



VANTHEON

A Protocol for Verifiable Healthcare

*Care that proves itself, trust computed from results, and value that settles
to whoever produces the outcome, human or machine.*

Alejandro S. Méndez Valentín

x/acc labs • San Juan, Puerto Rico • amendez@xacc.io

Whitepaper v1.1 (draft) • October 2026

Abstract

Proof, not permission. Healthcare has learned to compute its fragments—claims, codes, forms—but not the complete event: what was required, what was done, by whom, under whose authority, and what happened to the patient. Because the complete event is never computed, the system pays for proxies of care and spends a large share of its resources checking them. We specify VANTHEON, a protocol whose unit of state is the *verified event* $e = (R, V, A, D, X, O, C, L, S)$, and whose evolution is a deterministic state-transition function $\sigma_{t+1} = Y(\sigma_t, \mathcal{E}_t)$. From verified events the protocol computes a risk-adjusted, per-service trust index for every participant—clinician, organisation, software agent or robot—and acts on its lower credible bound. This yields *instant authorization*: care is approved at the moment it is ordered when rule, evidence, authority and trust predicates hold, and is otherwise decided by a person; denial is never automated. Value then settles on proof to the participant that produced it, and every result updates the rules through a versioned, human-approved process. We give the security model and prove determinism, monotonicity, concentration, cold-start safety, approval-only automation, value conservation and a solvency condition, bound the probability that fabricated events yield a net gain, which falls exponentially in the number of attempts, and give a working zero-knowledge proof that a participant clears the authorization threshold without revealing its record. We ground the design in public U.S. federal data: in 185,349 Medicare claim lines audited by CMS in 2024, 87.1% of improper-payment findings fail a predicate the protocol checks before payment; in the FY2026 Hospital Readmissions Reduction Program, 85.2% of condition-level results above expectation are statistically indistinguishable from expected. A seeded simulation evaluates the full mechanism. Throughout, we separate what is implemented, specified and open research.

Preface

For decades the public-health community has pursued one vision: a learning health system that generates real-world evidence, advances equity and coordinates care across a population [19]. Frameworks such as TEFCA and standards such as FHIR built the roads for data. Yet the system has reached a plateau. It can move a record from one institution to another; it cannot say, with proof, whether the care the record describes was needed, authorised, performed and effective. Because it cannot, it pays for proxies and spends a quarter of its resources checking them.

This paper is not a critique of the work that came before it. It is the next logical step: a specification of the missing piece that lets the existing frameworks reach their purpose. The Gray Paper (v1.67, October 2025) set out the architecture—seven layers, the trust primitives, Trust-Time and a treasury. This whitepaper is its formal successor: one primitive, one state-transition function, proofs where there were assertions, and measurements from public data where there were projections. Where the Gray Paper described a capability that is not yet built, this paper says so.

1. Introduction

In 2008 a nine-page paper showed that two parties who do not trust each other can exchange value without a trusted intermediary, by making the ledger itself verifiable [1]. Ethereum generalised that ledger into a state-transition machine on which any agreement could be expressed as code [2, 3]. Avalanche showed that such machines can reach finality in about a second through repeated random subsampling [4]; Polkadot showed that many such machines can share one security model [5]. Each of these protocols removed a specific class of intermediary by making a specific class of fact computable.

Healthcare has not yet had that moment. Its intermediaries do not order payments; they adjudicate *whether care should happen and whether it happened*. That adjudication is expensive. U.S. healthcare spends an estimated \$760–935 billion a year on waste, about a quarter of all spending, of which about \$266 billion is administrative complexity [9]. Administration absorbed 34.2% of national health expenditures in 2017, against 17.0% in Canada [10]. CMS estimates that 7.66% of Medicare fee-for-service payments in fiscal year 2024—\$31.70 billion—were improper [11]. In 2024 Medicare Advantage insurers received nearly 53 million prior-authorization requests, denied 4.1 million in whole or part, and

overturned 80.7% of the denials that were appealed [12]. The system asks permission it mostly grants, refuses care it often later approves, and pays claims it cannot verify.

The intellectual answer to this problem has existed since 2006. Porter and Teisberg argued that health care should compete on *value*—health outcomes achieved per dollar spent—measured for each medical condition over the full cycle of care [6, 7]. Value-based care, as it was implemented, kept the goal and lost the instrument: value is estimated months later from claims, for populations too small to measure reliably, and paid through contracts that rely on the same intermediaries. Porter’s equation was right; the system had no machine to compute it. We call the computation that was missing the *verified event*.

Thesis. If the complete event of care can be represented, checked and remembered, then trust can be computed from what participants actually achieve; authorization can follow computed trust instead of permission; payment can follow proof instead of claims; and every result can make the next decision better. VANTHEON is the protocol that makes this computable. It does not replace clinicians, payers or regulators. It gives them one verifiable record of what happened and one rule for what follows from it.

Contributions. (i) A formal model of healthcare as a state-transition system over verified events, in which the participant is entity-agnostic: a clinician, an organisation, a software agent and a robot are paid and governed by the same rules (§3–4). (ii) A Bayesian, risk-adjusted, per-service trust index with proofs of its safety properties, replacing the outcome-blind index of earlier drafts (§5). (iii) Instant authorization with approval-only automation (§6). (iv) Settlement on proof, a solvency condition for a member pool, and the Trust-Time allocation market (§8). (v) A security model, a fraud-economics calculation in the manner of Bitcoin’s §11, and post-quantum cryptographic agility (§10–12). (vi) An empirical grounding in public CMS data (§13) and a reproducible evaluation (§17).

Claim discipline. Each claim carries a tag: [Implemented] runs in the reference implementation on synthetic records; [Specified] is specified here; [Research] is open. Every number in this paper is produced by released code from public data or a seeded simulation.

2. Background: Value Without a Machine

Fee-for-service. pays for each visit, test and procedure. It pays for activity whether or not the patient improves, and it relies on documentation as the proxy for activity.

Value-based care. ties payment to quality and cost through shared savings, bundles and capitation. It depends on *attribution* (who is responsible for which patient), *risk adjustment* (accounting for how sick patients are, e.g. CMS-HCC) and *quality measures* (eCQM, HEDIS, CAHPS), computed months after care from claims.

Prior authorization. is a payer’s permission before certain services; CMS-0057-F now requires decisions within 72 hours (urgent) and 7 days (standard), a specific reason for every de-

Table 1. Porter’s value agenda [6, 8] and its computable counterpart.

Porter	VANTHEON object
Value = outcomes / cost	$\mathcal{V}_{p,k}$, eq. (1)
Outcomes per medical condition	trust per service k , Def. 5.1
Risk-adjusted results	outcome model $\pi(q, r)$, eq. (3)
Cost over the cycle of care	settlement ledger $\mathcal{P}$, Def. 8.1
Bundled payment	contract $\mathcal{C}$ over verified events
Information platform	the verified event, Def. 3.3

Table 2. What each protocol makes verifiable.

Protocol	Verifiable fact	Intermediary removed
Bitcoin	order of payments	payment processor
Ethereum	execution of agreements	contract enforcer
Avalanche	finality, at scale, in about a second	settlement delay
Polkadot	shared security across chains	per-chain security
VANTHEON	that care was needed, authorized, performed and effective	permission and paperwork

nial and FHIR APIs by 2027 [13]. It makes permission faster; it remains permission.

Interoperability. standards—FHIR R4, USCDI, CDS Hooks and the Da Vinci guides—move data. They do not make an event verifiable.

Porter’s framework names five things a value system needs: outcomes measured per condition, cost measured per patient over the cycle of care, bundled payment for the cycle, integrated delivery, and an enabling information platform [8]. The first four are economic designs; the fifth is infrastructure, and it is the one that never arrived. Table 1 maps each element to a protocol object. In VANTHEON, Porter’s ratio is computed per provider, per service, per event:

$$\mathcal{V}_{p,k} = \frac{\mathbb{E}[\text{verified outcome} \mid p, k]}{\mathbb{E}[\text{settled cost} \mid p, k]} = \frac{\text{PTI}_{p,k}}{\bar{c}_{p,k}}, \quad (1)$$

where $\text{PTI}_{p,k}$ is the risk-adjusted probability of a verified good outcome (§5) and $\bar{c}_{p,k}$ the mean value settled per event (§8).

In words. Porter defined value as outcomes divided by cost. The protocol computes both sides from verified events, for every provider and every kind of care, instead of estimating them a year later.

What earlier protocols made computable. Table 2 places VANTHEON beside the protocols whose method it follows. Each removed a class of intermediary by making one class of fact verifiable; none addressed whether care was needed, performed, or effective.

What the data say about value today. In the public CMS hospital files, heart-failure 30-day mortality and Medicare spending per beneficiary (MSPB) are nearly unrelated across 2,511 hospitals (Spearman $\rho = -0.067$; Fig. 1): spending more does not buy survival, and the market does not price the difference. Mortality ranges from 0.087 to 13.5% between the 10th and

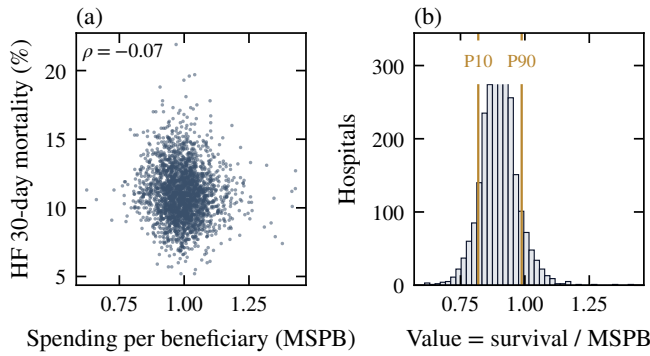


Figure 1. Porter’s value from public CMS data. (a) Heart-failure 30-day mortality vs. Medicare spending per beneficiary (MSPB, 1.0 = national median) for 2,511 hospitals. (b) Distribution of value = $(1 - \text{mortality})/\text{MSPB}$.

Table 3. Participant classes and their protocol rights.

Class	May
Clinician, organisation	request, decide, deny, attest, receive settlement
Agent (software)	gather evidence, evaluate rules, act within γ , receive settlement
Robot	perform physical actions within γ , receive settlement
Patient	grant and revoke consent κ ; attest experience
Pool	receive contributions; pay on proof
Steward	approve rule and parameter changes

90th percentiles while spending varies by a factor of 1.199. Porter’s value exists in the data but no payment flows on it.

3. System Model

3.1. Participants are entity-agnostic

A participant is anything that can hold a grant and be attributed an outcome. The protocol does not distinguish a person from an organisation, a software agent or a robot by what it *is*, only by what it is *authorized* to do and what it has *proven*.

Definition 3.1 (Participant). A participant is a tuple $a = (id, class, cred, key)$ with $class \in \{\text{clinician, organisation, agent, robot, patient, pool, steward}\}$, verifiable credentials $cred$ (e.g. licence and NPI for clinicians, device certification for robots, model and version attestation for agents) and a public key key .

Definition 3.2 (Grant). A grant is $\gamma = (holder, purpose, scope, owner, t_{exp})$. An action x by holder a at time t is *authorized* iff $x \in scope(\gamma)$, $t < t_{exp}$ and $owner(\gamma)$ is a clinician or organisation accountable for x .

In words. A doctor, a clinic, an AI agent and a robot are all participants. Each can do only what an accountable party allowed, for a stated purpose, until a stated time, and each is paid for what it proves. Capable is not authorized.

3.2. The verified event

Definition 3.3 (Verified event). A verified event is a tuple $e = (R, V, A, D, X, O, C, L, S)$: $R \in \mathcal{R}$ is the versioned rule that applied; $V = \{(v_j, src_j, t_j, sig_j)\}$ the evidence, each value signed by its source; $A = (\kappa, \gamma, role)$ the authority, including consent; $D \in \{\text{APPROVE, REVIEW, DENY}\}$ the decision and X the action, with its performing participant a_X ; $O \in \{0, 1\}$ the observed outcome; $C = (c_1, \dots, c_m)$ the verification results; L the proposed rule update, if any; S the settlement. Each event carries $H(e) = H(\text{canon}(e))$ for a collision-resistant hash H .

In words. One record holds everything needed to know that care happened, that it was right, who allowed it, who did it, what came of it, and what it is worth.

[Implemented] Events are built from FHIR R4 resources on synthetic patients (Synthea); rules are evaluated in CQL, evidence is cited by FHIRPath, authority is checked by relationship-based access control (OpenFGA) and policy (OPA), long-running events are orchestrated by Temporal, and events are appended to an immutable log (Fig. 2).

3.3. Network and adversary

We assume (A1) the event log is append-only and its commitments are anchored to a public ledger with the ledger’s safety (on Avalanche, safety holds with overwhelming probability while the Byzantine share of sampled stake is below the protocol’s threshold [4]); (A2) each event has at least one verification path independent of the participant whose trust it affects; (A3) the outcome model is correctly specified up to its risk factors; (A4) the hash and signature schemes are secure, including against quantum adversaries for the schemes of §10.2. The adversary controls any number of participant identities and chooses which patients to accept, what to document and when to re-enter under a new identity; it cannot forge source-signed evidence or break (A1), (A4).

4. State and the State-Transition Function

Following the Ethereum Yellow Paper [3], the protocol is the iterated application of one function.

Definition 4.1 (Protocol state). $\sigma = (\mathcal{W}, \mathcal{T}, \Pi, \mathcal{P})$, where $\mathcal{W} = (\mathcal{R}, \Gamma, K, \Theta)$ holds rules, grants, consents and parameters; $\mathcal{T} : \mathcal{A} \times \mathcal{K} \rightarrow \Delta[0, 1]$ maps each participant a and service k to a posterior over true quality; Π is the treasury balance; and $\mathcal{P}$ is the attribution ledger.

Definition 4.2 (State-transition function). For the batch $\mathcal{E}_t$ of events closed in epoch t ,

$$\begin{aligned} \sigma_{t+1} &= Y(\sigma_t, \mathcal{E}_t) \\ &= \Omega(\Lambda(\Phi(\Delta_\lambda(\sigma_t), \mathcal{E}_t^\vee))), \end{aligned} \quad (2)$$

with $\mathcal{E}_t^\vee = \{e \in \mathcal{E}_t : \mathbb{V}(e, \sigma_t) = 1\}$; Δ_λ decays trust evidence; Φ updates trust (Def. 5.1); Λ settles value (Def. 8.1); Ω applies steward-approved changes queued before t .

In words. Each epoch, the protocol takes the previous state and the cases that closed, keeps only the ones that pass every check, up-

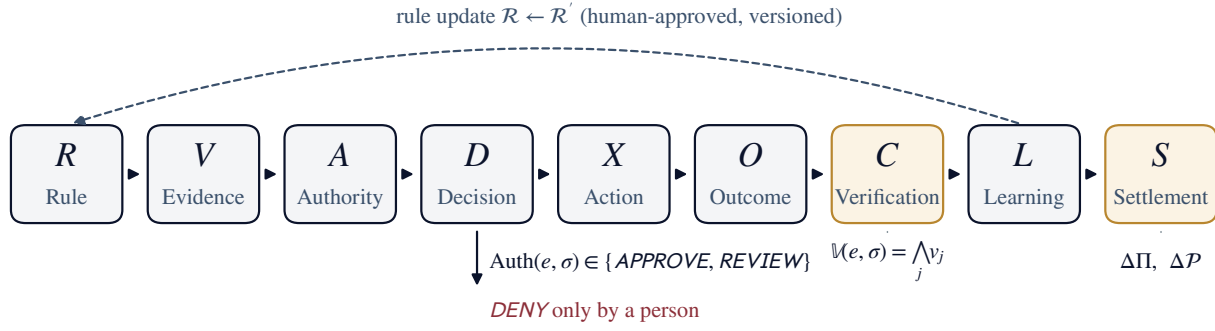


Figure 2. Life cycle of a verified event. Each component is produced by the participant entitled to it (Table 3). The decision function Auth has image $\{\text{APPROVE}, \text{REVIEW}\}$; DENY is reachable only by a person. Verification $\mathbb{V}$ is the conjunction of independent checks. Settlement updates the pool Π and the attribution ledger $\mathcal{P}$; learning proposes a human-approved rule version $\mathcal{R}'$.

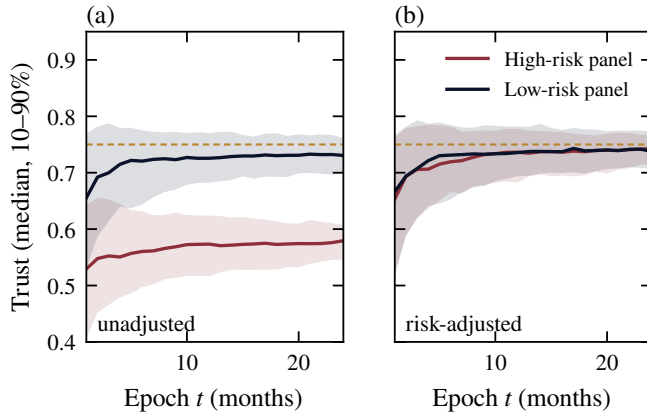


Figure 3. Fairness. 400 participants of equal quality ($q = 0.75$, dashed), half serving a high-risk panel. (a) Unadjusted; (b) risk-adjusted.

dates trust, settles value and applies the rule changes people approved. Nothing else changes the state.

Definition 4.3 (Verification predicate). $\mathbb{V}(e, \sigma) = \bigwedge_{j=1}^4 v_j(e, \sigma)$: v_1 , R is satisfied by V under its computable definition (CQL); v_2 , every evidence element required by R is present and source-signed; v_3 , A is valid at t_e (consent unrevoked, grant in scope and unexpired); v_4 , the outcome O is attested through a path independent of the performing participant (A2).

5. Computed Trust

5.1. Outcome model

For an event with patient risk $r \in (0, 1)$ (e.g. from CMS-HCC factors) performed by a participant of true quality q for service k ,

$$\pi(q, r) = \sigma(\text{logit } q - \beta_r(r - \bar{r})), \quad (3)$$

with σ the logistic function, $\bar{r}$ the mean risk and $\beta_r > 0$.

In words. Sicker patients are less likely to reach the goal whatever the provider does. The model expects that, so it does not blame the provider for it.

5.2. The trust index

Definition 5.1 (Trust index, VIP-0001). For participant a and service k with verified events $\mathcal{E}_{a,k}$,

$$\mathcal{T}_{a,k}(q) \propto \text{Beta}(q; \alpha_0, \beta_0) \times \prod_{e \in \mathcal{E}_{a,k}} [\pi(q, r_e)^{O_e} (1 - \pi(q, r_e))^{1-O_e}]^{w_e}, \quad (4)$$

with $w_e = e^{-\lambda(t-t_e)}$. The index and its lower credible bound are $\text{PTI}_{a,k} = \mathbb{E}_{\mathcal{T}_{a,k}}[q]$ and $\text{PTI}_{a,k}^- = F_{\mathcal{T}_{a,k}}^{-1}(\delta)$.

In words. Trust is what a participant's verified results show for this kind of care, after accounting for how sick the patients were, with older results counting less. The protocol acts on the careful end of that estimate.

Remark 5.2 (Relation to the Gray Paper). Gray Paper v1.67/v1.81 defined $\text{PTI} = \alpha_1 \text{OA} + \alpha_2 \text{DT} + \alpha_3 \text{MF} + \beta H e^{-\lambda \Delta t}$ over reported components. Executed on 2026-10-12, the reference implementation of that index gives 0.90 both to a provider with 0 of 3 successes and to one with 1 of 1, because outcomes enter only an unused interval. Definition 5.1 conditions on outcomes directly; OA, DT and MF are retained as explanations (CCS/CXS, §5.4), not as decision inputs.

5.3. Properties

Theorem 5.3 (Determinism). Y is a function: for any σ and any ordering of the same batch, Y yields the same σ_{t+1} .

Proof. $\mathbb{V}$ depends only on (e, σ_t) . The likelihood in (4) is a product over events, which commutes; Δ_λ scales log-likelihoods by $e^{-\lambda}$; Λ sums over $\mathcal{E}_t'$; Ω applies a queue fixed before t . Each stage is order-independent, so their composition is. $\square$

Theorem 5.4 (Monotonicity). Adding a verified success never decreases $\text{PTI}_{a,k}$ or $\text{PTI}_{a,k}^-$; adding a verified failure never increases them.

Proof. $\pi(q, r)$ is strictly increasing in q . The new posterior is the old one reweighted by $\pi(q, r)^w$, an increasing function of q . Reweighting by an increasing function induces a monotone likelihood ratio and hence first-order stochastic dominance, so

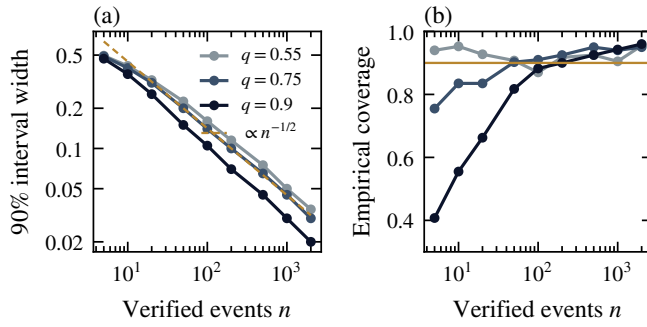


Figure 4. Concentration (Thm. 5.5), 400 replications per point. (a) Median 90% interval width; dashed $n^{-1/2}$. (b) Empirical coverage; the skeptical prior under-covers strong participants at small n by design and reaches nominal coverage by $n \approx 10^2$.

the mean and every quantile weakly increase. The failure case is symmetric with $1 - \pi$. $\square$

Theorem 5.5 (Concentration). *Under (A3), with $\lambda = 0$ and n events with i.i.d. risks, $\mathcal{T}_{a,k}$ concentrates at the true q and the width of its central $(1 - 2\delta)$ interval is $\Theta(n^{-1/2})$.*

Proof sketch. Model (3) is a regular one-parameter GLM in logit q with Fisher information $I(q) = \mathbb{E}_r[\pi(1 - \pi)/(q(1 - q))^2] > 0$. By the Bernstein–von Mises theorem the posterior is asymptotically $\mathcal{N}(\hat{q}, (nI(q))^{-1})$, so the width is $2z_{1-\delta}(nI(q))^{-1/2}$. With $\lambda > 0$ the effective sample size is bounded by $\bar{m}/(1 - e^{-\lambda})$, which keeps trust tied to current practice. $\square$

This replaces Appendix A.1 of the Gray Paper, whose logarithmic bound described the decay of an exponential moving average rather than convergence to the true quality. The $n^{-1/2}$ rate is confirmed empirically (fitted slope -0.507, Fig. 4).

Theorem 5.6 (Cold-start safety). *If $F_{\text{Beta}(\alpha_0, \beta_0)}^{-1}(\delta) < \theta_k$ for every k , no identity is eligible for instant authorization before accumulating verified evidence.*

Proof. With $\mathcal{E}_{a,k} = \emptyset$ the posterior is the prior, so $\text{PTI}_{a,k}^- < \theta_k$ and predicate a_4 of Rule 6.1 fails. $\square$

With $(\alpha_0, \beta_0, \delta) = (2, 3, 0.05)$ the prior bound is 0.098 against $\theta = 0.75$. Re-entry under a new identity therefore resets trust downward.

5.4. Explanatory scores

The Codex Clinical Score $\text{CCS} = \sum_c w_c s_c$ (effectiveness, safety, appropriateness, resource use) and the Codex Execution Score $\text{CXS} = \sum_x w_x s_x$ (adherence, timeliness, coordination) are retained from the Gray Paper [Specified]. They explain why PTI moves and are reported with every settlement; no decision depends on them, which keeps the decision variable a calibrated probability.

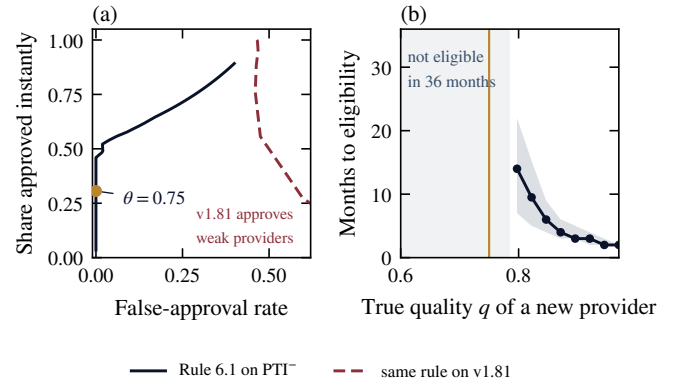


Figure 5. Instant authorization. (a) Share approved instantly vs. false-approval rate as θ varies; gold, $\theta = 0.75$. (b) Months until a new participant of quality q first satisfies a_4 (median, IQR; 200 replications).

6. Instant Authorization

Definition 6.1 (Authorization rule). For request e by participant a for service k :

$$\text{Auth}(e, \sigma) = \begin{cases} \text{APPROVE} & \text{if } a_1 \wedge a_2 \wedge a_3 \wedge a_4, \\ \text{REVIEW} & \text{otherwise,} \end{cases}$$

with a_1 : R satisfied by V ; a_2 : evidence complete; a_3 : authority valid; a_4 : $\text{PTI}_{a,k}^- \geq \theta_k$.

In words. When the guideline is met, the records are complete, the patient agreed and the provider has earned trust for this care, the answer is yes at once. Anything else goes to a person, with the reason.

Proposition 6.2 (Approval-only automation). $\text{DENY} \notin \text{im}(\text{Auth})$; every denial is a decision by an accountable person recorded in e .

Proof. Immediate from Definition 6.1; DENY is writable to D only under a clinician or organisation role. $\square$

This is consistent with CMS guidance that Medicare Advantage medical-necessity denials be reviewed by a qualified person and with laws such as California SB 1120. A fraction $\rho_{a,k} = \rho_0(1 - \text{PTI}_{a,k}^-)$ of instant approvals is sampled for retrospective audit; a failed audit enters $\mathcal{E}_{a,k}$ as a failure and, by Theorem 5.4, can only lower trust [Specified].

7. A Worked Event

Ethereum’s whitepaper explains its state-transition function by walking one transaction through it [2]. We do the same for one event.

Setting. A 62-year-old patient with type 2 diabetes (synthetic) has not had a diabetic retinal exam in 14 months; her most recent HbA1c is 7.9%. The quality measure (HEDIS Eye Exam for Patients with Diabetes; eCQM CMS131) is open. Her primary-care clinic operates an autonomous retinal-imaging system with FDA marketing authorization, billed in Medicare under HCPCS 92229 (“imaging of retina for disease detection, with automated review and report at point of care”). In

2024 Medicare paid for 2,633 such services by 69 practitioners across 8 specialties at a mean of \$27.5—a public record that autonomous AI already performs, and is paid for, clinical work.

Event. The coordinator’s request creates e with:

- R : CMS131 v13, “diabetic patient with retinal exam in the measurement period”;
- V : Condition E11.9 (signed by the EHR), Observation HbA1c 7.9% (signed by the laboratory), last exam date (signed by the eye clinic);
- A : consent κ to share records for treatment; grant γ for the imaging agent, scope {capture, grade, report}, owner the clinic medical director, expiring end of day;
- $D = \text{APPROVE}$: a_1 – a_3 hold and $\text{PTI}_{\text{agent}, 92229}^- = 0.81 \geq \theta = 0.75$;
- X : imaging performed and graded by the agent α_X ; referral to ophthalmology if positive;
- $O = 1$: the report is filed and the patient informed; attested by the reading record and the patient’s portal acknowledgment (v_4);
- C : v_1 .. v_4 pass;
- L : none;
- S : measure closed for the period; settlement of the contract value to the clinic, with the agent’s share attributed to its owner.

Transition. In epoch t , Φ multiplies the agent’s posterior for service 92229 by $\pi(q, r)$ for this patient’s risk r ; by Theorem 5.4 its trust can only rise. Λ debits the pool by $v_e(e)$ and credits $(H(e), \alpha_X, v_e(e))$ to $\mathcal{P}$. The commitment $H(e)$ joins the epoch’s Merkle tree. A payer, auditor or regulator can later verify the event from the anchor with 18 hashes and recompute $\mathbb{V}(e)$, without seeing the patient’s record.

What changed. No prior-authorization request was filed; no claim was adjudicated against documentation after the fact; the quality measure was computed from the event, not from a claim months later; and an autonomous agent was paid for verified work exactly as a person would be, under a grant owned by an accountable clinician.

8. Settlement, Treasury and Trust-Time

8.1. Settlement on proof

Definition 8.1 (Settlement). For a contract $\mathcal{C}$ mapping verified outcomes to values $v_e(e) \geq 0$, epoch settlement Λ is

$$\begin{aligned}\Pi &\leftarrow \Pi + C_t(1 - f) - \sum_{e \in \mathcal{E}_t'} v_e(e), \\ \mathcal{P} &\leftarrow \mathcal{P} \cup \{(H(e), \alpha_{X,e}, v_e(e))\}_{e \in \mathcal{E}_t'},\end{aligned}$$

where C_t is contributions, f the protocol fee and $\alpha_{X,e}$ the performing participant—human, organisation, agent or robot.

Proposition 8.2 (Value conservation). *In every epoch the total debited from Π equals the total credited in $\mathcal{P}$, and every credit references exactly one verified event hash.*

Proof. Both sides of Λ sum over $\mathcal{E}_t'$ with identical values; unverified events contribute to neither. $\square$

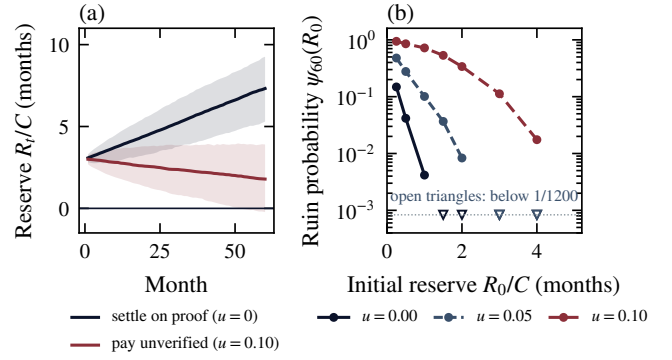


Figure 6. Pool solvency, 1,500 members, loading $\eta = 0.10$, fee $f = 0.02$. (a) Reserve in months of contributions (median, 5–95% over 2,000 paths). (b) 60-month ruin probability by initial reserve and unverified-payment rate u ; triangles, no ruin in 1,200 paths.

In words. Value leaves the pool only when an event is proven, and every unit paid points to the exact event it paid for, whoever—or whatever—performed it.

A robot that completes an authorized task with a verified outcome is credited exactly as a clinician would be; the credit is held by the accountable owner of its grant (Def. 3.2) under the operating agreement [Specified]. [Implemented] Reconciliation statements are generated from verified events on synthetic data; the reference implementation documents and recommends, it does not move funds. [Specified] Holding pooled funds for care requires a licensed structure or partner, and funds movement a licensed money-transmission partner; stablecoin rails are designed for and operate only through such partners. No patient information is written to any ledger.

8.2. Solvency

Let claims per epoch be compound Poisson $L_t = \sum_{j \leq N_t} Y_j$, $N_t \sim \text{Poisson}(\nu)$, $\mathbb{E}Y = \mu$, with contributions $C = (1 + \eta)\nu\mu$. If a fraction u of unverified claims is also paid, $R_{t+1} = R_t + C(1 - f) - (1 + u)L_t$.

Proposition 8.3 (Solvency condition). *The reserve has positive drift iff $(1 + \eta)(1 - f) > 1 + u$.*

Proof. $\mathbb{E}[R_{t+1} - R_t] = \nu\mu[(1 + \eta)(1 - f) - (1 + u)]$. $\square$

With $\eta = 0.10$ and $f = 0.02$, settlement on proof ($u = 0$) has drift +7.8% of expected claims per epoch; paying 10% unverified turns it to −2.2% and raises 60-month ruin from a three-month reserve from below $1/1200$ to 0.113 (Fig. 6). Medicare’s measured improper-payment rate (7.66%) is of the same order as u . Verification is the solvency mechanism of the treasury.

8.3. Participation and surplus

The pool is open to three roles: *members*, who contribute C_t^m to fund their own care; *depositors*, who provide capital K_t and share in surplus; and *producers*, any participant class of Definition 3.1, who are credited by Λ and neither contribute nor deposit. With reserve target R^* , the epoch surplus is

$$\Sigma_t = \max(0, R_t + C_t^m(1 - f) - \sum_{e \in \mathcal{E}_t'} v_e(e) - R^*), \quad (5)$$

distributed as $\omega\Sigma_t$ to members, as reduced contributions, and $(1 - \omega)\Sigma_t$ to depositors pro rata to K , with ω a governed parameter.

In words. The pool first pays for proven care and refills its reserve. Only what is left over is shared, part back to members and part to depositors.

Proposition 8.4 (Surplus requires verification). $\mathbb{E}[\Sigma_t] > 0$ in steady state only if the solvency condition of Proposition 8.3 holds; each unit of unverified payment u reduces expected surplus one-for-one.

Proof. By (5) and the drift of Proposition 8.3, $\mathbb{E}[\Sigma_t] \leq \nu\mu[(1 + \eta)(1 - f) - (1 + u)]^+$ once $R_t = R^*$, and the bracket decreases linearly in u . $\square$

The return to depositors is therefore not produced by leverage or a token price but by the claims the pool does not pay: verification is the yield. A share in the pool that pays a return is an investment contract and is offered only through a registered or exempt securities offering; paying members' care from the pool is risk pooling and requires an insurance licence, mutual or licensed carrier [Specified].

8.4. Tokens and the stablecoin

Settlement on L0 uses two instruments, carried from the Subnet and Treasury specifications [21] (working names) [Specified].

- **VVAL, the VANTHEON stablecoin.** A payment stablecoin issued by VANTHEON, redeemable one-to-one for U.S. dollars and backed by reserves held under the issuer rules of the GENIUS Act (2025). It is the unit in which C_t , K_t , $v_c(e)$ and Σ_t are denominated, so the value of care is never exposed to the price of a speculative asset. Becoming a permitted issuer is part of the protocol's own build, not a dependency on a third party.
- **VHTH, the network token.** Staked by validators and verifiers that run consensus and the checks $v_1..v_4$; paid as fees for verification and agent execution; used for governance of Θ under the rules of §19. A verifier whose attestation is later shown false loses stake, which adds an economic channel to the detection probabilities of §12.

Where the yield comes from. Three sources, and only three: (i) a depositor's pro-rata share $(1 - \omega)\Sigma_t$ of the pool surplus, paid in VVAL (§8.3); (ii) staking rewards and fees in VHTH for verification work; (iii) members' share $\omega\Sigma_t$ as reduced contributions. Holding VVAL itself pays nothing: U.S. law prohibits a permitted payment stablecoin issuer from paying holders "any form of interest or yield ... solely in connection with the holding" of the coin (GENIUS Act, 2025, §4(a)(11)). The yield is earned by funding verified care, not by holding money.

In words. Care is paid in a dollar stablecoin. People who deposit that stablecoin into the pool earn a share of what the pool does not lose to unproven claims. Validators earn the network token for checking the work.

Genesis distribution. VHTH is distributed at a genesis sale, following the pattern Ethereum set in 2014: a public sale of

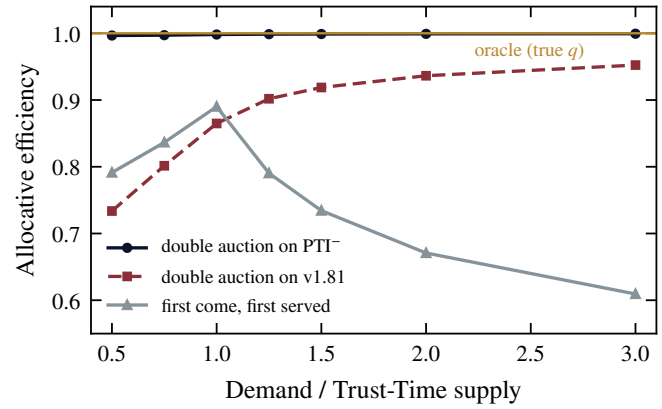


Figure 7. Trust-Time allocation: urgency-weighted expected success relative to the oracle allocation as demand varies (40 rounds per point).

fixed length, priced in an existing currency, documented in advance by the whitepaper, a development plan, an intended use of revenue and purchase terms, with an issuance cap and endowments for early contributors and a non-profit foundation stated before the first unit is sold [22]. Units are not usable until the genesis block. The parameters below are design targets for governance review, not commitments [Specified]:

Allocation of genesis supply	Share
Genesis sale	50%
Verification rewards (validators, verifiers, emitted over time)	20%
Foundation endowment	15%
Contributors (four-year vesting, one-year cliff)	15%

Sale proceeds fund the build and seed the pool reserve R^* ; they never fund the payment of care, which is paid only from contributions and deposits.

In words. The network token is first sold publicly, the way Ethereum did it: every rule is published before the sale, and nothing is usable until the network starts.

8.5. Trust-Time

The Gray Paper defines Trust-Time $\tau(a, t) = T_{avail}(a, t) \cdot PTI(a, t)$, verified capacity weighted by trust, cleared in a continuous double auction where bid b and ask a match when $b \geq a$ [21]. Ranking capacity by PTI^- recovers 0.999 of the oracle allocation at demand twice supply, against 0.936 for the earlier index and 0.671 for first-come-first-served (Fig. 7) [Research].

9. The Economy of Verified Work

Fee-for-service prices activity; value-based care prices populations, after the fact. The protocol prices the *verified event*, at the moment it closes, to the participant that produced it. This changes who can earn, what can be earned on, and what earning requires.

Entity-agnostic payment. Settlement (Def. 8.1) credits the performing participant a_X whatever its class. A clinician, a

community health worker, a clinic, an autonomous diagnostic system and a robot are paid by the same function for the same kind of proof. What differs is not the rule but the grant: a non-human participant acts and is credited under a grant whose owner is accountable for it, and it accumulates its own trust history $\mathcal{T}_{a,k}$. The 2024 Medicare data already record payment for autonomous AI work (HCPCS 92229, §7); the protocol generalises that precedent from one code to any service with a computable rule and an attestable outcome.

Earning trust is earning access. Because Rule 6.1 depends on $\text{PTI}_{a,k}^-$, trust is an asset: it converts directly into instant authorization, shorter cash cycles and eligibility for higher-value contracts. It cannot be bought, transferred or reset upward (Thm. 5.6), and it decays without new evidence ($\lambda > 0$). It is therefore a reputation that is costly to fake and valuable to keep, which is the condition under which honest behaviour is the equilibrium strategy (§12).

Price of an event. For a contract $\mathcal{C}$ the value of an event can be anchored on public fee schedules (e.g. the Medicare Physician Fee Schedule: $\text{RVU} \times \text{conversion factor} \times \text{geographic index}$) and adjusted by verified outcome, so that the price of the same service differs by what it achieved. In Porter’s terms, payment follows $\mathcal{V}_{a,k}$ rather than volume [Specified].

Where money moves. The reference implementation calculates and documents settlement; funds move only through licensed pools, carriers or money-transmission partners, and on stablecoin rails only through such partners [Specified]. The protocol defines what is owed and why; it does not require any participant to hold a token.

10. Security

10.1. Fraud, waste and abuse

The Medicare Program Integrity framework distinguishes *fraud* (intentional deception), *abuse* (practices inconsistent with sound practice) and *waste* (overuse without benefit). The protocol addresses each by construction: a claim of care that is not a verified event cannot settle (Prop. 8.2); fabricated evidence fails source signature (v_2); documentation inflation does not raise trust because trust conditions on outcomes, not documentation (Fig. 8a); upcoding fails the rule predicate (v_1); and services without benefit lower trust (Thm. 5.4). Table 4 gives the full model; §12 bounds what an attacker can gain.

10.2. Post-quantum resistance

Shor’s algorithm breaks the elliptic-curve and RSA signatures used by most ledgers today on a sufficiently large quantum computer [16]; hash functions lose only a square-root factor to Grover’s search [17]. The protocol therefore separates two layers. *Commitments*—event hashes and Merkle anchors—use SHA-256 or BLAKE3 at 256-bit output, which retain about 128-bit security against quantum search. *Signatures*—evidence sig_j , grants and settlement attestations—are specified with NIST’s post-quantum standards, ML-DSA (FIPS 204, lattice-based) as the default and SLH-DSA (FIPS 205, stateless hash-based) for long-lived keys, finalised in August 2024 [18]. Because event integrity rests on hashes and

Table 4. Threats, mechanisms and evidence. **T** theorem, **P** proposition, **F** figure.

Attack	Mechanism	Evidence
Phantom billing	no settlement without $\forall = 1$	P8.2, §12
Upcoding	rule predicate v_1 over signed evidence	§13, F13
Documentation inflation	trust conditions on outcomes	F8a
Cherry-picking	risk adjustment, eq. (3)	F8b, F3
Sybil re-entry	skeptical prior; credential-bound identity	T5.6, F8c
Collusive verification	independent path (A2); graph review	[Research]
Agent / robot misuse	scoped, expiring grants; deny-by-default	Def. 3.2
Automated denial	image of Auth	P6.2
Treasury drain	settle only $\mathcal{E}^\vee$	P8.3
Ledger tampering	hash anchoring (A1),(A4)	§16
Quantum adversary	PQ signatures, hash-based commitments	§10.2
Trust disclosure	zero-knowledge threshold proof	§11

the signature scheme is a versioned parameter in Θ , the protocol is cryptographically agile: a scheme can be replaced by a VIP without rewriting history. [Specified] The current reference implementation uses classical signatures (Ed25519 and secp256k1); migration to ML-DSA is VIP-0006. Zero-knowledge proofs of trust are treated in §11.

11. Zero-Knowledge Verification

The anchored log lets anyone verify that an event exists; it does not let them verify a participant’s trust without seeing the participant’s record. The Gray Paper proposed zero-knowledge proofs for exactly this gap [21]. We specify the statement, prove a sufficient certificate, and report a working circuit.

Statement. A participant proves to a payer, regulator or patient that it is eligible for instant authorization for service k —that $\text{PTI}_{a,k}^- \geq \theta_k$ —without revealing its tallies, its patients or its exact score. Public inputs: a commitment $C = \text{Poseidon}(s, f, t, \rho)$ to its verified successes s and failures f at epoch t with blinding ρ ; the Merkle root of the anchored batch; θ_k ; and the prior (α_0, β_0) . Witness: (s, f, ρ) and the Merkle path of C .

Lemma 11.1 (Integer certificate). *Let $n = \alpha_0 + \beta_0 + s + f$ and $\theta = \theta_n / \theta_d$. If $(\alpha_0 + s)\theta_d - \theta_n n \geq M$ with $M^2 \geq n\theta_d^2$, then $\mu - 2\zeta \geq \theta$, where μ and ζ are the mean and standard deviation of $\text{Beta}(\alpha_0 + s, \beta_0 + f)$.*

Proof. The hypothesis gives $\mu - \theta = ((\alpha_0 + s)\theta_d - \theta_n n) / (n\theta_d) \geq M / (n\theta_d) \geq 1/\sqrt{n}$. Since $\zeta^2 = \mu(1 - \mu)/(n + 1) \leq 1/(4n)$, $2\zeta \leq 1/\sqrt{n} \leq \mu - \theta$. $\square$

The certificate uses only integer arithmetic and no square root, so it is cheap to arithmetise. It is conservative: in 200,000 random tallies, it accepted 42,359 participants and 0 whose

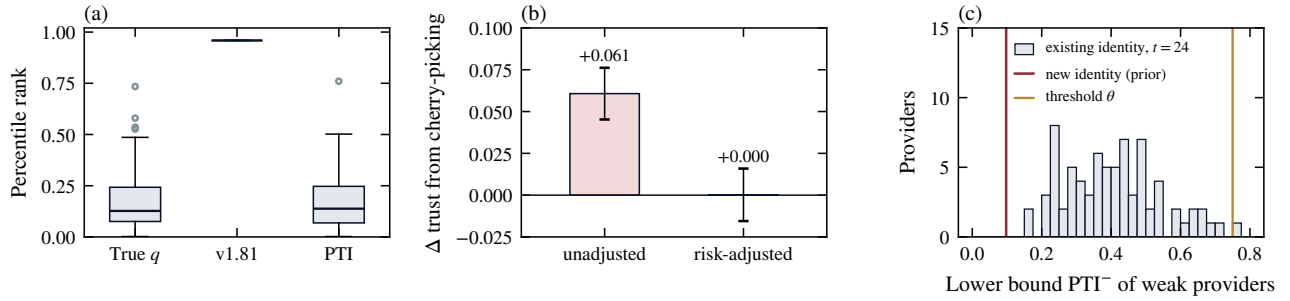


Figure 8. Adversarial strategies. (a) Percentile rank of gaming participants: true quality, earlier index, PTI. (b) Trust gained by declining the sickest third of patients (mean, 95% CI). (c) Lower bounds of weak participants against a fresh identity (red) and θ (gold).

exact 5% Beta bound was below θ , recovering 95.4% of those eligible under the exact rule.

Circuit. [Implemented] The statement is implemented as a Groth16 circuit in Circom over BN254: a Poseidon commitment, an 8-level Poseidon Merkle inclusion proof (a batch of 256 commitments), 20-bit range checks, and the two inequalities of Lemma 11.1. It has 5,055 constraints. On one laptop core, witness generation takes 54 ms, proving 505 ms and verification 298 ms; the proof is 256 bytes on chain (three group elements). A participant below threshold cannot produce a witness, and a proof checked against a modified commitment is rejected; both tests are part of the released suite.

Limits and the post-quantum path. Groth16 rests on pairings over elliptic curves; it is succinct and cheap to verify on Ethereum-compatible chains, but it is *not* post-quantum, and its setup ceremony here has a single contributor, sufficient for research and not for production. The Gray Paper’s choice of STARKs—hash-based, transparent, no trusted setup and plausibly post-quantum—remains the production target [Specified]: the same statement compiles to a STARK by replacing the backend, at the cost of larger proofs (tens to hundreds of kilobytes). Extending the statement from tallies to the full risk-adjusted posterior of Definition 5.1, and proving correct computation of the tallies from the anchored events themselves, is [Research].

12. Calculations

Bitcoin’s §11 computes the probability that an attacker, z blocks behind, ever catches up [1]. The analogous question here is: what is the probability that an attacker who submits z fabricated or inflated events ends with a net gain?

Each fabricated event is checked by m independent channels—source-signature verification (v_2), independent outcome attestation (v_4), patient attestation and retrospective audit—detecting it with probabilities $d_1, \dots, d_m$. It settles undetected with probability $u = \prod_j (1 - d_j)$ and pays v ; if detected it is clawed back and costs the attacker $c = (1 + \kappa)v$, where κv is the value of lost instant-authorization standing. With $U \sim \text{Bin}(z, u)$ undetected events the net gain is $G = vU - c(z - U)$.

Table 5. Probability that z fabricated events yield a net gain. Exact binomial tail; the last column is the bound of Theorem 12.1.

z	one channel ($u = 0.5$)	four channels ($u = 0.135$)	bound
1	0.5000	0.1350	0.3769
5	0.1875	0.0015	0.0076
10	0.0547	3.8×10^{-6}	5.8×10^{-5}
20	0.0059	3.4×10^{-11}	3.3×10^{-9}
50	0.0002	2.0×10^{-23}	6.5×10^{-22}
100	9.1×10^{-8}	2.1×10^{-45}	4.2×10^{-43}

Theorem 12.1 (Fraud bound). *Let $\beta^* = c/(v + c)$. If $u < \beta^*$, then $P(G > 0) \leq \exp(-z \text{KL}(\beta^* \| u))$, where $\text{KL}(x \| y) = x \ln \frac{x}{y} + (1 - x) \ln \frac{1 - x}{1 - y}$.*

Proof. $G > 0 \iff U > z\beta^*$. Since $\mathbb{E}U = zu < z\beta^*$, the Chernoff–Hoeffding bound for the binomial upper tail gives the result. $\square$

In words. Each additional fabricated case multiplies the attacker’s odds of coming out ahead by a constant less than one. Fraud at scale is not merely risky; its chance of paying off vanishes exponentially, as a forked chain’s chance of overtaking the honest one does in Bitcoin.

With conservative detection probabilities $d = (0.50, 0.60, 0.25, 0.10)$, $u = 0.135$ and $\kappa = 2$ (Table 5, Fig. 9). With a single check ($m = 1$, as in a pre-payment documentation review) the probability of profit after 20 fabricated events is 0.0059; with four independent channels it is 3.4×10^{-11} . Two caveats hold: the bound requires the channels to be independent (A2), and collusion between the performing and the attesting party breaks independence, which is why collusion detection is listed as open research.

13. Evidence from Public CMS Data

The protocol’s design choices respond to measurable failures of the current system. We analyse six public CMS datasets (Table 6); all files, hashes and retrieval dates are released with the code.

13.1. Improper payments fail checkable predicates

CMS’s Comprehensive Error Rate Testing program audits a random sample of Medicare claims. Of the 29,276 line-level improper-payment findings in the 2024 public file (excluding

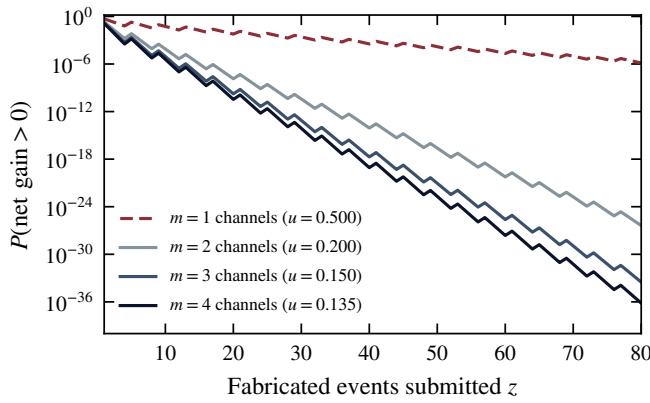


Figure 9. Probability of net gain from z fabricated events as independent verification channels are added (exact binomial; $\kappa = 2$).

Table 6. Public datasets used (data.cms.gov).

Dataset	Records
CERT 2024 Part A/B improper-payment review	185,349 lines
Hospital Readmissions Reduction Program FY2026	8,037 hosp.-cond.
HCAHPS patient survey (hospital)	3,949 hospitals
Medicare Part B by provider and service 2024 (E/M 99211–15, FP+IM)	267,170 rows
Complications and deaths; unplanned visits; MSPB	hospital level

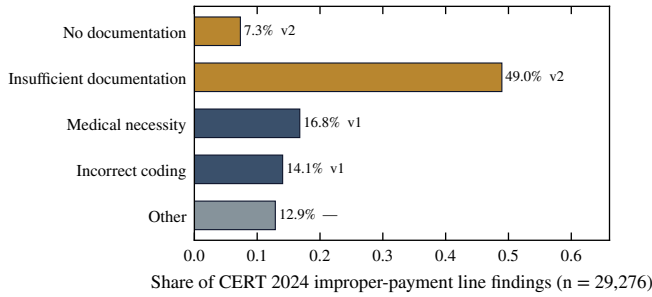


Figure 10. Causes of improper-payment findings in the CMS CERT 2024 public file, mapped to verification predicates (Def. 4.3).

findings later overturned), 56.3% are *no* or *insufficient documentation*, 16.8% *medical necessity* and 14.1% *incorrect coding* (Fig. 10). Missing documentation fails predicate v_2 ; medical necessity and coding are rule predicates v_1 over signed evidence. In total 87.1% of findings correspond to a predicate the protocol evaluates *before* settlement. Today they are found by sampling after payment; in VANTHEON the claim does not settle.

13.2. Point estimates misclassify performance

The HRRP reduces payments to hospitals whose excess readmission ratio (ERR) exceeds 1. Applying the protocol’s interval rule to the published discharges, readmissions and CMS’s own expected rates, 53.7% of 8,037 hospital–condition results have $ERR > 1$, but the 90% interval excludes the expected rate for only 8.0%; 85.2% of results above expectation are statistically indistinguishable from it, and the share falls with volume (Fig. 11). At hospital level, 81.9% of hospitals have at least

one condition above 1, against 20.2% under the interval rule. Split-half reliability of readmission performance is 0.061 in the lowest volume tercile (below 145 discharges) and 0.374 in the highest. These results echo the finding that HRRP penalties fell disproportionately on safety-net hospitals [14] and the debate over its effect on mortality [15]. A payment rule that acts on point estimates rewards and penalises noise; Rule 6.1 acts only on the lower bound.

13.3. Patients are an independent attestation channel

HCAHPS collects standardised patient-reported experience; the public file reports 2,411,406 completed surveys across 3,949 hospitals. The share of patients who would definitely recommend a hospital is associated with lower heart-failure excess readmission (Spearman $\rho = -0.188$, $n = 2,293$, $p = 1.1 \times 10^{-19}$; Fig. 12). Patient attestation is weakly but independently informative about outcomes, which is the property (A2) requires of a verification channel; it is included as the third channel of §12.

13.4. Documentation-driven intensity

Office visits for established patients are billed at levels 99213–99215 according to documented complexity. Among 29,935 Family Practice and Internal Medicine practitioners in 2024 reporting all three levels, the median share billed at 99215 is 9.3%; the 90th percentile is 31.5% and the 99th 64.0% (Fig. 13), at a mean Medicare payment of \$119 versus \$80.4 for 99214. Variation alone is not evidence of abuse: case mix differs. It shows that when payment follows documentation, intensity varies by a factor of 6.891 between median and 99th percentile, and the only check is retrospective sampling. In VANTHEON the level is a rule over signed evidence (v_1) and the trust a participant earns depends on outcomes, not on the level billed.

14. Learning: the Computed Learning Health System

The Gray Paper’s central claim was that learning should become *computational rather than institutional*. In the protocol, learning is the operator Ω acting on the rule set $\mathcal{R}$.

Definition 14.1 (Rule proposal). A proposal is $\ell = (R, R', \mathcal{E}_\ell, \Delta, \text{rollback})$: the current rule, the candidate version, the verified events that motivate the change, the estimated effect Δ on outcomes and cost with its interval, and the condition under which R' is withdrawn.

A proposal is applied only if stewards approve it and the lower bound of Δ on the primary outcome is non-negative; it is applied at an epoch boundary, kept with its predecessor, and monitored by the same estimator as participants. Learning is therefore subject to the same discipline as care: a change to the rules must prove itself.

Causal attribution. The Gray Paper proposed structural causal models and Pearl’s hierarchy to separate the effect of a coordination decision from the case mix that accompanies it [20]. The trust estimator of \$5 adjusts for measured risk; adjustment for unmeasured confounding—the condition under

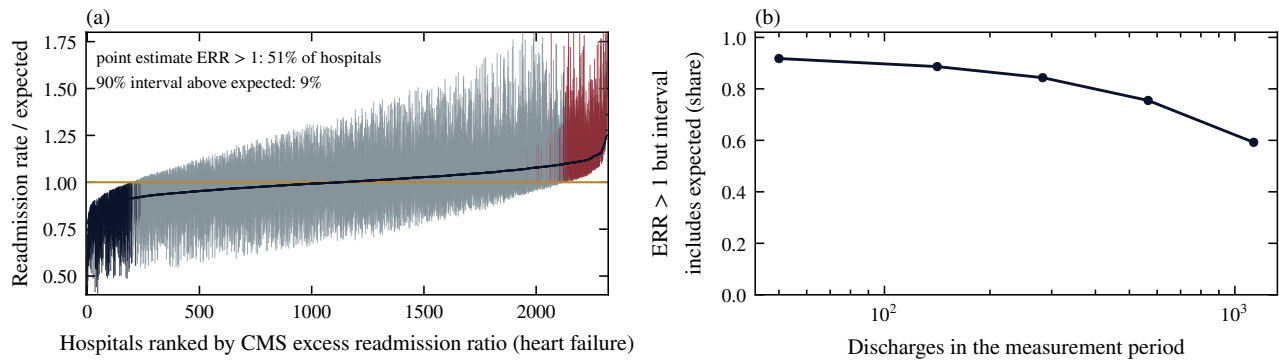


Figure 11. HRRP FY2026. (a) Heart failure: CMS excess readmission ratio (dots) and the 90% interval of the readmission rate relative to expected (lines; red, interval above expected; black, below; grey, includes expected) for 2,316 hospitals. (b) Share of all hospital-condition results with ERR > 1 whose interval includes the expected rate, by discharge volume.

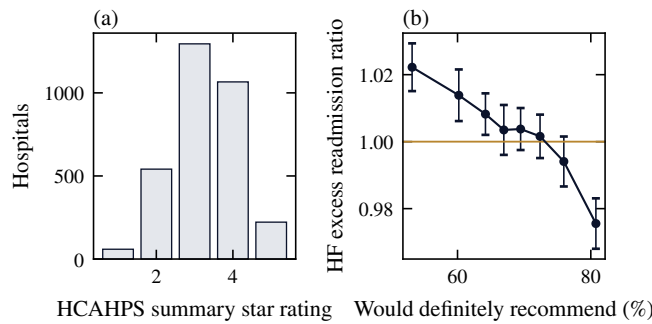


Figure 12. HCAHPS. (a) Distribution of summary star ratings. (b) Heart-failure excess readmission ratio by octile of “would definitely recommend” (mean, 95% CI).

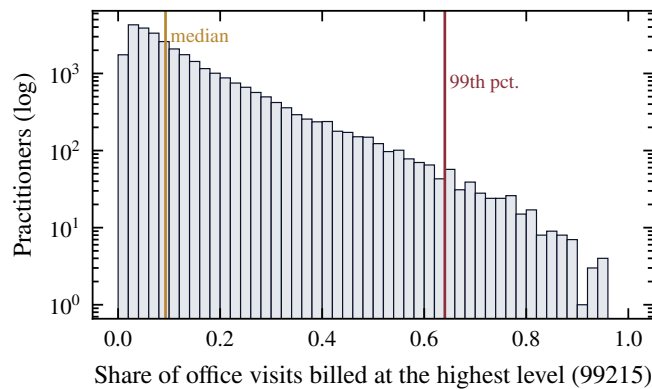


Figure 13. Medicare Part B 2024: share of established-patient office visits billed at 99215, per practitioner (FP and IM, all three levels reported; log scale).

which a learned change is causal rather than correlational—requires designs such as staggered roll-out of R' across sites, which Ω supports by applying proposals to a randomised subset of epochs or sites [Research].

Sovereignty. Participants keep control of their identity and history. Credentials are W3C verifiable credentials bound to decentralised identifiers; a participant can present its trust as a range proof (“ $\text{PTI}^- \geq \theta$ ”) rather than disclose its record [Spec-

ified]. Patients hold consent as a revocable grant; revocation propagates to v_3 on every subsequent event.

15. Applications

The protocol is a substrate. The following applications are specified on it; the first is in operation on synthetic data [Implemented], the rest are [Specified].

Clinical AI workspace. An agentic environment for clinicians, community health workers and quality coordinators: agents assemble evidence and draft requests; people decide. Every action is a verified event, so the workspace produces trust, quality measures and settlement statements as a by-product of work.

Instant authorization. Rule 6.1 replaces prior authorization for any service with a computable guideline, with denials reserved to people. It is a strict generalisation of “gold-carding” laws such as Texas HB 3459 (2021), which exempt clinicians with high historical approval rates; here the exemption is earned per service, adjusted for risk and revocable by outcome.

Quality and value-based contracts. Measures (eCQM, HEDIS, CAHPS) are computed from events at the moment of care, attributed to the participant who acted, and settled under the contract $\mathcal{C}$. Shared-savings and bundled-payment arrangements become contracts over verified events instead of annual reconciliations of claims.

Pools. Self-insured employers, community health programmes and member cooperatives fund care into a pool Π that pays only on proof. Proposition 8.3 and Fig. 6 show why verification is the condition of solvency.

Payment to non-human participants. An autonomous AI or a robot that performs authorized work with a verified outcome is credited per Definition 8.1. The settlement is held by the accountable owner of its grant; the participant accumulates its own trust history, which is what allows it to be trusted with more.

Research and public health. Because events are commitments with signed evidence, populations can be studied—for

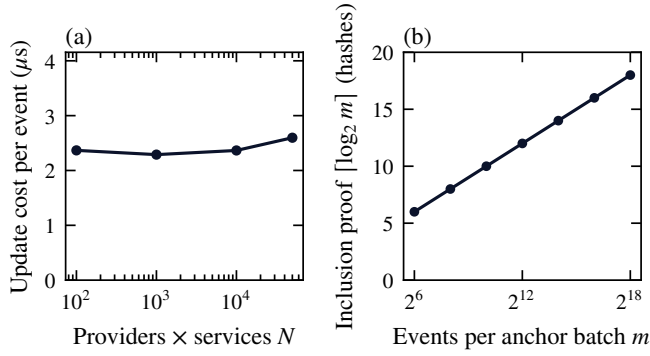


Figure 14. Scalability. (a) Measured trust-update cost per event. (b) Merkle inclusion-proof length per anchor batch.

Table 7. Executed invariant checks (seeded).

Invariant	Violations
I1 identical log $\Rightarrow$ identical state hash	0/20
I2 success never lowers trust (T5.4)	0/2,000
I3 no eligibility without evidence (T5.6)	0/1
I4 DENY $\notin$ im Auth (P6.2)	0/20,000
I5 pool debits = credits (P8.2)	0/2,000

effectiveness, safety or equity—without centralising records: the analysis is run where the data reside and its result is verified against the anchored commitments [Research].

16. Coherence and Scalability

Coherence. Every participant computing Y from the same log reaches the same state (Thm. 5.3); every credit references one event (Prop. 8.2); rules are changed only by Ω from a pre-epoch queue, so no participant can change the rule that judges an event after the event. Table 7 reports the invariants executed with the simulation.

Scalability. An event updates only $\mathcal{T}_{a_e, k_e}$, so trust computation is $\mathcal{O}(|\mathcal{E}_t|G)$ for a posterior grid of G points, independent of the number of participants. Measured per-event update cost is 2.289–2.597 μ s on one core from 10^2 to 5×10^4 participant–service pairs (Fig. 14a). Events are committed as a Merkle tree, one anchor per epoch; any event is proven with $\lceil \log_2 m \rceil$ hashes, 18 for $m = 262,144$ (Fig. 14b). Ledger load is $\mathcal{O}(1)$ per epoch; the Gray Paper’s target of 10,000 providers and 10^6 daily events corresponds to about 20 s of single-core trust computation per day under this model [Research].

17. Evaluation

Why simulate. Bitcoin supported its security claim with a calculation and a short program [1]; the Avalanche consensus paper with theorems and a measured deployment of 2,000 nodes [4]; the Ethereum whitepaper with worked examples and no simulation [2]. A protocol that has not yet run in the field must show that its mechanisms behave as specified under stated assumptions; we do so with a seeded simulation, and we ground the assumptions in the data of §13.

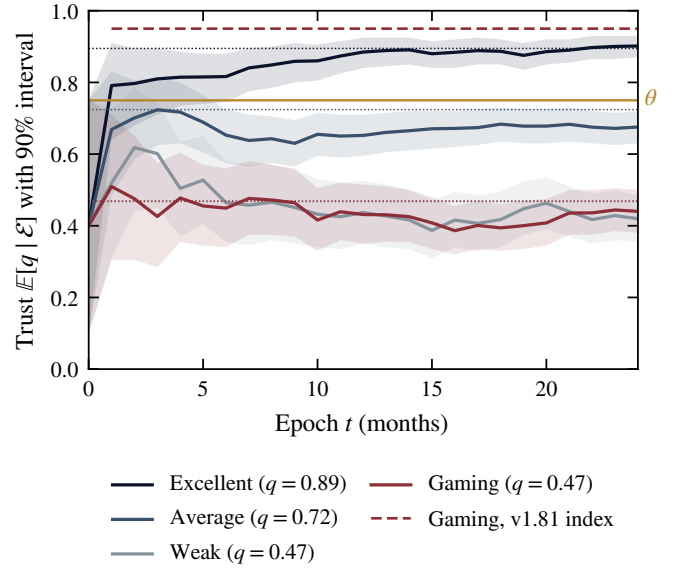


Figure 15. Trust trajectories for the median participant of four classes: posterior mean, 90% interval, true quality (dotted) and threshold θ (gold). Dashed red: the earlier index for the gaming participant.

Setup. 500 participants in five classes (excellent, average, weak, gaming, new entrant); Poisson event volume (mean ≈ 14 per month); patient risk $r \sim \text{Beta}(2, 5)$; outcomes from (3) with $\beta_r = 3$; $(\alpha_0, \beta_0, \lambda, \delta, \theta) = (2, 3, 0.04, 0.05, 0.75)$; 24 monthly epochs; seed 20261012.

Estimation. PTI tracks true quality within its interval while the earlier index stays flat (Fig. 15); correlation with true quality is 0.988 (mean absolute error 0.022) versus -0.273 (0.208), and 90% intervals cover the truth for 0.936 of participants (Fig. 16).

Fairness. Equal-quality participants with high-risk panels score 0.153 lower without risk adjustment and -0.002 with it (Fig. 3).

Adversaries. Gaming participants that report perfect documentation reach median percentile 0.959 under the earlier index and 0.138 under PTI, near their true 0.127; declining the sickest third of patients gains 0.061 unadjusted and 0.000 ± 0.016 adjusted; a new identity starts below every weak participant (Fig. 8).

Authorization. At $\theta = 0.75$ Rule 6.1 approves 30.6% of request volume instantly with false-approval rate 0.000; the same rule on the earlier index approves everything with false-approval rate 0.464. A new participant with $q = 0.9$ becomes eligible after a median of 3 months; those below θ never do within 36 months (Fig. 5).

Limitations. Simulation validates the estimator and mechanisms under the model, not clinical performance. Real outcomes are delayed, noisy and partly missing; risk models are imperfect; gaming is richer than simulated. Field evaluation precedes any use for payment.

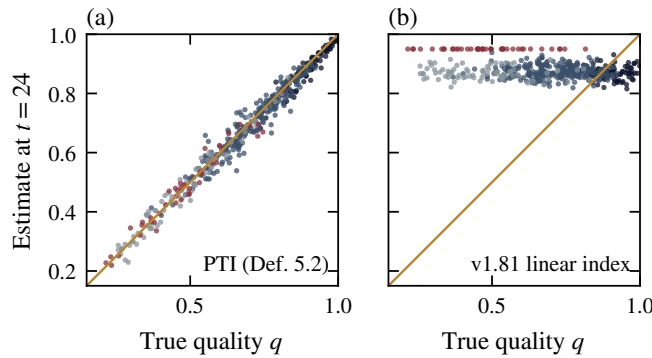


Figure 16. Estimated vs. true quality at $t = 24$ for 500 participants. (a) PTI; (b) earlier linear index.

18. Privacy, Consent and Law

VANTHEON is designed to operate under HIPAA (business associate agreements precede any patient data) and 42 CFR Part 2; CMS-0057-F [13]; CMS guidance on algorithmic coverage decisions and California SB 1120; the Anti-Kickback Statute and Stark Law wherever trust affects payment; state insurance law for any pooled-risk structure; and money-transmission and stablecoin law (the GENIUS Act, 2025) for funds movement. Only hashes and commitments are anchored on a public ledger.

19. Governance

Parameters $\Theta = (\alpha_0, \beta_0, \lambda, \delta, \theta_k, \rho_0, \eta, f, sig)$ and rules $\mathcal{R}$ change only through numbered VANTHEON Improvement Proposals (VIPs), each with a specification, a reference implementation and simulation results, approved by stewards and applied by Ω at an epoch boundary with the previous version retained. The Composite Governance Score and on-chain voting of the Gray Paper remain [Research].

20. Open Problems

A measure can be met while a patient is not helped; the verified event keeps *care given*, *measure met* and *health improved* distinct, but no score captures everything that matters. Rare services produce wide intervals and rely on people for longer. Outcomes arrive late; some never return. Trust can concentrate work among few participants, so access for new entrants must be protected. Collusion between performing and attesting parties, causal attribution of rule changes, privacy-preserving proofs of trust, and the legal form under which non-human participants hold settlement remain open.

21. Conclusion

Bitcoin made the ordering of payments verifiable and removed an intermediary. Ethereum made agreements executable. VANTHEON makes the complete event of care verifiable, so that trust can be computed, care authorized at the moment it is needed, value settled on proof to whoever produced it, and every result made into the next decision. Porter defined value as outcomes per dollar; this protocol is the machine that computes it. The future of healthcare will be programmable.

The question is who writes its rules, who owns what it learns and who claims the value it creates. We have specified our answer; the next step is to measure it in the field.

Reproducibility. Code and data: `research/whitepaper/(sim/fetch_cms.py, empirical.py, calculations.py, figures.py)`; dataset hashes in `data/cms/SHA256SUMS`.

References

- [1] S. Nakamoto. Bitcoin: a peer-to-peer electronic cash system. 2008.
- [2] V. Buterin. Ethereum: a next-generation smart contract and decentralized application platform. 2014.
- [3] G. Wood. Ethereum: a secure decentralised generalised transaction ledger (Yellow Paper). 2014–.
- [4] Team Rocket, M. Yin, K. Sekniqi, R. van Renesse, E. G. Sirer. Scalable and probabilistic leaderless BFT consensus through metastability. arXiv:1906.08936, 2019.
- [5] G. Wood. Polkadot: vision for a heterogeneous multi-chain framework. 2016.
- [6] M. E. Porter, E. O. Teisberg. *Redefining Health Care: Creating Value-Based Competition on Results*. Harvard Business School Press, 2006.
- [7] M. E. Porter. What is value in health care? *N. Engl. J. Med.* 363:2477–2481, 2010.
- [8] M. E. Porter, T. H. Lee. The strategy that will fix health care. *Harvard Business Review* 91(10), 2013.
- [9] W. H. Shrank, T. L. Rogstad, N. Parekh. Waste in the US health care system: estimated costs and potential for savings. *JAMA* 322(15):1501–1509, 2019.
- [10] D. U. Himmelstein, T. Campbell, S. Woolhandler. Health care administrative costs in the United States and Canada, 2017. *Ann. Intern. Med.* 172(2):134–142, 2020.
- [11] Centers for Medicare & Medicaid Services. Fiscal year 2024 improper payments fact sheet. 2024.
- [12] KFF. Medicare Advantage insurers made nearly 53 million prior authorization determinations in 2024. 2025.
- [13] Centers for Medicare & Medicaid Services. Interoperability and Prior Authorization Final Rule (CMS-0057-F). 2024.
- [14] K. E. Joynt, A. K. Jha. Characteristics of hospitals receiving penalties under the Hospital Readmissions Reduction Program. *JAMA* 309(4):342–343, 2013.
- [15] R. K. Wadhera, K. E. Joynt Maddox, J. H. Wasfy, et al. Association of the Hospital Readmissions Reduction Program with mortality among Medicare beneficiaries. *JAMA* 320(24):2542–2552, 2018.
- [16] P. W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. *Proc. 35th FOCS*, 124–134, 1994.
- [17] D. J. Bernstein, T. Lange. Post-quantum cryptography. *Nature* 549:188–194, 2017.
- [18] NIST. FIPS 203, 204, 205: post-quantum cryptography standards. August 2024.
- [19] Institute of Medicine. *The Learning Healthcare System: Workshop Summary*. National Academies Press, 2007.
- [20] J. Pearl. *Causality: Models, Reasoning, and Inference*, 2nd ed. Cambridge University Press, 2009.

- [21] A. S. Méndez Valentín. VANTHEON Gray Paper: the Computed Learning Health Protocol, v1.67 and v1.81. x/acc labs, 2025.
- [22] V. Buterin, “Launching the ether sale,” Ethereum Foundation Blog, Jul. 22, 2014.

A. From the Gray Paper to this Whitepaper

The Gray Paper (v1.67, October 2025; v1.81) is the architectural source of this work. Table 8 records how each of its claims is carried forward. Claims that could not be verified against a primary source or the reference implementation are removed rather than softened.

B. Data and Methods

Retrieval. All hospital files were downloaded from the CMS Provider Data Catalog (data.cms.gov/provider-data) through its metastore API, which records each file’s identifier and last-modified date: HCAHPS–Hospital (`dgck-syfs`), Complications and Deaths–Hospital (`ynj2-r877`), Unplanned Hospital Visits–Hospital (`632h-zaca`), Hospital Readmissions Reduction Program FY2026 (`9n3s-kdb3`), Medicare Spending per Beneficiary (`5hk7-b79m`) and Hospital General Information (`xubh-q36u`). The CERT 2024 Part A/B public file and Medicare Physician & Other Practitioners—by Provider and Service, 2024, were retrieved from data.cms.gov; the latter through its data API, restricted to HCPCS 99211–

99215 for Family Practice and Internal Medicine and to HCPCS 92229 for all specialties. SHA-256 hashes of every file are released.

CERT. Lines with review decision *Disagree* are improper-payment findings; lines whose error code is *Overturned* were reversed and are excluded. Error codes are mapped to predicates as follows: *No documentation* and *Insufficient documentation* to v_2 ; *Medical necessity* and *Incorrect coding* to v_1 ; *Other* is not mapped. Shares are of line-level findings, not of dollars; the public file does not report payment amounts per line.

HRRP. For each hospital–condition row with numeric discharges n and readmissions y , and CMS expected rate $\hat{p}$, we place a Beta prior on the non-readmission rate centred on $1 - \hat{p}$ with strength 50 discharges and compute the 90% posterior interval of the readmission rate. A result is *above expected* if the whole interval exceeds $\hat{p}$, *below* if it lies under it, and *indistinguishable* otherwise. “ERR > 1” uses CMS’s published excess readmission ratio. Rows reported as “Too Few to Report” are excluded. The analysis uses CMS’s own risk adjustment through $\hat{p}$; it does not re-estimate it. Split-half reliability thins each row’s discharges binomially into halves, draws readmissions hypergeometrically and reports the Spearman correlation of observed/expected between halves, by volume tercile.

HCAHPS. “Would definitely recommend” is measure H_RECMMND_DY (answer percent); star ratings are H_STAR_RATING. Hospitals are matched to HRRP heart-failure rows by CMS certification number.

Part B. CMS suppresses provider–code rows with fewer than 11 services. To avoid bias from suppression, the 99215 share is computed only for practitioners reporting all of 99213, 99214 and 99215 with at least 11 services each. Mean payments are service-weighted averages of *Avg_Mdcr_Pymt_Amt*.

Value. Porter value is computed as $(1 - m)/c$ with m the heart-failure 30-day mortality rate (MORT_30_HF) and c the hospital’s MSPB ratio (MSPB-1). This is a hospital-level proxy for illustration, not a condition-level cost measurement.

C. Simulation Protocol

Table 8. Reconciliation of Gray Paper claims.

Gray Paper claim	Status	In this paper
Seven-layer architecture L0–L6	Kept, recast	Layers become the stages of one transition function Y (eq. 2): L1 evidence $\rightarrow v_2$; L2 attribution $\rightarrow \mathcal{P}$; L3 Codex governance $\rightarrow \Omega$; L4 agents $\rightarrow$ grants γ ; L5 treasury $\rightarrow \Lambda, \Pi$; L0 $\rightarrow$ anchoring (§16).
CLHP = (A, T, C, E, L)	Kept, made precise	A: grants and sovereignty; T: Def. 5.1; C: coherence, §16; E: Ω with proposals; L replaced by Thm. 5.5.
PTI = $\alpha_1 OA + \alpha_2 DT + \alpha_3 MF + \beta He^{-\lambda \Delta t}$	Replaced	Outcome-conditioned Bayesian index (Def. 5.1); components retained as explanations (§5.4).
CCS, CXS	Kept	Explanatory scores, not decision inputs.
Trust-Time $\tau = T^{avail} \cdot PTI$, double auction	Kept	§8; allocation evaluated (Fig. 7); price formation [Research].
Appendix A.1, “logarithmic convergence” of PTI	Corrected	The proof bounded the decay of an EMA, not convergence to truth; replaced by Thm. 5.5 ($\Theta(n^{-1/2})$).
Appendix A.2, BFT $f < n/3$ inherited from Avalanche	Corrected	Avalanche’s guarantees are probabilistic and parameterised by its sampling threshold, not the classical $n/3$ bound [4]; stated as assumption (A1).
ZK-STARK proofs of PTI, “quantum-resistant”	Kept, built in part	Threshold proof implemented in Groth16 (§11); STARK backend [Specified]; full-posterior proof [Research].
SLOs: L0 finality < 1 s, L5 clearing < 2 s, L6 < 200 ms	Restated	Unmeasured targets removed; measured per-event cost and proof length reported (Fig. 14).
10,000 providers, 10^6 events/day	Kept as target	≈ 20 s/day of trust computation under the model [Research].
“43% of eligible providers do not join HIEs (2023)”	Removed	Source not identified.
“Barnett et al. 2019: referrals uncorrelated with outcomes”	Removed	Citation incomplete; replaced by public CMS analyses (§13).
Mandel & Valentín (2023), Tokenomics 2023	Removed	Not found in Crossref or DBLP.
“over 85% of EHRs support FHIR (2024)”	Removed	Source not identified.
VANTHEON as a QHIN under TEFCA	Removed	Designation requires ONC/RCE approval; not claimed.
Permissionless participation	Qualified	Participation is open to any credentialed participant (Def. 3.1); clinical actions remain licence-bound.

Element	Specification
Participant classes	excellent 20%, $q \sim \text{Beta}(18, 2)$; average 45%, $\text{Beta}(12, 5)$; weak 15%, $\text{Beta}(5, 6)$; gaming 10%, $\text{Beta}(5, 6)$ with perfect reported documentation; new entrant 10%, $\text{Beta}(12, 5)$
Volume	events per month $\sim \text{Poisson}(\mu_a)$, $\mu_a \sim 2 + \text{Gamma}(4, 3)$
Patients	risk $r \sim \text{Beta}(2, 5)$; fairness test adds +0.25 to half the panels
Outcomes	$O \sim \text{Bernoulli}(\pi(q, r))$, $\beta_r = 3$
Estimator	exact posterior on a 199-point grid; $(\alpha_0, \beta_0) = (2, 3)$, $\lambda = 0.04$, $\delta = 0.05$
Earlier index	$(0.4 OA + 0.3 DT + 0.2 MF)/0.9$ with reported inputs, as in the reference implementation
Authorization	$\theta = 0.75$; “unsafe” if $q < 0.7$; volume-weighted rates
Pool	1,500 members; claims compound Poisson, frequency 0.08, lognormal severity ($\sigma = 1.1$); loading 0.10; fee 0.02; 60 months; 2,000 paths (1,200 for ruin curves)
Trust-Time	capacity 4 slots per participant; urgency $\text{Beta}(2, 3)$; 40 rounds per demand level
Scalability	20,000 events per measurement, median of 5; Merkle tree over SHA-256
Seeds	base seed 20261012, fixed offsets per experiment

D. Version History

Version	Date	Change
Gray Paper v1.67	Oct 2025	Seven-layer architecture, PTI/CCS/CXS, Trust-Time, treasury
Gray Paper v1.81	2025	ZK-STARK and sovereignty chapters
Whitepaper v1.0	Oct 2026	One primitive and transition function; proofs; CMS evidence
Whitepaper v1.1	Oct 2026	Zero-knowledge threshold proof; Porter value; title “A Protocol for Verifiable Healthcare”; open participation and surplus (§8.3); tokens and stablecoin (§8.4)

E. Notation

$e =$ ($R, V, A, D, X, O, C, L, S$)	verified event (Def. 3.3)
α ; class	participant; clinician, organisation, agent, robot, patient, pool, steward
$\sigma = (\mathcal{W}, \mathcal{J}, \Pi, \mathcal{P})$	protocol state
$\Upsilon; \Delta_\lambda, \Phi, \Lambda, \Omega$	transition; decay, trust, settlement, governance
$\mathbb{V}; v_1..v_4$	verification predicate; rule, evidence, authority, outcome
$q, r, \pi(q, r)$	true quality, patient risk, outcome model
$\mathcal{J}_{a,k}; \text{PTI}, \text{PTI}^-$	trust posterior; mean, lower δ -bound
θ_k	instant-authorization threshold
$\mathcal{V}_{a,k}$	Porter value, eq. (1)
γ, κ	grant, consent
Π, C, f, η, u	pool, contributions, fee, loading, unverified rate
K, R^*, Σ_t, ω	deposits, reserve target, surplus, member share
VVAL, VHTH	stablecoin; network token (§8.4)
$\tau(a, t)$	Trust-Time
u, κ, z	undetected rate, standing penalty, fabricated events (§12)
H, <i>sig</i>	hash; signature scheme (ML-DSA / SLH-DSA)