AI Pilot Scorecard

Define one bounded workflow, collect useful evidence, and decide whether to continue, change, or stop.

SAFE USE

Operational learning only - not clinical validation. Do not use AI for patient-care decisions or enter patient-identifiable, confidential, credential, or restricted data.

01 DEFINE ONE BOUNDED WORKFLOW

WORKFLOW + BOUNDARY

Name one repeated task. State where the pilot starts and ends.

ALLOWED INPUTS

List only approved, minimum-necessary input types.

PROHIBITED INPUTS

Name confidential, identifiable, credential, or restricted data.

PILOT OWNER

Who can pause or stop the test?

HUMAN REVIEWER

Who checks every output before use?

TEST GROUP + DURATION

Who participates, how many, and for how long?

APPROVED TOOL + SETTING

Name the exact account, workspace, and configuration.

02 RECORD THE BASELINE BEFORE TESTING

TIME

Minutes per task / cycle

ERRORS + REWORK

Count or percentage

QUALITY + EXPERIENCE

Rubric score or feedback

EVIDENCE SOURCE

Where will each baseline and pilot result come from? Name the record, log, sample, or survey.

AI Pilot Scorecard

Compare the pilot with the baseline, check safeguards, then make a documented decision.

03 SCORE THE EVIDENCE

MEASURE	BASELINE	PILOT EVIDENCE	RESULT / NOTES
Time per task	Record number + unit	Sample size / source / date	Better / same / worse
Errors or rework	Count or percentage	Incidents + reviewer notes	Better / same / worse
Quality	Rubric or acceptance rate	Reviewed output sample	Better / same / worse
Team experience	Survey or observation	Feedback themes + count	Better / same / worse

04 SAFEGUARDS AND STOP CONDITIONS

- ☐ Only approved inputs and the approved tool/account were used.
- ☐ A named human reviewed every output before use or sharing.
- ☐ Privacy, security, quality, access, and incident checks were completed.

STOP CONDITIONS

Pause for any privacy, safety, security, policy, access, quality, or workflow failure. Record what happened and who was notified.

05 MAKE THE DECISION

○ CONTINUE

Evidence supports a bounded next cycle. Keep safeguards and name the next review date.

○ CHANGE

Revise the workflow, tool, training, scope, reviewer step, or evidence plan before retesting.

○ STOP

Risk, policy, quality, value, or evidence does not support continued use.

DECISION + RATIONALE

What evidence supports this decision?

OWNER + NEXT REVIEW

Name / date