FROM HEALTH SCORES TO ACTION:

A Practical Guide to Aligning Asset Health Indices and Condition Monitoring for Transformer Fleets

Clarifying purpose, governance, and alarm response pathways (with industry references)

G. Matthew Kennedy & Tony McGrail

ABOUT THIS eBOOK

This ebook is a practical guide for utility asset managers, reliability engineers, and maintenance leaders who use Asset Health Indices (AHIs) produced by Asset Performance Monitoring Systems (APMS) and transformer condition monitoring (CM) performed by condition monitors. It addresses some common misconceptions: treating an AHI like an operational trigger, or treating condition monitor alarms like "predictive maintenance," without a clear decision pathway from indication to accountable action.

How to use this guide

- If you are debating what your AHI should (and should not) represent, start with **Chapter 2** and **Chapter 4**.
- If you are struggling with alarms that do not result in timely action, go directly to **Chapter 5** and **Tool 3 (Decision plan)** in **Chapter 7.**
- If you are building (or refreshing) governance, read straight through and use the checklists in **Chapter 6**.

Chapter 1. The Problem in Plain Terms

Utilities increasingly rely on two tools to manage transformer fleets:

- **Asset Health Indices (AHIs)** summarize condition information to support long-horizon strategic decisions (e.g., repair/replace prioritization, deferral, lifecycle planning). They are often the output of Asset Performance Monitoring (APM) Systems.
- **Condition monitoring (CM) via condition monitors** detects abnormal behavior and emergent faults that may require rapid review and intervention, producing alerts/alarms intended to trigger a known response pathway.
- **The recurring problem:** AHIs are treated like intervention triggers, while CM alerts/alarms (and their related analytics) are treated like "predictive maintenance," and neither has a clear, owned pathway from indication to action.

KEY TAKEAWAY: Your organization may not need "more numbers" as much as it may need clarity on *what decision* each piece of data supports and *who* is responsible for acting when evidence indicates emerging or increasing risk.

Chapter 2. What Condition Monitoring Is (and Isn't)

Condition monitoring (CM) programs use condition monitors to (i) detect abnormal behavior that may require operational intervention and (ii) provide early warning of faults that develop despite time-based and predictive maintenance programs. Their value is greatest when they reduce time-to-awareness and give the organization enough lead time to verify, escalate, and act—i.e., when they expand the actionable interval between potential failure and functional failure (the "P–F interval") (Skog and Klempner, 2022).

Common misconception: "CM = predictive maintenance." A CM alarm that triggers offline verification followed by an unplanned repair is typically detection and response to an emergent failure mechanism—not a prediction-optimized maintenance plan. Some response can be scheduled to slowly developing degradation, but the most important CM outcomes often come from catching unexpected or fast-developing problems early. There is nothing predictive about preventing a transformer from failing due to being informed of a degrading state within the transformer.

Maintenance strategy evolves over the asset lifecycle. Maintenance planning for transformers cannot be static because dominant failure mechanisms, risk tolerance, and economic trade-offs change as assets age. In early life, issues are often dominated by commissioning-related defects, installation quality, and latent manufacturing or integration problems; the highest-value actions tend to be corrective: targeted adjustments, workmanship fixes, and focused investigation of abnormal indications. In mid-life, failure rates often stabilize and the emphasis shifts toward condition-informed inspection, monitoring, and selective intervention—aiming to sustain reliability while avoiding unnecessary intrusive work. In later life, age-related mechanisms may become more prominent, tolerance for unresolved defects decreases, and maintenance increasingly focuses on preventive actions, risk mitigation, and life-extension decisions.

Maintenance programs often “feel” different at different ages for good reason. Kim et al. (2020) reinforce a useful intuition: failure behavior may or may not follow a bathtub-curve pattern—early-life issues can be driven by commissioning/installation and latent defects, mid-life can be comparatively stable, and late-life risk can increase as aging mechanisms dominate. The practical takeaway is simple: don’t force one fixed maintenance approach across the entire fleet; calibrate your monitoring intensity, triggers, and intervention philosophy to each life-cycle stage for each asset (Kim et al., 2020).

Component quality can redefine what “predictive” means. In practice, the perceived need for “predictive maintenance” is often driven by the trustworthiness of specific component populations (notably bushings and OLTCs). Complicating matters, true component quality is not always known until issues surface in service, after installation, or across a population. Vendor advisories and industry forums can rapidly change risk perception by surfacing common-mode concerns (e.g., a specific bushing type or design family). In these cases, the program challenge is less about predicting degradation from first principles and more about governance: establishing enhanced monitoring, defining triggers for engineering review, and planning for staged replacement when confidence in a component population is low. (For practical implementation, see **Tool 2** and the **component-population risk lens** within **Tool 5** in Chapter 7.)

Quick checklist: Managing a high-concern component population

- **Define the population:** What units/serial ranges/locations are in scope?
- **Clarify the concern:** What failure mode is suspected, and what evidence exists (advisory, incidents, test data)?
- **Enhance monitoring:** What additional measurements, sampling frequency, or analytics are warranted (IEEE C57.143-2024)?
- **Set triggers:** What thresholds or patterns force engineering review and what is the required response time?
- **Decide the mitigation path:** Increased surveillance vs. staged replacement vs. immediate replacement for critical sites, all agreed and documented.
- **Document decisions:** Record rationale, risk posture, and next review date so knowledge survives staffing changes.

Chapter 3. What the Industry Already Warns Us About

The guidance is remarkably consistent across standards, technical brochures, and field experience:

- **Composite indices can mask urgent problems.** CIGRE TB 761 warns that averaging/weighting can hide critical failure modes (e.g., one very poor component inside an otherwise "healthy" transformer) and explicitly notes that transformer assessment indices are not intended for alarm management (CIGRE TB 761, 2019). Recent work proposes calculation methods specifically aimed at preventing masking (MOH'D et al., 2024).
- **Turning an AHI into PoF is non-trivial.** CEATI guidance highlights pitfalls such as non-monotonic indices, urgency dilution through weighting, and missing time calibration—problems that become severe when indices are used outside their intended purpose (McGrail and Kennedy, CEATI 30113, 2013/2018).
- **Monitoring only works when alarms are routed and owned.** CIGRE Technical Brochure 962 emphasizes that monitoring benefits depend on getting key signals (e.g., online DGA) into operational pathways (often via SCADA-type systems) and pairing them with defined and agreed alarm response (CIGRE TB 962, 2025). IEEE C57.143 provides a structured guide to applying monitoring equipment and interpreting parameters for liquid-immersed transformers and components (IEEE C57.143-2024).
- **Alarm response planning is a discipline.** Industry discussion calls for alarm response to be planned, agreed, risk-analyzed, documented, and auditable before alarms occur (McGrail and Kennedy, 2025).

Chapter 4. Why AHI Programs Fail in Practice

Where it goes wrong. AHIs are often interpreted as maintenance triggers: a single number that is expected to tell teams when to "do maintenance." But most AHIs aggregate condition information across multiple subsystems (tank, windings, bushings, DETC, OLTC, oil, etc.). That aggregation makes the score useful for screening and *prioritization*, but it can make the score ambiguous for *specific work planning*.

What gets lost. When an organization compresses multiple signals into one value, it risks losing the "why." Weighting/averaging can hide the worst condition (masking), and some index designs are not monotonic—meaning the score can move in unintuitive ways even when engineering intuition says risk is rising (CIGRE TB 761, 2019; MOH'D et al., 2024; McGrail and Kennedy, CEATI 30113, 2013/2018). These effects are manageable when the AHI is used as a planning aid, but they become hazardous when teams treat the AHI as an operational trigger.

Chapter 5. When Scores and Alarms Collide

AHIs are attractive because they appear to enable fleet screening and leadership reporting. But when decision-makers see only a composite score, the organization can miss (or delay) specific, actionable interventions—especially when the risk is driven by a replaceable component such as a bushing or surge arrester.

Mini example (the masking trap). A transformer can look "acceptable" in an overall index while one component is in poor condition—TB 761 explicitly cautions that aggregation can mask worst-case conditions and that urgent findings should be addressed outside the index output (CIGRE TB 761, 2019). If the organization lacks a pathway to elevate the component-level story to decision-makers, it may miss an opportunity for a targeted replacement that materially reduces risk.

Common symptoms you'll recognize

- Alarms arrive, but there is no clear owner to verify, escalate and act on agreed response plans.
- Leadership dashboards show a single health number, but the "why" is buried or unavailable.
- Teams debate whether an issue is "real" because the index is fine while monitoring is concerning (or vice versa).
- Online monitoring data exists, but condition monitor alarms are not routed into operational systems where they are seen and acted upon (CIGRE TB 962, 2025; IEEE C57.143-2024).

Chapter 6. The Three Questions to Answer Before You Change Anything

1. **When should we use an overall AHI vs. explicit (component-level) AHIs vs. maintenance indices?**
2. **How should those metrics interact with condition monitor alarms intended to prompt intervention?**
3. **What governance, processes, and system architecture do we need so that indices and alarms lead to timely, accountable action?**

What this guide does. It provides a practical governance and decision-pathway framework grounded in industry guidance (CIGRE, IEEE, CEATI) and common program failure modes (masking, unclear escalation ownership, and alarm fatigue).

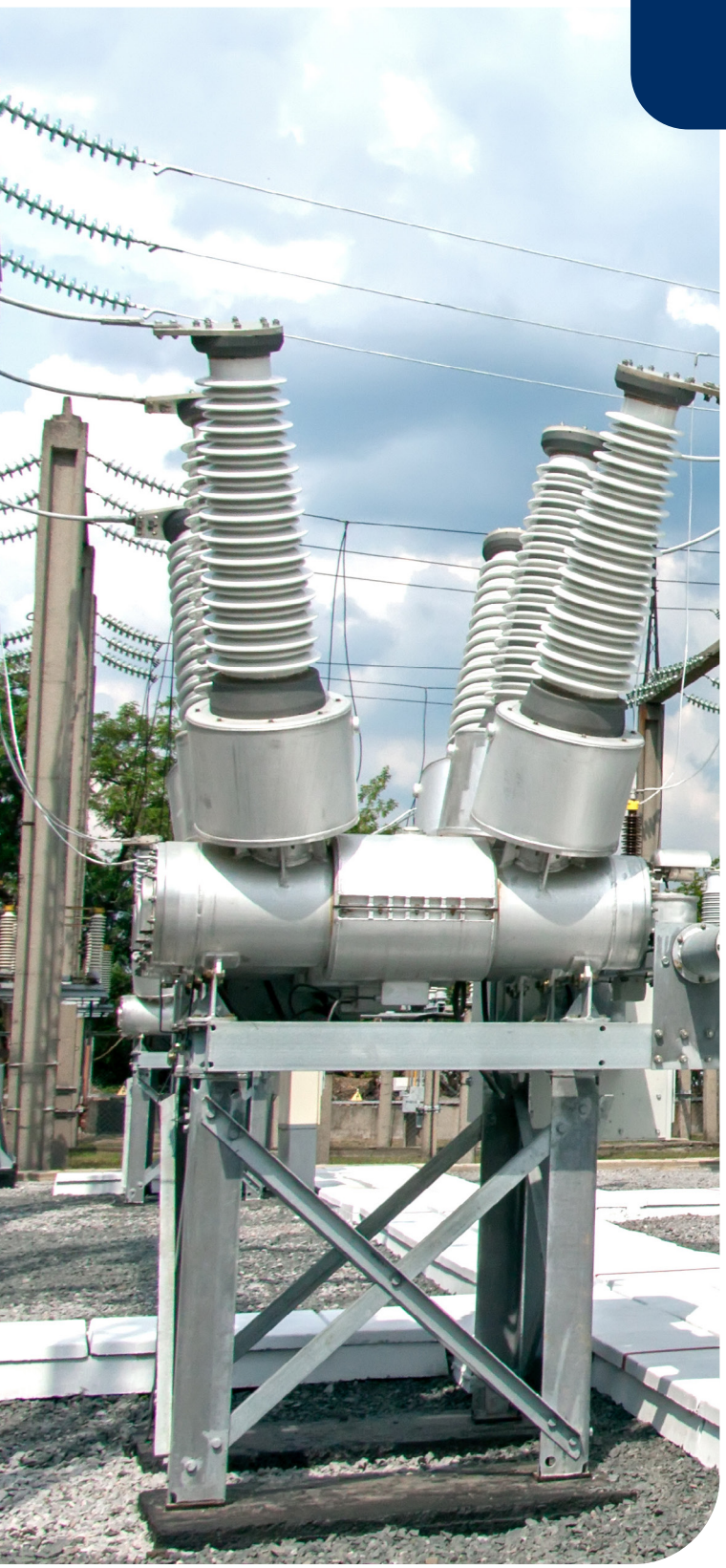
What this guide does not do. It does not prescribe a single "best" AHI formula or setpoint library. Instead, it shows how to connect whatever analytics you choose to explicit decisions, owners, and response timelines.

QUICK TIP: KPIs to confirm your AHI or CM alerts/alarms program is actually working

AHI (from your Asset Performance Monitoring System) and **CM alerts/alarms (and their related analytics)** are only valuable if they support decisions in practice and reduce risk (or cost) in measurable ways. A quick way to keep programs honest is to track a small set of KPIs that (i) reflect business outcomes, (ii) reveal whether the workflow is functioning (routing, triage, verification, decision, closure), and (iii) discourage "vanity metrics" such as alarm volume or index coverage alone. Furthermore, think of what benefit an AHI could bring if you only had one asset: the AHI won't tell you anything you couldn't work out from first principles on paper, so what is it doing for you?

AHI-focused KPIs (planning and portfolio decisions)

- **Decision adoption rate:** Percent of annual plan decisions (repair/replace/refurbish/defer) that explicitly reference AHI drivers "reason codes" related to specific components rather than a single composite score.
- **Stability / churn:** Percent of assets that change priority band per refresh cycle (e.g., Top 20 / Watchlist / Routine). High churn often indicates an overly sensitive index or inconsistent inputs.
- **Explainability completeness:** Percent of assets whose AHI includes the top contributors and supporting evidence links (tests, inspections, work history) sufficient for an engineer to explain the score.
- **Work-package alignment:** Percent of planned interventions that match the dominant drivers (e.g., bushing-driven risk leading to bushing replacement, not generic "maintenance").
- **Backlog risk exposure:** Count (or consequence-weighted count) of assets in the highest-risk band that remain without an approved mitigation plan past a defined time limit.
- **Post-intervention movement:** Median AHI (and driver) change after a completed intervention—used as a sanity check that the index responds to real work and is not "stuck."
- **Sudden changes:** Percent of cases where an asset suddenly changes AHI with no apparent reason.

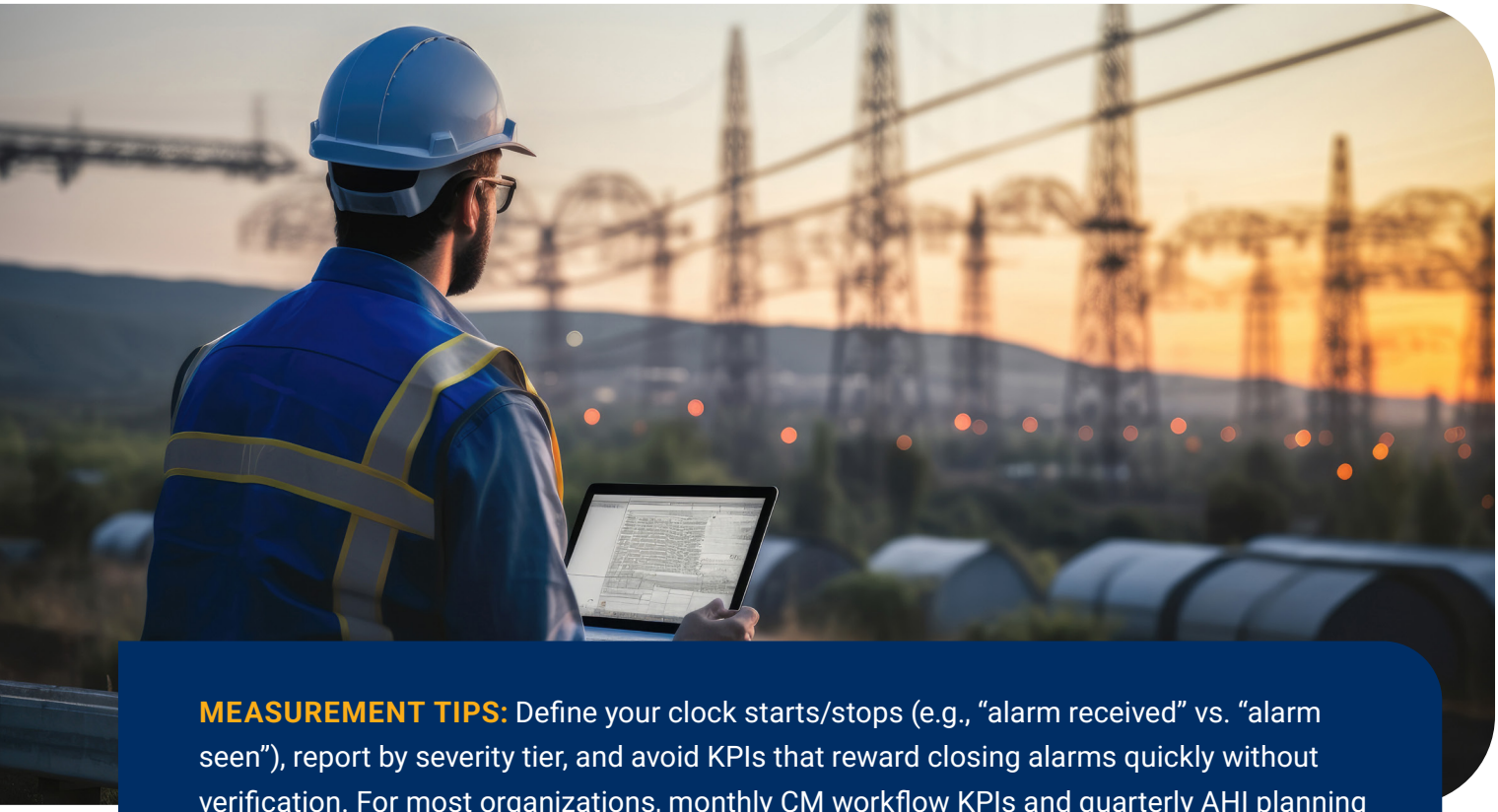


CM-focused KPIs (alarm performance and response pathway)

- **Time to awareness:** Time from condition onset (or first abnormal measurement) to when the accountable individual is notified (routing + visibility).
- **Time to triage:** Time from alarm receipt to initial classification (data quality check, severity tier, immediate actions).
- **Time to verification:** Time from alarm to completion of agreed confirmatory step (e.g., repeat sample, offline test, visual inspection).
- **Time to decision:** Time from alarm to an explicit decision (monitor / constrain / schedule outage / remove from service) with recorded rationale.
- **Actioned-alarm rate:** Percent of alarms that result in a documented action beyond acknowledgement (including “no action” with justification).
- **Nuisance (non-actionable) alarm rate:** Percent of alarms closed as invalid/ expected/noise (tracked by alarm type) to drive threshold and data-quality improvement.
- **Repeat-alarm rate:** Percent of alarms that recur for the same asset within X days after closure—often a sign of ineffective mitigation or poor closure criteria.

End-to-end outcome KPIs (tied to business impact)

- **Prevented consequence events (leading indicator):** Count of cases where CM/AHI led to an intervention that removed a credible, documented defect before it progressed (use engineering review to classify).
- **Unplanned transformer outage rate:** Trend of forced outages/failures for the monitored population (normalized and, ideally, categorized by failure mode).
- **Outage quality:** Percent of CM-driven outages that found a confirmed defect consistent with the alarm hypothesis (a practical proxy for "signal quality").
- **Response compliance:** Percent of alarms handled within agreed timelines by severity tier (triage/verification/decision).
- **Risk reduction per dollar (program economics):** Track the cost of monitoring + response effort versus quantified/estimated risk reduction or avoided cost (even if ranges for dollars are used in preference to point/single value estimates).



MEASUREMENT TIPS: Define your clock starts/stops (e.g., "alarm received" vs. "alarm seen"), report by severity tier, and avoid KPIs that reward closing alarms quickly without verification. For most organizations, monthly CM workflow KPIs and quarterly AHI planning KPIs are frequent enough to drive improvement without creating noise.

Chapter 7. The Alignment Toolkit (Definitions, Governance, and Checklists)

TOOL 1: Define terms and intended decisions. Write down what each artifact is for (APMS AHI, component AHI, maintenance index, condition monitor alarm) and what decision or decisions it supports. This aligns with repeated guidance that indices should be designed for a specific purpose and should not be repurposed casually (CIGRE TB 761, 2019; McGrail and Kennedy, CEATI 30113, 2013/2018).

TOOL 2: Start from business objectives (then engineer the pathway). Begin by stating what you are trying to prevent or optimize, then build the sensing, routing, and response around that objective. Practical monitoring guidance emphasizes that alarm routing and response are critical—e.g., online DGA benefits depend on transmission into operational pathways (often SCADA-type systems) and a proper alarm response process (CIGRE TB 962, 2025; IEEE C57.143-2024).

Lifecycle lens (so your objectives stay realistic). The "right" monitoring, inspection, and intervention strategy changes across the transformer lifecycle. Use lifecycle stage as an explicit input when defining objectives and response timelines:

- **Early life:** Focus on commissioning/installation quality and latent manufacturing or integration defects. High-value actions are often corrective and investigative (e.g., targeted adjustments, workmanship fixes, abnormal-condition follow-up).
- **Mid-life:** Failure rates often stabilize. Emphasize condition-informed inspection, monitoring, and selective intervention to sustain reliability while avoiding unnecessary intrusive work.
- **Late life:** Aging mechanisms dominate and tolerance for unresolved defects decreases. Emphasize preventive actions, risk mitigation, and life-extension decisions (including de-risking critical components and defining end-of-life criteria).

QUICK TIP: Run the checklist below once per lifecycle band (early/mid/late) or per asset class; you will often find that triggers, verification steps, and acceptable response time differ materially by stage.

Economics lens (calibrate effort to consequence). Preventive maintenance reduces risk, but you don't get the same benefit from every additional dollar or every additional inspection. Kim et al. (2020) show this as a "diminishing returns" effect: as you increase preventive effort, the incremental reliability gain gets smaller. Use this as a practical rule: invest more (and react faster) where the consequence of failure is higher, and be cautious about adding intrusive work on low-consequence assets when the risk reduction is marginal (Kim et al., 2020).

- **Objective:** Which failure mode(s) are we trying to reduce?
- **Detection:** What evidence indicates increasing risk, and how reliable is it?
- **Routing:** Where will the indication appear so that the right person sees it in time?
- **Verification:** What offline tests or inspections confirm the condition?
- **Decision:** Who can approve outage/repair/replacement?
- **Response time:** What is the maximum acceptable time from alarm to decision?
- **Review:** What can we do better next time?

TOOL 3: Build a "decision plan" (alarm response playbook). Alarm response should be planned, agreed, risk-analyzed, documented, and auditable before alarms occur (McGrail and Kennedy, 2025). This is especially important for fast-moving conditions (e.g., rapidly increasing acetylene in DGA), where delayed decision-making can have disproportionate consequences. It is also consistent with critiques that indices lacking time calibration can be misinterpreted and cannot reliably support risk statements such as Probability of Failure (PoF) without explicit modeling assumptions (McGrail and Kennedy, CEATI 30113, 2013/2018).

1. **Trigger definition:** What constitutes an alarm (thresholds, rate-of-change, persistence, data quality rules)?
2. **Primary owner:** Who receives the alarm first and is responsible for initial triage?
3. **Verification pathway:** What confirmatory tests are required, and who performs them?
4. **Escalation ladder:** Who is called next (engineering, operations, management), and in what order?
5. **Decision rights:** Who can approve an outage, load reduction, or removal from service?
6. **Time expectations:** What are the maximum times for triage, verification, and decision?
7. **Documentation:** Where are decisions recorded, and how is learning fed back into thresholds and indices?

TOOL 4: Design for real-world topology (not the ideal diagram). Utilities operate across heterogeneous architectures—from "islanded" substations with limited digital integration to fully networked environments. Treat sensor access, data latency, and alarm routing constraints as first-order design inputs. If the alarm cannot reliably reach the decision-maker, the analytic sophistication of the index will not matter.

TOOL 5: Ensure the AHI is focused on decisions. The purpose of an AHI should constrain what it includes. If the goal is to manage bulk transformer systems (e.g., tank and windings) rather than replaceable accessories, consider excluding components whose risk can be mitigated through targeted replacement. If the goal is PoF estimation for financial exposure analyses, inclusion may be broader—but then the time basis and modeling assumptions must be explicit (McGrail and Kennedy, CEATI 30113, 2013/2018).

Component-population risk lens (when trust changes after installation). Some of the hardest programs to run are those affected by a "high-concern" component population (e.g., a particular bushing or arrester design family) where confidence drops due to field experience, advisories, or emerging industry findings. In these situations, treat the component population as its own risk-managed program: define the population, establish enhanced monitoring and decision triggers, and plan staged replacement where warranted. Importantly, keep the overall AHI interpretable: preserve "reason codes" (top drivers) and avoid letting a composite score hide a population issue that requires executive visibility and funding.

- **Don’t wait for perfect certainty:** set “engineering review” triggers that are conservative when a population has low trust.
- **Separate views:** keep a fleet-planning AHI stable, and add a parallel population advisory flag/list for high-concern components.
- **Make it fundable:** translate the population issue into a staged replacement plan with clear criteria and decision rights.
- **Make it visible:** ensure dashboards show both the composite score and the population advisory status so leaders see the story, not just the number.
- **Planning index:** Optimize ranking clarity and repeatability; avoid hiding worst-case conditions.
- **Component action index:** Keep it directly actionable (clear work types) and owned.
- **Leadership view:** Always include “reason codes” (top drivers) alongside any single score.

TOOL 6: Fuse evidence to increase confidence—not to delay action. In selected cases, combining independent indicators (CM outputs plus point-in-time tests) can increase confidence in a specific failure mode and support outage justification. But don’t let “more data” become a reason to ignore the first credible warning. Masking and urgency dilution are first-order hazards in aggregation and weighting choices (CIGRE TB 761, 2019; MOH’D et al., 2024), and these hazards grow when indices are used outside their intended scope (McGrail and Kennedy, CEATI 30113, 2013/2018). **Also remember:** optimization and reliability models can be valuable decision inputs, but they do not replace engineering review and clear decision rights—especially when you are dealing with rare, high-consequence events (Kim et al., 2020).

Chapter 8. Closing Thoughts

AHIs (as outputs of Asset Performance Monitoring Systems) and condition monitor alarms serve different purposes and operate on different timescales. Treating AHIs as intervention triggers—or treating condition monitor alarm response workflows as "predictive maintenance"—creates governance ambiguity and slows escalation during rapidly developing faults. The references reviewed here converge on two practical imperatives: avoid masking and misuse of composite indices (CIGRE TB 761, 2019; MOH'D et al., 2024; McGrail and Kennedy, CEATI 30113, 2013/2018), and ensure condition monitors are paired with fit-for-purpose alarm routing and response (CIGRE TB 962, 2025; IEEE C57.143-2024; McGrail and Kennedy, 2025).

- **Keep AHIs strategic.** Use them for screening, prioritization, and planning—not as real-time triggers.
- **Keep alarms operational.** Make sure every condition monitor alarm has an owner, a verification step, and a decision path.
- **Design against masking.** Preserve visibility of worst-case and most urgent conditions and provide "reason codes" with any composite score.
- **Document the decision plan.** If you can't explain what happens after an alarm occurs in one page, the program will fail under stress.

BIBLIOGRAPHY

[1] CIGRE, *Technical Brochure 761: Condition Assessment of Power Transformers*, Working Group A2.49, March 2019.

[2] CIGRE, *Technical Brochure 962: Guide for Transformer Maintenance*, 2025.

[3] IEEE, *IEEE Std C57.143-2024: IEEE Guide for Application of Monitoring Equipment to Liquid-Immersed Transformers and Components*, 2024.

[4] T. McGrail & G. Matthew Kennedy, "Deriving Probability of Failure (PoF) from an Asset Health Index (AHI)," CEATI International Inc., Station Equipment Asset Management (SEAM) Program, Technical Brief 30113 (Project 30/113), 2013 (revised 2018).

[5] M. MOH'D, H. Schnittker, and P. Werle, "A new method for health index calculation using power transformers as an example," CIGRE Session 2024, Paper A2-PS2-11726, 2024.

[6] CIGRE USNC, "Asset Health Index (AHI) and Probability of Failure (PoF)," *Grid of the Future Symposium*, presentation PDF, 2017. Available: https://cigre-usnc.org/wp-content/uploads/2017/10/4-Reynold-C-2017-GOTF-Presentation-AHI-PoF-Final.pdf

[7] J. E. Skog and D. Klempner, "Southern California Edison's business case for transformer online monitoring," *Transformers Magazine*, vol. 9, no. 3, 2022.

[8] T. McGrail and G. M. Kennedy, "Condition monitoring: 'Remind me… what do we do when the alarm goes off?'," *Transformers Magazine*, vol. 12, no. 1, Jan. 2025.

[9] S.-Y. Kim, M. Choi, D.-W. Kim, and M.-K. Kim, "Optimal preventive maintenance: Balancing reliability and costs in the electricity market," *Energy Economics*, 2020.