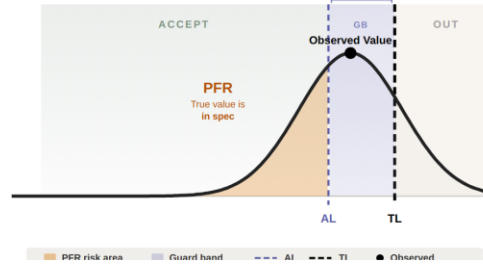
SIMPLE ACCEPTANCE · MEASURED NEAR USL

The true value is **outside** the specification, but the measured value falls inside the tolerance. The item is **accepted when it should be rejected**. The shaded red area is the probability this happens.

This risk is borne by the **customer / end-user** who receives a non-conforming item.

AL = TL · No guard band · Risk highest near limits

PFR — Probability of False Reject PRODUCER'S RISK



GUARD-BANDED · MEASURED JUST OUTSIDE AL

The true value is **inside** the specification, but the measured value falls outside the acceptance limit (in or beyond the guard band). The item is **rejected when it should be accepted**. The shaded amber area is the probability this happens.

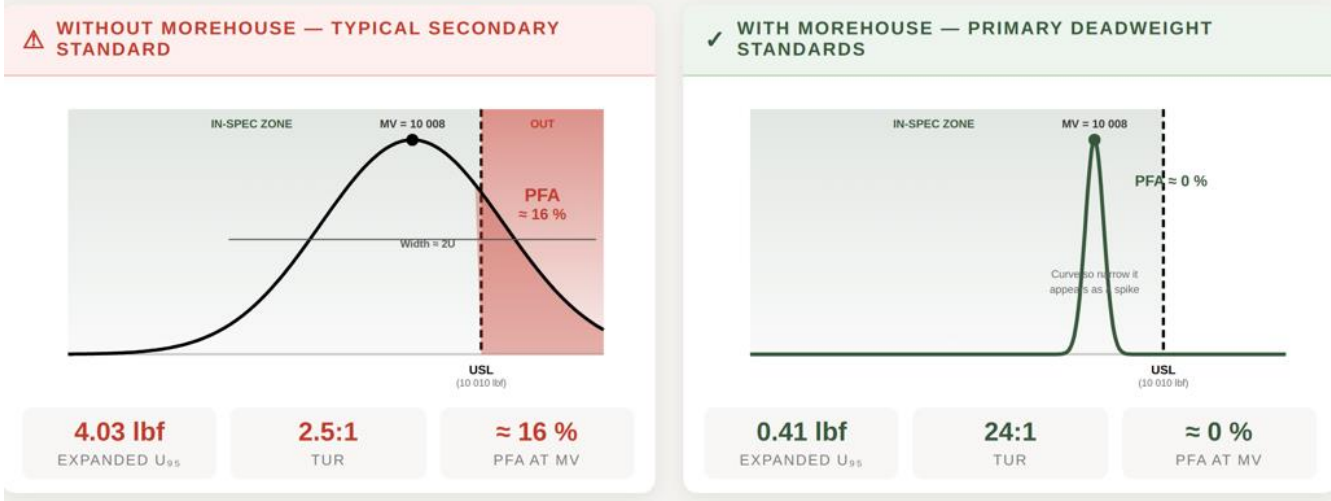
This risk is borne by the **calibration lab / producer** who incurs unnecessary rework or recall costs.

AL = TL - w · Guard band creates PFR trade-off

Figure 5 Understanding PFA and PFR

Same Instrument, Different Lab Uncertainty → Different Risk

Both charts show the same 10 000 lbf load cell with ± 0.1 % tolerance (±10 lbf), measured value = 10 008 lbf. Only the calibration source changes.



PARAMETER	TYPICAL LAB	MOREHOUSE	DIFFERENCE
Expanded U_{95} at 10 000 lbf	≈ 4.03 lbf	≈ 0.41 lbf	≈ 10× lower
CMC (lab reference)	4.00 lbf	0.16 lbf	25× lower
Repeatability	0.379 lbf	0.057 lbf	6.6× lower
TUR (±10 lbf tolerance)	2.5:1	24:1	10× higher
PFA at MV = 10 008 lbf	≈ 16 %	≈ 0 %	Far lower risk
Guard band (2.5 % specific risk)	3.95 lbf	0.40 lbf	10× smaller
Usable acceptance zone	±6.05 lbf (60 %)	±9.60 lbf (96 %)	+36 % recovered

Figure 6 TUR Ratios' Specific Risk Example

Picture the measurement result as a probability distribution, a bell curve centered on the observed value. The width of that bell curve is set by your measurement uncertainty. It represents what you don't know about where the true value actually lies.

Both charts show the same load cell, the same ± 0.1 % tolerance, and the same measured value of 10 008 N. The only thing that changes is the calibration source and with it, the uncertainty. This is where TUR is only indicative of the guard banded acceptance limits.

On the left, a typical secondary-standard calibration carries $U \approx 4.03$ lbf (0.040 % of applied force), giving a TUR of only 2.5:1. The bell curve is wide enough that a substantial portion of the distribution spills past the upper spec limit. That area under the curve beyond the USL is PFA: the probability that the true value is actually out of spec, even though the measured value appears to pass. Here it is approximately 16 %. A guard band of about 3.95 lbf would be needed just to pull PFA below 2.5 %, and that guard band consumes nearly 40 % of the tolerance, leaving a usable acceptance zone of only ±6.05 lbf. That comes at the cost of increased PFR, good units rejected unnecessarily. It is important to note that the end user may be okay with a PFA > 2 %. The decision rule and amount of risk should be agreed upon during contract review.

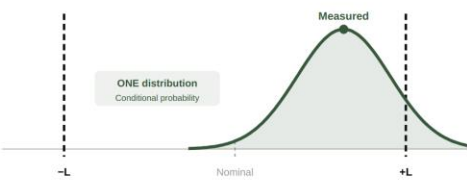
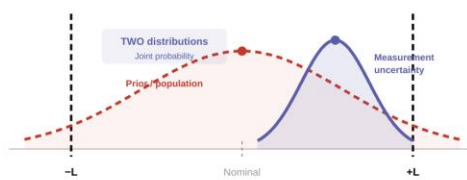
On the right, the same load cell calibrated against Morehouse primary deadweight standards carries $U \approx 0.41$ lbf (0.004 % of applied force), yielding a TUR of approximately 24:1. The bell curve is so narrow it appears as a spike well inside the tolerance zone. Virtually none of the distribution reaches USL. PFA drops to effectively zero. The guard band shrinks to approximately 0.40 lbf, leaving a usable acceptance zone of ± 9.60 lbf — nearly the full specification range. Decision confidence approaches 100 %.

The lesson is that **the lower uncertainty compresses the bell curve**, giving you a better chance of meeting your PFA requirements. Your pass/fail decisions start meaning what they say.

4. Global and specific risk

Specific vs. Global Risk in Conformity Assessment

Two distinct frameworks for evaluating measurement decision risk — one evaluates risk for a single measurement result, the other across a population of instruments.

Specific (Conditional) Risk BENCH-LEVEL	Global (Unconditional) Risk PROCESS-LEVEL
 <p style="margin-top: 10px;">"Given this measurement result, what's the probability this specific item is actually out of spec?"</p> <ul style="list-style-type: none"> ● Based on a single measurement result at the bench ● Uses conditional probability — risk depends on where the observed value fell ● Characterized by one probability distribution (the measurement uncertainty) ● Appropriate when no prior population data exists ("meagre prior information" per JCGM 106) 	 <p style="margin-top: 10px;">"Across all instruments we calibrate, what percentage of truly bad ones do we accept?"</p> <ul style="list-style-type: none"> ● Based on a population of instruments and future measurement results ● Uses joint probability — combines population behavior with measurement uncertainty ● Characterized by two probability distributions (population + measurement process) ● Requires historical reliability data (EOPR, interval analysis) to model the prior distribution
<p style="color: #c00000; font-weight: bold;">⚠ Critical distinction (per ILAC-G8, Section 5.3): An instrument that passes a 2 % global PFA requirement can still present specific risk for false acceptance approaching 50 % on individual measurements near the tolerance boundary. Global risk is a population average — it masks what happens at the edges.</p>	

Morehouse Instrument Company • mhforce.com • (717) 843-0081
ISO/IEC 17025 Accredited

Figure 7 Specific Versus Global Risk

Global (Unconditional) Risk

Global PFA asks: *"Across all the instruments we calibrate, what percentage of truly bad ones do we accept?"*

This is an average across your entire population of instruments and calibration events. It depends on two things: your historical reliability data (how often do instruments actually drift out of spec?) and the uncertainty of your calibration process. You must know how your instruments behave over time to properly calculate global risk.

In programs that contractually invoke ANSI/NCSL Z540.3 [5], a common target is controlling consumer's risk (PFA) for acceptance decisions to $\leq 2\%$ under the specified decision rule. This is typically framed as an unconditional, population-level requirement. Achieving that target generally requires not only a measurement uncertainty model but also a defensible model for the population of items being calibrated, meaning prior information supported by reliability and interval data such as end-of-period reliability (EOPR) or similar historical evidence.

Where such data are limited, some Z540.3 [5] handbook approaches use managed guard banding to control reported risk under stated assumptions, at the cost of increased false rejects, without claiming full population characterization.

The catch: an instrument that passes a 2% *global* PFA requirement can still present much higher *specific* risk on individual measurements near the tolerance boundary. The global number is an average; it masks what's happening at the edges. ILAC-G8:09/2019 [3] makes clear that an instrument passing a global risk criterion of 2% PFA may have a specific risk for false acceptance that can be as high as close to 50% as shown in the previous section.

Specific (Conditional) Risk

Specific PFA asks: "Given this particular measurement result, and given my uncertainty, what's the probability this specific instrument is actually out of spec?"

This is a conditional probability; it depends on where the observed value landed relative to the tolerance limits. A result right in the middle of the tolerance range? Very low specific PFA. A result right at the edge of the acceptance limit? The specific PFA can be substantial, even if the global number looks fine.

Specific risk is the right framework when you don't have prior population data, what JCGM 106:2012 [4] (clause 7.2.2) calls 'meagre prior information,' or when a more cautious approach to false accepts is required, regardless of available data.

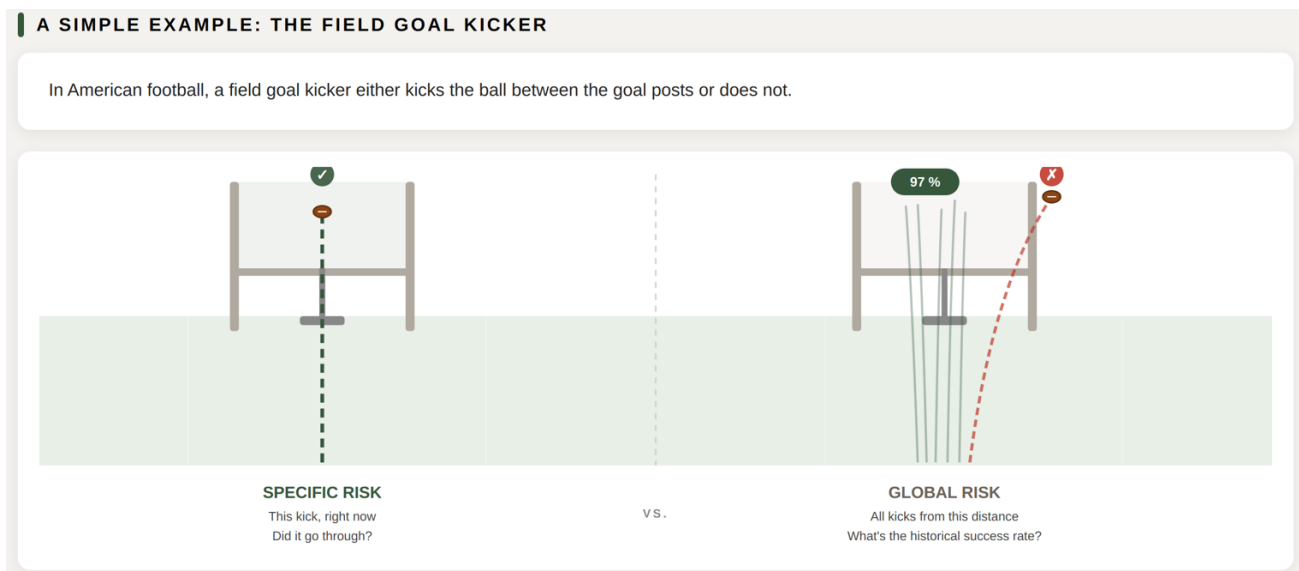


Figure 8 Field Goal Example

Specific (Conditional) Vs Global (Unconditional) Risk, A Simple Example: The Field Goal Kicker

In American football, a field goal kicker either kicks the ball between the goal posts or does not.

Specific risk is the single kick in front of you right now. The ball either traveled between the goal posts or it didn't. Pass and points or fail and no points for your team (sorry, Ravens fans). When we make a measurement and record the observed value, it is either between the acceptance limits or it is not. One kick, one result, one decision.

Global (average) risk steps back and looks at all the kicks. Say the Steelers kicker makes 97 % of his field goals from the 35-yard line. This kick is from the 35-yard line, yet it didn't go in. Based on population data alone, you would say it is likely that it did go in. And the Steelers won the game, hooray!

That is exactly the tension between the two approaches. Global risk uses historical population data to assess the probability of an incorrect decision, such as a false accept or false reject, across a population. Specific risk evaluates the single result in front of you. One tells you how reliable the process is over time. The other tells you what happened right now.

Now shift your focus from the kick to the goal posts themselves. The distance between them represents your tolerance. A kicker with a tight, repeatable trajectory effectively has more room to succeed because their variability is low. In contrast, a kicker with inconsistent performance needs nearly perfect alignment every time. The same principle applies to measurement: as uncertainty decreases, the "usable" portion of the tolerance increases, and confidence in the result improves.

With that perspective in mind, decision rules become the method for determining whether a result truly falls within those goal posts. Most practical decision rules stem from three foundational approaches, each offering a different way to balance risk and confidence.

5. The three core decision rule approaches

Most practical decision rules are derived from three foundational approaches. Understanding each helps a laboratory select the most appropriate one for its application.

APPROACH 01 · HIGHEST RISK

Simple Acceptance No guard band · AL = TL · uncertainty ignored in decision

Pass if the measured value falls within the tolerance limits. Measurement uncertainty plays **no role** in the acceptance/rejection decision. With the observed value near the USL, a large portion of the uncertainty distribution extends beyond the limit — yet the result is still declared a PASS.



PFA RISK

High — **uncontrolled near limits**. When the measured value is near the tolerance boundary, the PFA can approach 50%. Only acceptable when TUR is very high ($\geq 10:1$).

AL = TL · NO GUARD BAND

WHEN TO USE

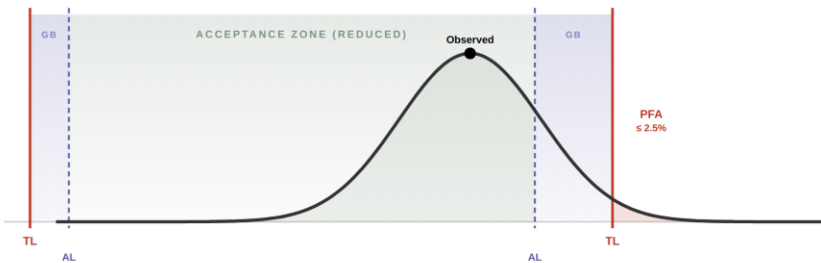
When TUR is very high ($\geq 10:1$), making PFA negligible. Or when explicitly requested by the customer with documented understanding of the risk. Must still document the decision rule and associated risk per ISO/IEC 17025:2017.

Figure 9 Simple Acceptance

APPROACH 02 · CONTROLLED RISK PER POINT

Specific Risk Fixed guard band · PFA $\leq 2.5\%$ per measurement

A fixed guard band is applied inward from each tolerance limit to create a tighter acceptance limit (AL). The guard band is sized so that PFA at any single measurement point does not exceed a defined threshold — typically 2.5% per side.



PFA RISK

Controlled — $\leq 2.5\%$ per side. The fixed guard band ensures that at every measurement point the probability of false acceptance is bounded. This is point-by-point risk control.

AL = TL \pm (GB Multiplier \times U)
For $\leq 2.5\%$ risk per side: GB Multiplier ≈ 0.980

KEY CHARACTERISTICS

AL < TL — acceptance limits are tighter than tolerance. The guard band is **fixed as a function of U**. This provides point-by-point control of consumer's risk, at the cost of increased producer's risk (more false rejects of borderline-good items).

Figure 10 Specific Risk

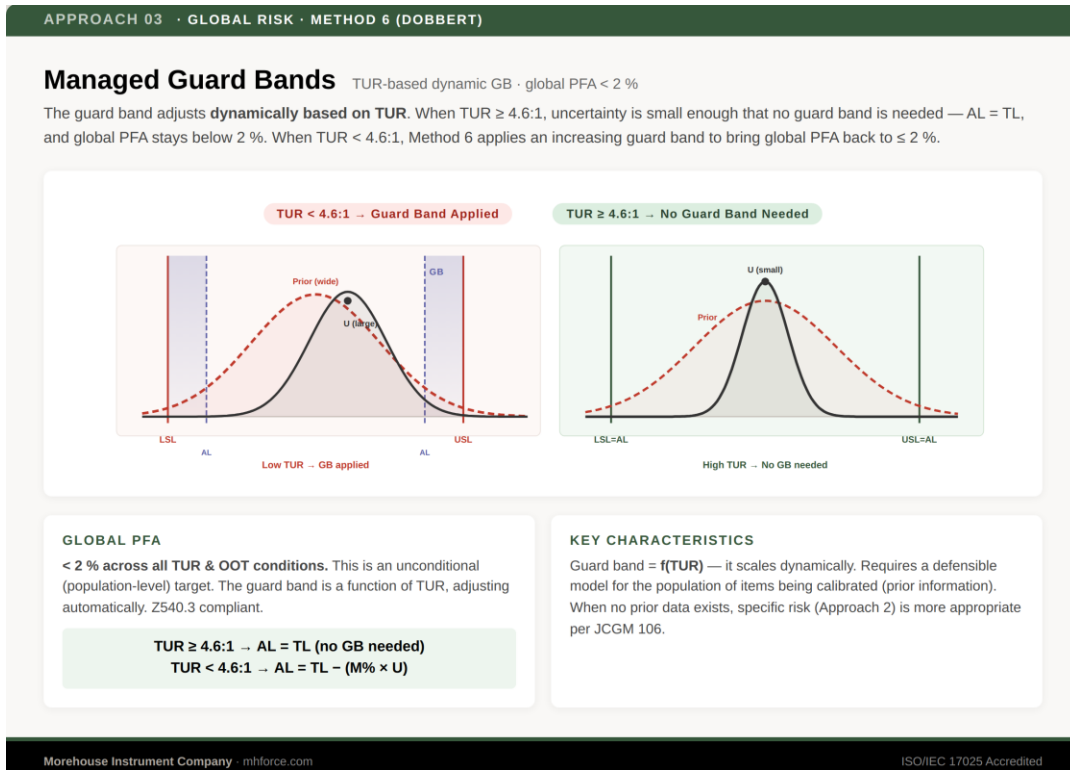


Figure 11 Global Risk: Managed Guard Bands

5.1 Simple acceptance (no guard band)

A result is passed if the measured value falls within the tolerance limits. Measurement uncertainty plays no role in the acceptance/rejection decision itself. This is the highest-risk approach, particularly when the TUR is low (i.e., measurement uncertainty is large relative to tolerance).

When might Simple Acceptance be used?

Simple acceptance may be appropriate when (a) the TUR is very high (e.g., greater than 4.6:1), making PFA negligible, (b) it is explicitly requested by the customer with documented understanding of the risk, or (c) the consequences of a false accept are minimal. The laboratory must still document the decision rule and the associated risk.

ILAC-G8 [3] Requirement: ILAC-G8:09/2019 [3] (definition 1.9) defines simple acceptance as $AL = TL$. The broader guidance in the document indicates that, without a precondition constraint on measurement uncertainty, this approach alone may not satisfy the Clause 3.7 requirement to account for measurement uncertainty in the decision rule. To use simple acceptance as a compliant decision rule, it should be paired with a precondition constraint on measurement uncertainty, either by specifying a minimum TUR (e.g., $TUR > 4:1$) or by setting an upper limit on allowable expanded measurement uncertainty. Without such a constraint, the laboratory cannot demonstrate that it has accounted for measurement uncertainty in the decision rule.

5.2 Specific risk (guard-banded acceptance)

A guard band is applied to the tolerance limit to create an acceptance limit that is tighter than the specification. The guard band is calculated so that the probability of false acceptance at any given measurement point does not exceed a defined threshold (commonly 2.5 % per side for a 95 % confidence requirement).

The guard band is calculated as: Acceptance Limit = Tolerance Limit \pm (Guard band Multiplier \times Expanded Measurement Uncertainty)

For a 2.5 % maximum specific risk on a two-sided tolerance, the Guard band Multiplier is approximately 0.980, calculated using $\text{NORM.S.INV}(0.975)/2$ in Excel.

5.3 Global risk - managed guard bands (Method 6 – Dobbert [11])

Developed by Michael Dobbert [11], this method aims to achieve a global unconditional PFA of less than 2 % across all combinations of TUR and out-of-tolerance rates. It uses a dynamic guard band that adjusts based on TUR and is particularly useful for meeting ANSI/NCCL Z540.3 [5] requirements. When the TUR is 4.6:1 or greater, no guard band is required under Method 6, the measurement uncertainty is sufficiently small that the global PFA remains below 2 % with AL = TL.

5.4 Non-binary decision rules (conditional pass/fail)

ILAC-G8:09/2019 [3] describes non-binary decision rules as an alternative to the binary pass/fail approach. While the standard presents a four-outcome model (Pass, Conditional Pass, Conditional Fail, Fail), a practical three-outcome approach, as illustrated in UKAS LAB 48 [2], may be more useful in many calibration applications.

The three-outcome model uses:

Pass — The measured value falls within the acceptance limits. The Specific (Bench Level) PFA is at or below the agreed threshold (e.g., $\leq 2\%$). Conformity is stated with confidence.

Possible Pass — The measured value is within the tolerance limits but outside the acceptance limits. The specific PFA exceeds the desired threshold but has not reached 50 %. Conformity cannot be assured at the desired confidence level, but the item is not definitively non-conforming. The customer receives the measurement result, the uncertainty, and the calculated PFA to make an informed decision.

Fail — The measured value produces a specific PFA greater than 50 %, or the measured value falls outside the tolerance limits by more than the expanded uncertainty. Non-conformity is stated.

This three-outcome approach is particularly useful when applying specific risk, and a maximum PFA has not been specified by the customer. Rather than forcing a binary decision in a region of genuine ambiguity, it gives the customer actionable information about where the measurement landed relative to the risk boundary. The "Possible Pass" category avoids both the false precision of declaring conformity and the unnecessary cost of declaring failure when the evidence is inconclusive.

Note: Conformity with a requirement is inherently connected to the decision rule employed; therefore, it is expected that the decision rule is agreed upon before measurements are taken (ISO/IEC 17025:2017 [1], Clause 7.1.3).

6. How to calculate and apply guard bands for Specific Risk

6.1 Two-sided tolerance example

Objective: Limit specific risk to no more than 2.5 % per side (97.5 % conformance confidence).

- Tolerance: ± 1 unit
- Standard Measurement Uncertainty ($k = 1$): 0.125 units
- Expanded Measurement Uncertainty ($k = 2$): 0.250 units
- Guard band Multiplier (from $\text{NORM.S.INV}(0.975)/2$ in Excel): 0.980

Calculation:

1. Upper Acceptance Limit = $1 - (0.980 \times 0.250) = 1 - 0.245 = 0.755$ units
2. Lower Acceptance Limit = $-1 + (0.980 \times 0.250) = -1 + 0.245 = -0.755$ units

Any measured value between -0.755 and $+0.755$ is accepted. Values outside this range but within the ± 1 unit tolerance are rejected (even though they might technically meet the specification) to maintain the required confidence level.

6.2 One-sided tolerance example

For a one-sided tolerance with 5 % maximum risk:

- Guard band Multiplier = $\text{NORM.S.INV}(0.95)$ in Excel = 1.645
- Acceptance Limit = Tolerance Limit $- (1.645 \times \text{Expanded Uncertainty})$

Excel Note

The formula always simplifies to: Acceptance Limit = Tolerance Limit \pm Guard band Multiplier \times Expanded Measurement Uncertainty. The sign is always applied such that the acceptance limit moves inward from the tolerance limit.

6.3 Guard bands: the balancing mechanism

So how do labs manage the PFA/PFR trade-off in practice? With guard bands.

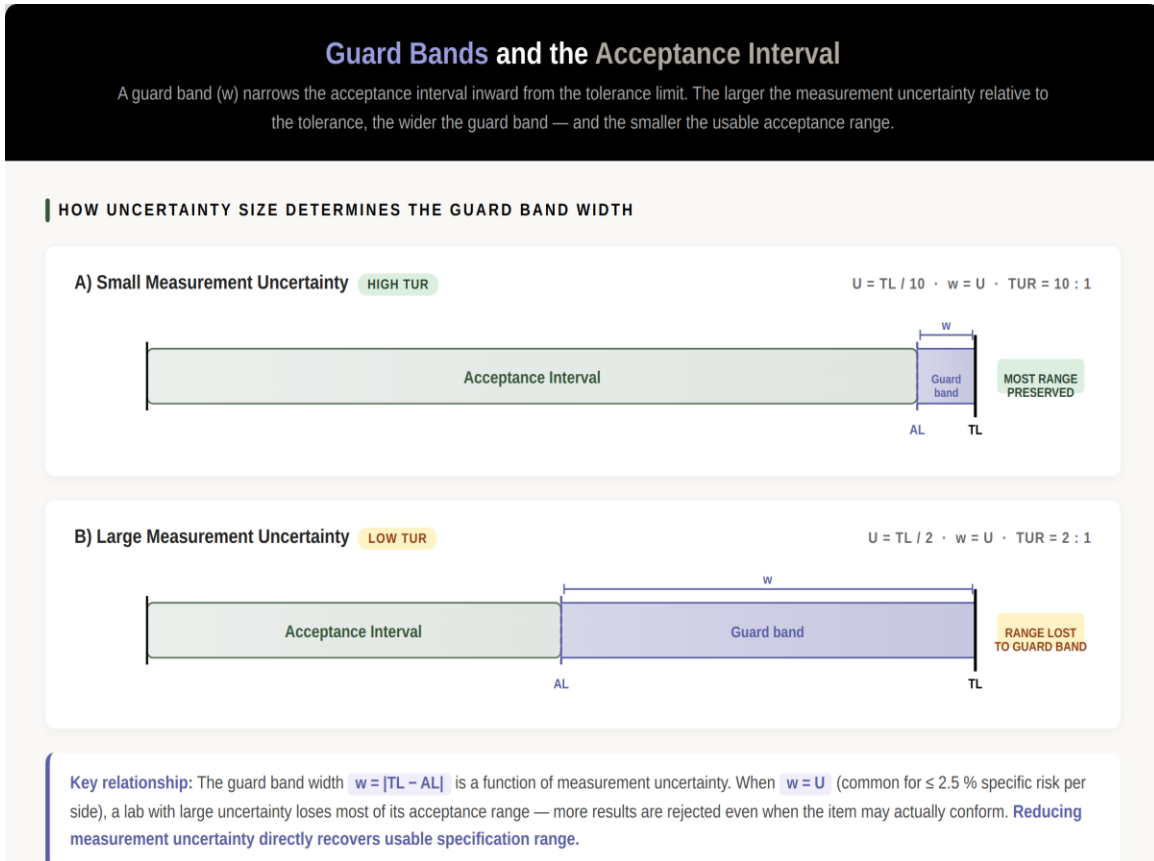


Figure 12 Acceptance Limits (Intervals)

To review, a guard band is a buffer zone inside the specification limit. Instead of accepting any result that falls within the full spec range (LSL to USL), you only accept results that fall within a tighter acceptance limit (AL). The gap between the AL and the spec limit is the guard band.

Guard bands deliberately trade some PFR (you'll reject more borderline-good instruments) in order to reduce PFA (you'll accept fewer borderline-bad ones). The width of the guard band is set based on the measurement uncertainty and the desired maximum risk level.

7. What must appear on a calibration certificate

Per ISO/IEC 17025:2017 [1], Clause 7.8.6.2, every calibration certificate that includes a statement of conformity must clearly identify:

- To which results the statement of conformity applies.
- Which specifications, standards, or parts thereof are met or not met.
- The decision rule applied (unless inherent in the requested specification or standard).

In practice, this means the certificate should state something equivalent to: "Pass/fail determination was made using [specific decision rule], with a guard band of [value] applied to the tolerance limits of [specification]. Measurement uncertainty is [U] at a 95 % confidence level ($k = 2$). Any result within [AL lower] to [AL upper] was accepted."

⚠ Common Deficiency

Many calibration certificates simply state "PASS" or "IN TOLERANCE" without identifying the decision rule applied, what guard band (if any) was used, or acknowledging the role of measurement uncertainty. This is non-compliant with ISO/IEC 17025:2017 [1], Clause 7.8.6.2, and may also be non-compliant with accreditation body requirements.

8. Frequently asked questions

Q1: Our customer says they just want a simple PASS/FAIL and don't care about uncertainty. Can we do that?

You can use simple acceptance if the customer explicitly requests it and understands the associated risk – but you must still document the decision rule on the calibration certificate (ISO/IEC 17025:2017 [1], Clause 7.8.6.2) and ensure the decision rule is "communicated to, and agreed with, the customer" (Clause 7.1.3).

The critical point from UKAS LAB 48 [2]: it is not acceptable to claim you are 'taking account of uncertainty' by simply ignoring it. A decision rule that ignores uncertainty must be explicitly identified as such, and both parties must understand what that means for the reliability of the conformity statement. ILAC-G8:09/2019 [3] further indicates that simple acceptance without a precondition constraint on measurement uncertainty (such as a minimum TUR requirement or an upper limit on allowable U) may not satisfy the Clause 3.7 requirement to account for measurement uncertainty.

Q2: Does UKAS LAB 48 [2] prescribe a specific decision rule that must be used?

No. UKAS LAB 48 [2] is a guidance document that explains decision rules and their implications. It does not mandate a specific rule. However, it does set out clear expectations for how laboratories should document, communicate, and apply whatever decision rule they use. Accreditation bodies may have additional requirements through their policies.

Q3: What is the difference between global risk and specific risk?

Specific risk is assessed at each individual measurement point: what is the probability that this particular measurement is a false accept? Global (or unconditional) risk considers the overall population of measurements and their distribution relative to tolerance. Methods like Dobbert's Method 6 [11] are designed to control global risk across the full distribution of likely measurement values, whereas specific risk guard bands control the risk at any single measurement outcome.

Q4: We are a calibration laboratory calibrating other calibration labs' reference standards. Are we required to apply a decision rule?

Only if you provide a statement of conformity (i.e., a pass/fail or in-tolerance/out-of-tolerance determination). If your calibration certificate reports only the measured values and uncertainties without a conformity statement, no decision rule is required. However, if any conformity statement is made, even informally, the requirements of Clauses 7.1.3, 7.8.6.1, and 7.8.6.2 [1] apply.

Q5: Our measurement uncertainty is larger than we would like. Can we still issue a statement of conformity?

Yes, though the guard band will consume more of the specification range, resulting in a narrower acceptance zone. This means more results may be rejected even though they fall within the tolerance limits. This is the correct outcome: when uncertainty is large, the lab has less ability to confidently distinguish conforming from non-conforming items. Investing in lower-uncertainty measurement capability directly expands the usable acceptance zone and reduces both false accept and false reject rates.

Q6: What happens if the expanded uncertainty exceeds the tolerance?

If the expanded uncertainty U equals or exceeds the tolerance TL , a guard band calculated to limit specific risk would equal or exceed the full tolerance – meaning no measurement value could be accepted. This is an extreme case that indicates the measurement capability is fundamentally inadequate for the specification being tested. The laboratory should either improve its measurement capability or decline to issue a conformity statement, and instead report only the measured value with uncertainty.

9. Key reference documents

The following documents form the normative and guidance framework for decision rules and statements of conformity:

Document	Summary
[1] ISO/IEC 17025:2017	The primary standard for calibration and testing laboratory competence. Clauses 3.7, 7.1.3, 7.8.6.1, and 7.8.6.2 govern decision rules and conformity statements.
[2] UKAS LAB 48	"Decision Rules and Statements of Conformity" – A practical guidance document from the UK Accreditation Service with worked examples. Highly recommended reading for all calibration laboratories. Edition 5, 2024.
[3] ILAC-G8:09/2019	"Guidelines on Decision Rules and Statements of Conformity" from the International Laboratory Accreditation Cooperation. Covers guard band calculation, PFA, and reporting requirements.
[4] JCGM 106:2012	"Evaluation of measurement data – The role of measurement uncertainty in conformity assessment." The statistical foundation document for understanding risk in conformity decisions.
[5] ANSI/NCSL Z540.3 Handbook	Handbook for the Application of ANSI/NCSL Z540.3-2006. Covers the 2 % rule for PFA and practical methods for achieving it.
[6] ASME B89.7.3.1-2001	"Guidelines for Decision Rules: Considering Measurement Uncertainty in Determining Conformance to Specifications."
[7] ASME B89.7.4.1-2005	"Measurement Uncertainty and Conformance Testing: Risk Analysis."
[8] ISO 14253-1:2017	"Decision rules for verifying conformity or nonconformity with specifications."
[9] NCSLI RP-18	"Estimation and Evaluation of Measurement Decision Risk."
[10] JCGM 100:2008 (GUM)	"Evaluation of measurement data — Guide to the Expression of Uncertainty in Measurement." The foundational document for expanded uncertainty (U), coverage factor (k), and combined standard uncertainty.
[11] Dobbert, M. (2008)	"A Guard-Band Strategy for Managing False-Accept Risk." Proc., NCSL Workshop & Symposium, 2008. The source for Method 6 managed guard bands referenced in Section 5.3.
[12] Decision Rule Guidance V1.4	An in-depth guidance document on decision rules written by Henry Zumbrun, Greg Cenker, and Dilip Shah. Highly accessible treatment of specific and global risk with worked examples.

10. Quick reference checklist

Use this checklist to verify compliance before issuing any calibration certificate that includes a statement of conformity:

#	Checklist Item
1	Have you determined the customer's requirements for a conformity statement (pass/fail, in-tolerance, etc.)?
2	Have you selected and documented a decision rule that accounts for measurement uncertainty?
3	Have you communicated the decision rule to the customer and obtained agreement (when required by Clause 7.1.3)?
4	Have you evaluated and documented the measurement uncertainty for the calibration?
5	Have you calculated the TUR and assessed whether the measurement capability is adequate for the specification?
6	Have you calculated any required guard bands and applied them consistently?
7	Does the calibration certificate clearly identify: (a) which results the conformity statement applies to, (b) which specifications are met/not met, and (c) the decision rule applied?
8	Is the expanded uncertainty reported on the calibration certificate?
9	Have you retained records of the decision rule documentation and customer agreement?
10	If simple acceptance is used, have you documented the associated PFA risk and communicated it to the customer?

11. Conclusion

Decision rules are not a bureaucratic formality; they are the mechanism by which a calibration laboratory honestly communicates the reliability of its conformity determinations to its customers. When measurement uncertainty is not properly accounted for, pass/fail statements can be misleading and, in safety-critical applications, dangerous.

The core obligations under ISO/IEC 17025:2017 [1] are clear: define the decision rule, apply it consistently, communicate it to the customer, and report it on the calibration certificate. UKAS LAB 48 [2] and ILAC-G8:09/2019 [3] provide excellent worked examples to help laboratories understand how to do this in practice.

The investment in understanding and properly implementing decision rules can drastically increase margins: it reduces the risk of costly false accepts reaching the field, demonstrates technical competence to customers and assessors, and supports a culture of measurement integrity throughout the calibration chain.