30-Day Test Worksheet

Turning one workforce-design diagnosis into one small test of change, with the balancing measure built in.

SETTING _____ TEAM OR POD _____ OWNER _____ DATE _____

> One test. One setting. Thirty days. **Do not leave with five ideas. Leave with one test.**
> Thirty days is a practical default, not a validated test duration. Adjust the window to the workflow, clinical risk, available data and training schedule. In residency settings, avoid spanning a block change when that would materially change the system being tested.

Step 1 See

Describe what is visible without assuming the cause.

Observable problem

Why does this matter, and to whom?

Step 2 Diagnose

Which workforce-design domain looks most relevant?

☐ Capacity ☐ Control ☐ Workflow ☐ Team ☐ Leadership

What data support this hypothesis?

What additional data do we need before we intervene?

What competing explanation should we rule out?

Step 3 Redesign

Which single modifiable condition will we change?

Hypothesis

If we _____

then we expect _____

because _____

Step 4 Test

WHERE WILL WE TEST IT?

WHO NEEDS TO TAKE PART?

START DATE

REVIEW DATE

Who will collect or extract the data?

What would make this test unsafe or inappropriate to continue?

Step 5 Measure

Four questions, and the fourth is the one most teams skip.

1. What happened to the workforce?

After-hours EHR time, perceived workload, engagement, turnover intention.

2. What happened to the process?

Touches per message, first-touch resolution, turnaround time, routing accuracy, rework.

3. What happened to care?

Access, continuity, response time, quality, safety, patient experience.

4. What happened elsewhere in the system?

RN and MA workload, routing errors, escalation delays, complaints, new constraints.

> **Improvement for one role is not system improvement if the burden simply moves somewhere else.** If physician inbox time falls and RN workload rises, the system has not improved. It has moved the burden.

Step 6 Decide

WHAT HAPPENED

- ☐ Improved
- ☐ No meaningful change
- ☐ Mixed result
- ☐ Worse, or unintended consequences
- ☐ Insufficient data

NEXT DECISION

- ☐ Scale it
- ☐ Adapt and retest
- ☐ Stop
- ☐ Collect more information

What did we learn about the system, not just the intervention?

Which assumption did the test confirm or challenge?

Step 7 Sustain

Only if the test is being adopted or spread. Sustaining the gain is a separate job from scaling the change. This is the step most programs skip, and it is why improvements decay.

What becomes standard work, and who owns it?

Which single measure stays on a dashboard, reviewed how often and by whom?

What would tell us it is slipping, and what happens then?

What did we stop or simplify to make room for this?