From lagging to leading: How one health plan drove $13.4M in inpatient savings without sacrificing quality

By combining AI-powered clinical decision support with continuous review, quality analytics, and auditing, a regional health plan (~1M members) transformed its inpatient utilization management program, improving decision accuracy and consistency while saving $13.4M in annualized medical expenses.

The Challenge

For this health plan, the complexity of the source material itself was central to the challenge. Inpatient medical records often span 50 to 150 pages, inconsistently formatted, filled with repetitive and sometimes conflicting documentation. A single readmission review–calculating date gaps, confirming facility and condition matches, checking policy exclusions, and parsing two full clinical records for avoidability–could routinely consume 30 to 45 minutes. The plan's own leadership acknowledged that the core issue wasn't process speed. The real question was whether they had the right technology to make consistent, defensible clinical decisions at scale.

Inpatient utilization management wasn't failing for lack of review–it was failing for lack of consistency. Before partnering with Cohere Health, the plan had no AI-powered clinical decision support for inpatient reviews. The process followed a standard UM workflow, in which nurses made different calls on similar cases, resulting in inconsistent outcomes among reviewers.

Two opportunities emerged to help define the approach to this challenge:

> **Almost nothing was consistently downgraded to a lower level of care when appropriate.** The clearest signal was in admission-to-observation (A2O) conversion cases, where a nurse determined that a lower level of care was clinically appropriate. Industry benchmarks for comparable populations suggest an expected conversion rate of roughly 10%. This plan was converting closer to 1%, reflecting inconsistent application of clinical policy. In practice, that inconsistency meant nearly every case was approved as submitted, regardless of whether the clinical evidence fully supported it.

> **Preventable readmissions were slipping through the cracks.** A second, related gap involved avoidable readmissions–specifically, 30-day readmissions to the same facility for the same or similar condition, where more thorough discharge planning could reasonably have prevented a return visit. Identifying these cases is manual, tedious work: reviewers must cross-reference clinical details across multiple admission records–often checking up to 7 distinct data points across two full inpatient charts–before they can even apply the readmission policy. The result was a two-fold problem–readmissions that should have been flagged were being missed entirely, while those caught were subject to inconsistent policy application.

COHERE'S OPPORTUNITY

Improve consistency and accuracy of inpatient-to-observation level-of-care decisions.

Improve identification and correct policy application for avoidable readmissions.

Reduce administrative burden, so nurses and clinicians can focus on clinical judgment, not paperwork.

The Solution

The plan partnered with Cohere Health to introduce two things simultaneously: an AI tool for review nurses, and an AI system to continuously check the nurse's work.

AI-powered clinical decision support built into the review workflow

In-app AI guidance provides nurses with case-overview support, review note generation, and a clinical evaluation tool that analyzes the records against medical policy. A purpose-built readmissions workflow automates the full review–from date-gap calculation to policy application–resolving in minutes what once required manual cross-referencing of multiple clinical records. AI is changing how nurse reviewers navigate inpatient records—helping them quickly locate and synthesize the clinical information they need, improving efficiency while supporting high-quality decision-making and documentation.

Auditing every case for ongoing quality assurance

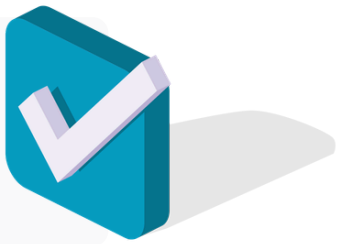
Alongside its AI tooling, Cohere Health audits every case–not just a sample–comparing the AI's recommendation to the nurse's clinical review and resulting case outcome. Any disagreements are flagged for manual audit, with insights fed directly into nurse and MD education. That's a level of oversight that manual audit programs, typically limited to a handful of cases per week, simply can't match.

Cohere Health also tracks whether decisions hold up through peer-to-peer review and appeals. A wide gap between gross and net conversion rates usually signals decisions aren't standing up to scrutiny. For this plan, that gap has stayed tight, meaning the improvement reflects more accurate decisions, not just more denials.

THE COHERE DIFFERENCE

AI-powered decision support

In-app clinical guidance for case review, evidence-to-policy evaluation, and automated readmission workflows–so nurses can quickly find the specific clinical documentation required to support decision-making.

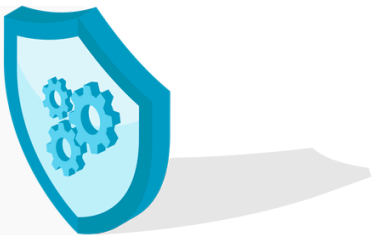


Continuous quality auditing

Every case reviewed for AI/nurse concordance, with discordant cases manually audited and insights fed directly into ongoing nurse and MD education.

Durable, defensible decisions

Outcomes are tracked beyond the initial determination, through peer-to-peer reviews and appeals, to confirm that improvements reflect accuracy, not just a shift in approval rates.

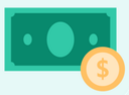


The Results

Measured across the 6 months before vs. the 6 months after implementation:

Metric	Before	After	Change	Why This Matters
Cases escalated to physician review	10%	25%	+15 pts.	A modest rise suggests a more appropriate escalation of complex cases, rather than blanket referral.
MD concordance with nurse A2O recommendations	30%	60%	+30 pts.	Physician agreement with nurse recommendations improved sharply, reflecting more consistent, defensible decision-making.
Inpatient-to-observation conversion rate	1%	10%	+9 pts.	The gap was consistency, not case mix–this plan now converts cases in line with the industry benchmark for comparable populations.
Avoidable-readmission impact rate	3%	5%	+2 pts.	Reflects the readmissions workflow, catching and correctly adjudicating cases nurses previously missed.
Post-decision overturn rate	4%	4%	No change	Held flat–evidence the accuracy gains came from more precise decisions, not more aggressive ones.

In just six months, those gains translated directly into medical expense savings–and against the plan's broader goals, the impact goes even further:



$13.4M
in annualized incremental medical expense savings, in just six months

>50%+
of this plan's total ~10M vendor-wide savings target for the year, beating Cohere Health's own internal projections

By pairing AI-powered clinical decision support with continuous, case-level quality assurance, Cohere helped this health plan move from inconsistent, under-benchmark inpatient decisions to a faster, more consistent, and more defensible program–with the savings to show for it.

See how Cohere Health combines AI-powered clinical intelligence with AI decision support and continuous quality auditing to strengthen inpatient utilization management.

Visit: **www.coherehealth.com/utilization-management-suite**