

TRAINING EFFECTIVENESS MEASUREMENT MODEL

TRAINING EFFECTIVENESS MEASUREMENT MODEL

A Manufacturing-Linked View of Training Results

Cell Therapy | Gene Therapy | Biologics Manufacturing

GxP Frame

gxpframe.com

September 2026

performance and selected manufacturing outcomes change after a training intervention. Linked trends do not by themselves establish that training caused the change.

1.3 Regulatory Foundation

The sources below support qualification, continuing training, quality-system review, and appropriate evaluation. They do not prescribe the six metrics, a five-layer taxonomy, or the numerical thresholds in this customer template:

Regulation	Requirement	Model Implication
21 CFR 211.25(a)	Personnel performing drug manufacturing functions must have the necessary education, training, experience, or combination; training in assigned operations and CGMP must be continuing and sufficiently frequent.	Supports qualification and continuing training. It does not prescribe this model or any metric threshold.
FDA Quality Systems Approach guidance (2006), Personnel Development	Describes evaluation of training effectiveness and verification that learned skills are used in day-to-day performance. The guidance is nonbinding.	Supports evaluating training beyond completion; the metric design remains a site decision.
ICH Q10, Sections 1.1, 2.4 and 3.2	Describes a pharmaceutical quality system model and management responsibilities, including resources and monitoring. ICH Q10 states it does not add expectations beyond regional GMP requirements.	Useful quality-system context, not a mandate for a fifth layer or six indicators.
ICH Q9(R1)	Describes quality-risk-management principles.	May inform which tasks warrant closer monitoring; does not set this package's cadence or thresholds.
EU GMP Annex 1 (2022), Personnel	Contains additional personnel, training, and qualification expectations for sterile medicinal-product manufacture.	Apply only where Annex 1 is relevant; it does not require this analytics model.

2. Four Kirkpatrick Levels in a GMP Setting

The Kirkpatrick model (Reaction, Learning, Behavior, Results) is a useful organizing framework for evaluating a learning intervention. Its Results level can include manufacturing outcomes. This package adds a defined way to examine selected GMP process data within that level, alongside direct evidence of task performance and the site's existing quality-system measures.

2.1 What Each Level Establishes and Its Limits

Level	What It Measures	Manufacturing-Context Limitation
1. Reaction	Learner response to the training experience	Useful for improving delivery; does not establish task performance.
2. Learning	Knowledge or skill demonstrated in an assessment	A controlled assessment may not represent every production condition.

- **Separate association from causal judgment.** A trend after a curriculum change is a reason to investigate. Examine task exposure, process or equipment changes, staffing, reporting practice, and other plausible explanations before drawing a conclusion.
- **Tier measurement effort to risk.** Not every curriculum needs Layer 5 instrumentation. Reserve it for GMP-critical, product-quality-impacting, and patient-safety-impacting task training, consistent with ICH Q9(R1).

4. Layer 5: Manufacturing-Linked Indicators

Six candidate indicators form the instrument set. Select only those with a defensible denominator, reliable source records, and a documented interpretation. Some sites will not have all necessary data; a metric should be omitted rather than estimated from incomplete joins.

4.1 The Core Metric Set

Training-Attributable Deviation Rate (TADR)

Count of deviations with an approved investigation finding that training or qualification contributed, divided by a site-defined eligible exposure such as comparable task opportunities or hours. State the denominator and exclusions; do not assign a shared batch to multiple cohorts without a defined rule.

Interpretation: A sustained change may justify review of curriculum, qualification, process changes, and investigation coding. It is not by itself evidence that training caused the change.

Right-First-Time Rate by Cohort (RFT-C)

Percentage of comparable process-step opportunities completed without a defined rework, rejection, or deviation event, segmented only where a cohort can be linked reliably to the opportunity.

Interpretation: Differences between cohorts may reflect task mix, equipment, product, shift, experience, or reporting practice. Compare like with like before acting.

Time-to-Independent-Work (TTIW)

Elapsed time from training completion to supervisor sign-off for unsupervised task performance, by role and curriculum version.

Interpretation: A longer interval can prompt review of the training path, assessment, staffing, task availability, and supervisor sign-off practice. It does not identify the cause alone.

Qualification-to-First-Deviation Interval (QFDI)

Elapsed time or eligible task opportunities from qualification to a first deviation with an approved training-contribution finding, when such an event occurs. Report censoring and limited event counts.

Interpretation: A cluster of early events warrants investigation of shared curriculum and process conditions; it does not establish a curriculum defect.

Retraining Recurrence Rate (RRR)

Percentage of personnel with an approved training-related intervention who have a subsequent qualifying event during a site-defined follow-up period. State the eligible population and follow-up window.

3. **Implement the analysis in a controlled reporting layer appropriate to its intended use. QA and IT should determine access controls, change control, testing, and any computer-system validation needed under the site's procedures.**
4. **Verify linkage and denominator logic before reporting. Reconcile a risk-based sample to source records, investigate unmatched or duplicate records, and document the result and owner.**

SYSTEM CONTROL NOTE: QA and IT must determine the reporting layer's intended use, access controls, record retention, testing, change control, and any applicable computer-system validation under site procedures. Do not assume that internal trend analysis is exempt, or that this template is a validated system.

6. Statistical Method and Cohort Analysis

This section sets limits for interpreting sparse, heterogeneous data. Record the metric definition and denominator, show event counts and exposure, and account for process changes and other plausible explanations.

6.1 Minimum Cohort Size

Do not use a fixed 8–10-person threshold as a universal rule. Define minimum exposure and event-count criteria for each measure with QA or an appropriate analyst. Show counts and uncertainty where possible; suppress or aggregate small cells when privacy or unstable estimates make direct reporting unsuitable.

6.2 Control Charting, Not Threshold Alarms

Choose the display from the data structure: for example, a p-chart for a proportion with a defined denominator or a u-chart for event counts per varying exposure, if assumptions and baseline stability are met. A run chart may be more appropriate with sparse data. Predefine review rules and confirm a signal against case mix, coding practice, and process changes before attributing it to training.

6.3 Distinguishing Curriculum, Trainer, and Cohort Effects

When a metric moves, review curriculum version, trainer, and learner cohort together with product, process, equipment, shift, task exposure, supervision, and reporting changes. Use comparisons only where groups are reasonably comparable. The observation may remain unexplained; a causal conclusion requires a supported investigation, not a dashboard contrast.

POTENTIAL CONFOUND: A recent hiring or process ramp-up may change task exposure, supervision, and event reporting. Compare groups only after documenting these differences; a change is not automatically a curriculum failure.

7. The Effectiveness Scorecard

Define a scorecard owner and a review cadence before publication. Include data provenance, numerator and denominator, exclusions, uncertainty or data sufficiency, and the action taken on each signal. A site may use more than one view where operations differ.