



Build for the real world. Test in a virtual one.

Find implementation failures
before a healthcare pilot.

A three-month test of whether synthetic
workflow experiments help builders improve a design.

team@alivehealthtech.com





The same technology. Different workflows. Different outcomes.

Video visits were popular when the clinical and practical conditions fit. Yet the same study found sharply different outcomes across four services.

YOUNG ADULT DIABETES + CANCER SURGERY

About 20% of follow-ups moved to video.

Remote consultations became part of routine service delivery.

PRENATAL DIABETES + COMMUNITY HEART FAILURE

Abandoned in one.
Put on hold in the other.

Multidisciplinary handoffs and missing records strained one workflow; examination needs constrained the other.

A working interaction is not a working care pathway.

Before a live pilot, builders need to ask how their product behaves when staffing, handoffs, data and patient participation interact.



An empty alert queue can hide a broken product.

A monitoring startup tests the classifier together with device gaps, service latency and nurse shifts. 200 synthetic wearers, seven days and 30 paired runs expose a failure hidden by a quiet queue.

monitoring / experiment workspace

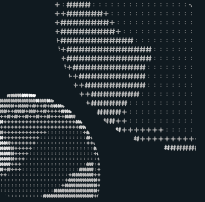
Local synthetic run

EXPERIMENT

world.json

models.lock.json

results.zip



Four variants.
Same world.
Matched seeds.

world.json / configuration excerpt

```
// Existing mock model + operating constraints
adapter_timeout_seconds: 5
latency_seconds: 20
fallback_to_threshold: false

staff: 1
shift_hours: 8
workdays: 5
```

TERMINAL

```
> alive run --world world.json --output results.zip
SUCCEEDED · 4 variants × 30 replications
Open metrics, assumptions and event traces.
```

Extracted results / modeled outcomes

Ideal conditions

98.8% detection

Among received abnormal samples; 50.3% of alerts are false alarms.

Latency stress

0% alerting

Replies arrive after the deadline. No alerts reach triage; its empty queue masks the failure.

Fallback retest

5.6 hours to review

Against an assumed 2-hour target. Mean queue wait for alerts starting review; unfinished waits excluded.



Put the workflow around the product.

Compose synthetic people, processes, staff capacity and service behavior around a proposed change. Run matched variants, inspect the event trail, change the design and retest.

Plan the right test.

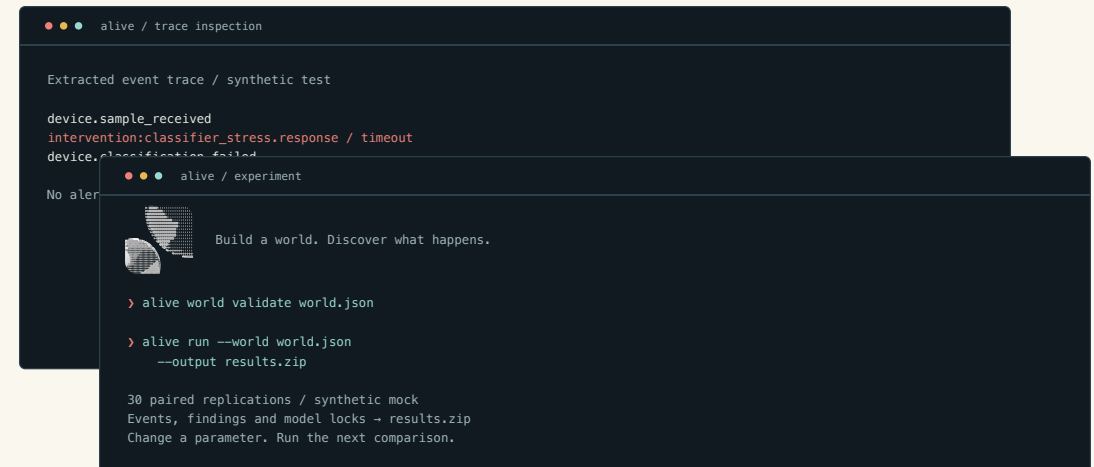
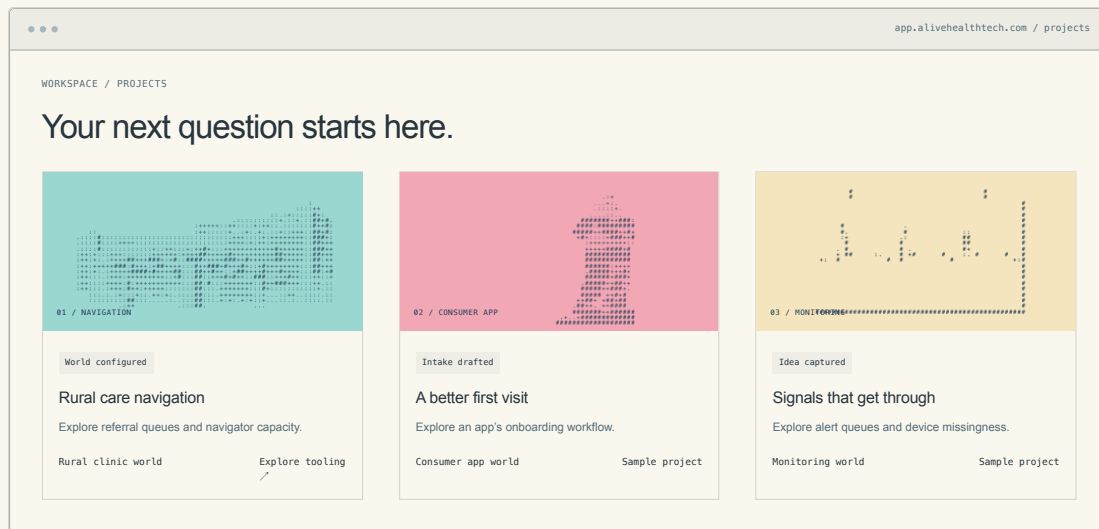
Specify the workflow and decision.

Build its environment.

Set demand, capacity and failure conditions.

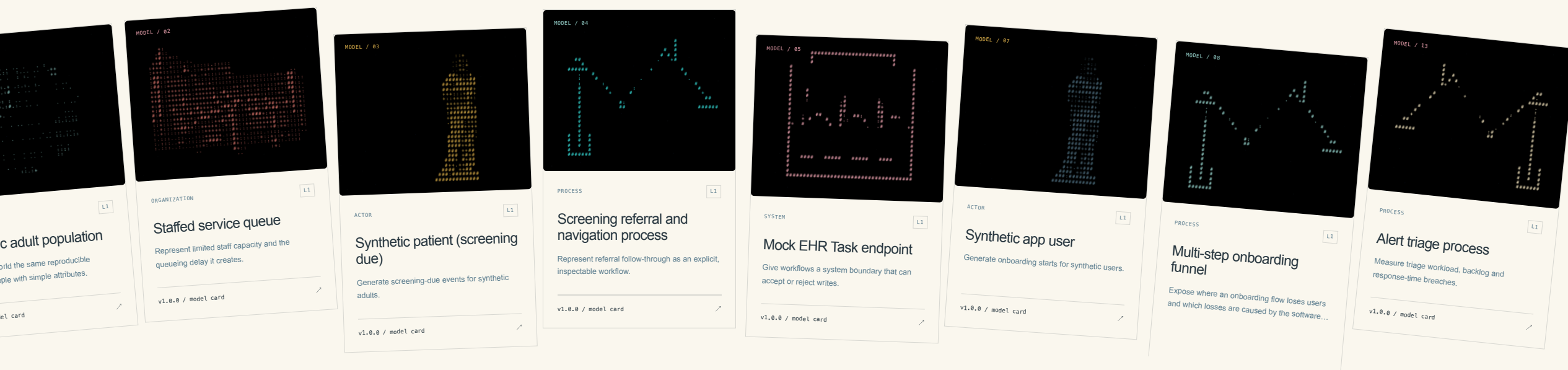
Inspect. Change. Retest.

Compare variants and preserve the trace.





Reusable models. Evidence you can inspect.



L1 • Executed

Observed behavior, event trace and assertion. The cards shown start here.

L2 • Internally verified

Replay, conservation, boundary and distribution tests pass for the stated use.

L3 • Calibrated

Fit to an external dataset; publish comparators, tolerances and applicability.

L4 • Externally validated

Independent back-test or prospective comparator; disclose predictive error.

Advance only with a versioned benchmark report and reviewer sign-off. Repeated use creates test cases; it does not establish accuracy.



Care is changing settings. The workflow has to keep up.

More tools in clinical work

72%

Report using at least one AI application.

AMA survey, 2026: 1,342 responses to the use-case question, up from 48% in 2024. Self-reported use; survey wording evolved.

More care outside hospital walls

Hospital-level care. Home-level constraints.

CMS's hospital-at-home study describes selection based on patients' circumstances, home environments and local resources. Technology must fit those conditions.

Evidence is an adoption requirement

88%

Say safety and efficacy validation matters.

AMA, 2026. Workflow experiments can inform what to test; external validation and live evaluation remain separate evidence gates.

The immediate opportunity: test the interactions around a new care workflow while its design can still change.



Bring implementation risk into the development loop.

Repeatable development-loop test ↑

Software behavior Antithesis Faults, execution and reproducible software tests. Synthea Repeatable synthetic patient-record generation.	Alive's intended opening Product + operating environment Reusable workflow models, technical failure paths and human/resource constraints in a run → inspect → retest loop.
Study inputs and assurance Synthea supplies records to studies. ORCHA assesses digital-health products.	Systems and behavior studies AnyLogic · Simul8 · FlexSim Operational and capacity simulation. Synthetic Users LLM-persona user research.

Technical faults / data

Operational / human behavior →

One-off study ↓



Start with one team preparing a monitoring pilot.

USER #1

A digital-health builder designing remote monitoring.

A founder or research engineer with an alerting prototype, a defined staff workflow and a near-term pilot. Their question: will a service failure or shift boundary silently break escalation?

Begin with synthetic inputs and the local CLI. Capture the decision, the experiment and the resulting design change.

3 teams

Proposed three-month cohort

Two early-stage digital-health product teams.

One university research team developing a monitoring workflow.

One workflow owner per team for assumptions and result review.

Recruit through labs and accelerator cohorts. Interview at least five builders before onboarding; record needs, consent and baseline decisions.



Built at the intersection of research and delivery.



Isaac Wilkins

Founder & President · Alive

Founder-led pilot: scenario design, engineering, onboarding and evaluation.

Harvard Medical School & Wheatley Institute

Co-investigator & Lead Researcher. Building a 250-intervention dataset on digital-health adoption and sustainment; translating implementation questions into testable assumptions.

MGH Healthcare Transformation Lab

Research Fellow. Led development and launch of two digital tools; AHA software collaboration and hospital-at-home technology-priority work.

Wildfire Partners VC

Founding Team; Advisory Board. Fund formation, investor materials and digital-health investment strategy.

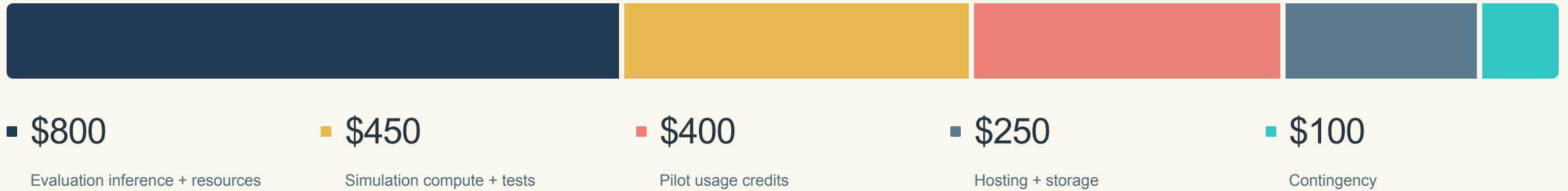
McKinsey & Company

Incoming 2027 Summer Business Analyst. Future role, after this proposed three-month pilot.

Scope the pilot to three teams and one workflow. The \$2,000 funds operating expenses, with \$0 personnel; recruiting roles are outside the funded delivery plan. Weekly availability and contributor arrangements must be confirmed before launch.



\$2,000 to test whether builders come back.



Month 1 · Establish the benchmark

Replay + explain.

Publish a locked monitoring benchmark, boundary tests and parameter provenance; interview ≥5 builders.

Month 2 · Make it usable

Onboard 3 teams.

Each runs a baseline and changed variant, inspects a trace and records a design decision.

Month 3 · Test the value

Change. Return. Rerun.

≥1 team changes a design after a finding; ≥2 teams return to test a later version.

Deliver the evidence, even if the pilot misses its targets.

Report runs, decisions, return use, actual spend and barriers. This grant tests usefulness; external calibration is a separate workstream.