

Composite Health Score (CHS): Executive Summary

A Statistical Framework for Measuring Team Health in Jira Environments

The Problem

When organizations measure team performance using **percentile rankings**, a fundamental problem emerges: if all teams improve equally, all percentiles stay the same. An organization might invest significantly in training and improvement initiatives, see genuine progress across every team, yet dashboard metrics show zero change.

This is not a bug—it is inherent to percentile-based systems. Percentiles only measure *relative position*, not absolute improvement. When everyone moves up together, nobody's position changes.

The Result: Stakeholders lose trust in metrics. Teams feel their improvement efforts are invisible. Improvement programs are questioned despite genuine success.

The Solution: Composite Health Score (CHS)

CHS is a **multi-component scoring system** that makes real improvement visible while maintaining meaningful peer comparison. It combines three complementary measures:

1. Current State Score (CSS) — 50% weight

What it measures: Where the team is RIGHT NOW compared to fixed baseline standards

- Uses z-score standardization against baseline norms established at calibration
- Not affected by what other teams are doing today
- Enables true longitudinal comparison

Key benefit: If all teams improve, all CSS scores increase. Real improvement is visible.

2. Trajectory Score (TRS) — 35% weight

What it measures: How the team is TRENDING during the assessment period

- Uses effect-size methodology (Cohen's d) to measure improvement magnitude
- Compares early-period performance to recent-period performance

- TRS = 50 means stable; TRS = 70 means strong improvement

Key benefit: Captures momentum and direction, not just current level.

3. Peer Growth Score (PGS) — 15% weight

What it measures: Growth compared to teams that started at a similar level

- Groups teams by baseline, then ranks growth within groups
- Applies statistical shrinkage to stabilize small-group estimates
- A team improving from 30th to 50th percentile gets more credit than one going from 80th to 85th

Key benefit: Fair comparison that accounts for starting position.

Interpretation Guide

CHS Score	Category	What It Means
70-100	Excellent	Significantly above baseline with strong trajectory
55-69	Good	Above baseline with positive direction
45-54	Average	Near baseline norms, stable
30-44	Below Average	Room for improvement
0-29	Needs Attention	Significant gaps, intervention needed

Key Features

Uncertainty Quantification

Every CHS score includes a **90% confidence interval**. A score of "65 [58, 72]" means we're 90% confident the true score falls between 58 and 72. Small score differences should not be over-interpreted.

Sensitivity Analysis

The system automatically tests scores under alternative weight configurations. If scores change dramatically with different weights, users are warned that results are sensitive to methodology choices.

Graceful Degradation

- Fewer than 20 teams? PGS is automatically excluded; 2-component model used
- Less than 8 weeks of data? TRS weight is reduced; score marked "provisional"
- Missing indicators? Available data is reweighted if coverage exceeds 70%

Comparison: CHS vs. Pure Percentile

Aspect	Percentile Scoring	CHS
When all teams improve equally	No change visible	All scores increase
Accounts for trajectory	No	Yes (35% weight)
Accounts for starting position	No	Yes (via PGS)
Uncertainty quantified	No	Yes (SE and CI)
Gaming resistant	Low	Higher

How CHS Works with CPS

The **Composite Health Score (CHS)** measures health at a single point in time.

The **Composite Progress Score (CPS)** measures progress between two points.

Typical workflow:

1. Assessment T1: Calculate CHS_before = 42
2. Assessment T2: Calculate CHS_after = 58
3. Progress: Calculate CPS to measure quality of improvement

CPS and CHS share the same statistical foundations but answer different questions.

Minimum Requirements

Requirement	Minimum	Recommended	If Not Met
Teams	20	50+	PGS unavailable
Measurement periods	4 weeks	12+ weeks	TRS unreliable
Indicator coverage	70%	90%	Reweighting applied

Statistical Foundation

CHS is built on established statistical methods:

- **Student Growth Percentiles** (Betebenner, 2009) — peer comparison methodology
- **Cohen's d effect sizes** (Cohen, 1988) — trajectory measurement
- **Empirical Bayes shrinkage** (Morris, 1983) — small-sample stabilization
- **T-score conventions** — scale centered at 50 with SD = 15

The methodology has undergone multiple rounds of statistical review and is designed for production deployment.

Bottom Line

CHS solves the fundamental problem of percentile-based measurement: it makes genuine improvement visible even when all teams improve together.

For leaders: You can now track whether improvement initiatives are working.

For teams: Your improvement efforts are visible in your scores.

For stakeholders: Metrics reflect reality, building trust in the measurement system.

For complete methodology details, see the full academic paper: `chs-academic-paper.md`

For threshold configuration guidance, see: `chs-threshold-adjustment-guide.md`

Document Version: 1.0 | January 2026