

# The Financial Plumbing: Claims, Systems, and Incentives

## Study Guide

by Richard J. Young, Ph.D.

[richard@deepneuro.ai](mailto:richard@deepneuro.ai)

Chapter 1 of "Healthcare Analytics and AI: Building Systems That Actually Work"

*The essentials: what to know, why it matters at the bedside, and how models fail.*

**One-sentence thesis.** In U.S. healthcare the billing system is not background infrastructure, it is the operating environment every deployed AI system has to survive, and whoever writes the check defines the model's objective function.

## PART 1: The Claim Is the Data

### The 837 and 835

Before any model, ask where the data came from. Almost every dollar in U.S. healthcare moves through two HIPAA-mandated EDI transactions, and the data you train on is their residue. The 837 is the electronic claim, the invoice that says here is what we did, how we coded it, and what we want paid. The 835 is the remittance advice, the payer's marked-up answer listing what it allowed, reduced, denied, and what the patient owes. In symbols, the money makes a round trip:

Provider  $\xrightarrow{837}$  Payer  $\xrightarrow{835}$  Provider

Between these two transactions sits the entire provider-payer financial relationship, and most of the data healthcare AI trains on. The point builders miss first is that claims data records what was billed, not what clinically happened: diagnosis codes are the codes that maximize defensible reimbursement, not a physician's best judgment, and confusing billing data with clinical truth is the first mistake most AI teams make.

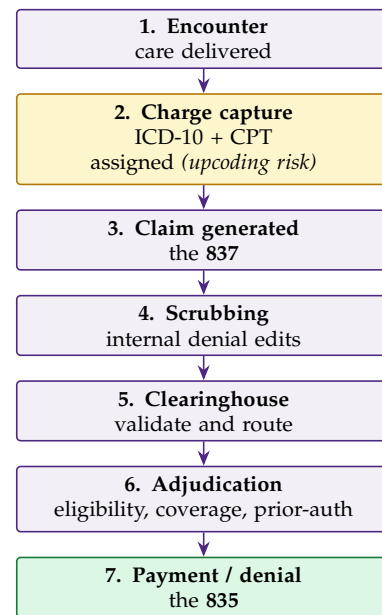
#### ANALOGY

The 837 is a restaurant bill you hand the insurance company; the 835 is that bill returned with half the line items crossed out and a note explaining why.

### The Claim Lifecycle

Seven steps turn care into money: (1) the encounter, (2) charge capture, where a coder or AI assigns ICD-

10 and CPT codes, (3) claim generation, the 837, (4) scrubbing, where internal edits flag likely denials, (5) clearinghouse transmission, (6) adjudication, where the payer runs eligibility, coverage, prior-auth, and medical-necessity checks, and (7) payment or denial, the 835. Each check is increasingly automated by AI, and the whole thing moves like a parcel through a sorting facility, scanned, validated, and routed at every belt with a chance of being kicked off the line at any point. **Watch out.** Charge capture is where upcoding risk lives, so the dataset that lands in your training pipeline is the residue of these seven decisions, each shaped by financial incentive.



*The seven-step claim lifecycle. Care becomes money through the 837 (step 3) and returns as the 835 (step 7); your training data is the residue of all seven decisions.*

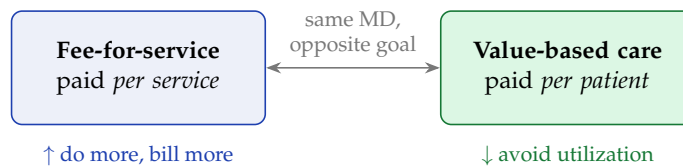
### Fee-for-Service vs. Value-Based Care

Under fee-for-service (FFS), providers are paid per service, so the incentive is to do more and bill more. Under value-based care (VBC), payers pay a fixed amount per patient (capitation), a bundled payment, or shared savings, so the incentive inverts to keeping patients healthy and avoiding utilization. The payment model, not the technology, determines what AI gets funded: FFS funds coding, CDI, and scheduling

tools that capture revenue, while VBC funds risk stratification, readmission predictors, and SDOH models that avoid cost. Before you write code, ask under what payment model the AI will operate.

#### ANALOGY

Fee-for-service is a taxi meter that rewards the long route; value-based care is a flat-rate delivery contract that rewards getting there efficiently.



*The incentive inversion. The payment model, not the technology, decides which AI gets funded: revenue capture under FFS, utilization avoidance under VBC.*

### The Hybrid Reality

Most large systems run both models at once: some patients on FFS Medicare, some on Medicare Advantage, some commercial, some Medicaid. The same physician on the same day is rewarded to do more for one patient and less for another. A readmission-reduction model saves money under VBC but can reduce revenue under FFS, and this dual-incentive tension, not the technology, is why adoption is slow and fragmented. A tool must show value under both structures or it gets adopted by some departments and ignored by others in the same building.

### The Clearinghouse

A clearinghouse is the intermediary that validates, formats, and routes claims between roughly 900,000 provider organizations and 1,000-plus payers, a combined postal sorting hub and switchboard for health-care payments. Change Healthcare (Optum, acquired 2022 for \$13 billion) processes about 15 billion transactions a year, touching an estimated one in three patient records. That gives clearinghouse data cross-payer visibility no single EHR has, but it is billing data, not clinical data. The stakes of that concentration became clear when the February 2024 ransomware attack on Change Healthcare compromised an estimated 192.7 million records and froze national claims processing for weeks, one point of failure cascading across the whole system.

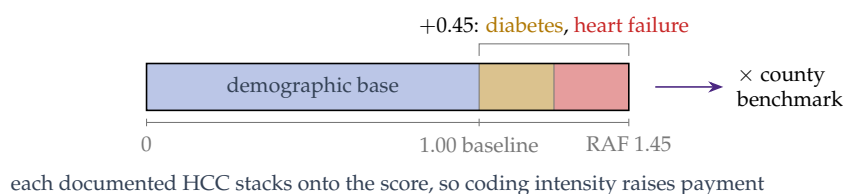
## The Four Parts and HCC Risk Adjustment

Medicare has four parts: A (hospital), B (physician), C (Medicare Advantage), and D (drugs). Traditional Medicare (A and B) pays fee-for-service, while Medicare Advantage (Part C, now **54 percent** of the 67 million beneficiaries) pays private plans a risk-adjusted per-member-per-month capitation. The question every builder should ask is what raises that monthly check, and the answer is documented diagnoses. In symbols, the payment is a risk score times a local benchmark:

MA payment = RAF × county benchmark

RAF = demographic base +  $\sum$  HCC weights

The risk score sums a demographic baseline with Hierarchical Condition Category (HCC) weights; diabetes plus heart failure can yield a score near 1.45, meaning CMS pays 45 percent above baseline. Because every documented diagnosis raises the risk score and the payment, HCC coding optimization is one of the largest internal AI applications at every MA plan, and it draws federal fraud liability: in January 2026 Kaiser Permanente paid **\$556 million** to settle False Claims Act allegations of inflated risk scores, and MedPAC estimates \$84 billion per year in excess MA payments, much of it from coding intensity.



*HCC risk adjustment. The RAF sums a demographic base with condition weights, then multiplies the county benchmark: about 45 percent above baseline here.*

## Star Ratings and Dual Eligibles

CMS rates each MA contract 1 to 5 stars on about 40 quality measures. Reaching 4 stars unlocks a Quality Bonus Payment worth 5 percent of the county benchmark, 10 percent in double-bonus counties; for a 500,000-member plan, moving from 3.5 to 4.0 stars is worth roughly \$300 million a year. But Star Ratings reward closing measurable care gaps, not necessarily

caring for the sickest patients, so quality-measure AI optimizes a proxy that can diverge from health. The problem sharpens for dual eligibles, the 12.5 million people on both Medicare and Medicaid who are 19 percent of Medicare enrollees but 35 percent of Medicare spending: they span two data systems with different formats and identifiers, so a model trained only on Medicare claims misses Medicaid-covered services and fails for exactly the highest-need population.

### ■ The \$300 Billion Waste Pile

A 2019 JAMA study put U.S. healthcare waste at about 25 percent of spending, with administrative complexity the single largest category at \$265.6 billion, likely over **\$300 billion** today, six times the entire NIH research budget spent every year on paperwork that treats no one. Prior authorization is the most-cited burden: staff spend about 13 hours a week on it, 93 percent of physicians say it delays care, and 34 percent report it caused a serious adverse event. This kind of waste is high-volume, rule-based work on structured data, exactly what ML and NLP do well, and CFOs have budget authority, which is why revenue cycle and administrative AI attract the most venture capital. The wins are real: Hackensack Meridian's prior-auth automation cut turnaround from 15 days to 1 to 2 days across 1,200-plus clinicians, and ambient scribes

like DAX Copilot (in 400-plus systems) save 5 to 7 minutes per encounter.

### ■ RCM and the Upcoding Line

Revenue cycle management AI suggests codes before submission (computer-assisted coding), prompts physicians to document higher-acuity codes (clinical documentation improvement), predicts denials, and prioritizes appeals by dollar value and overturn probability. **Watch out.** The line between capturing revenue legitimately owed and upcoding, assigning a higher code than documentation supports, is not bright. A CDI tool tuned only to maximize severity becomes an upcoding engine with a physician's signature laundering its output, exposing the organization to False Claims Act liability. Build guardrails, auditability, and physician-override tracking.

### The Algorithmic-Denial Case

MA capitation gives plans a direct incentive to deny or delay care. A 2023 STAT News investigation reported that several major Medicare Advantage plans used algorithms, including one developed by NaviHealth (an Optum subsidiary), to predict how long post-acute patients should need skilled nursing care; when the predicted timeline expired, coverage was denied regardless of the patient's actual status. Case managers overrode the algorithm less than 2 percent

of the time, which made its output a de facto coverage decision. 2024 CMS guidance now says MA organizations cannot use an algorithm as the sole basis for a medical-necessity denial and that human clinical review is required, but enforcement is inconsistent and the financial incentive is unchanged. When the payer's objective (spend less) conflicts with the patient's (get care), the AI optimizes for whoever paid for it, so it is worth laying out plainly who funds what.

| Stakeholder    | Success metric                   | AI they fund                              |
|----------------|----------------------------------|-------------------------------------------|
| Payer (FFS)    | Detect fraud and waste           | Fraud detection, utilization review       |
| Payer (VBC/MA) | Reduce total cost of care        | Risk stratification, care management      |
| Provider (FFS) | Maximize reimbursement           | RCM, CDI, coding optimization             |
| Provider (VBC) | Quality scores, fewer readmits   | Readmission prediction, population health |
| CMS            | Control spend, ensure access     | Quality measurement, program integrity    |
| Patient        | Timely, affordable, quality care | <b>(Usually not the buyer)</b>            |

The most consequential row is the last one: the patient wants care but does not write the check, so the entity that pays defines the objective function.

### Should You Build This?

**Build it when** the system attacks administrative waste both sides want gone: auto-approving clean prior-auth requests, catching coding errors before submission, shrinking the denial-and-appeal churn that consumes staff time without changing care. The claim pipeline is standardized, an 837 out and an 835 back, so it supports automation better than almost any other healthcare workflow. **Think twice when** the model's

value depends on winning the payer-provider arms race rather than ending it, or when a capitated buyer's savings come from denied care, the NaviHealth pattern, where algorithmic recovery timelines became de facto coverage decisions overridden less than 2 percent of the time. **Who bears the risk.** Errors land first on patients, who are rarely the buyer and never define the objective function; delay and denial reach them as pain, lost mobility, and adverse events long before any consequence reaches the payer.

### Quick Self-Test

1. What is the difference between an 837 and an 835, and why does "billed, not clinical" matter for your training data?
2. A readmission-reduction model saves money under one payment structure and loses money under another. Which is which, and why does that fragment adoption?
3. Write the Medicare Advantage payment identity and explain why every extra documented diagnosis carries fraud risk.
4. Why can a plan that never improves the care of its sickest members still climb the Star Ratings?
5. Where in the claim lifecycle does upcoding risk live, and what one guardrail most reduces it?
6. In the NaviHealth case, what made the algorithm's output a de facto coverage decision despite no rule requiring it?

## If You Remember One Thing

Follow the money. The claim lifecycle determines what data exists, the payment model determines what AI gets funded, and the payer's objective function becomes your model's objective function. A technically excellent model pointed at the wrong objective, like predicting cost when you meant to predict need, will faithfully optimize the wrong outcome. Understand the financial plumbing first, or you will build systems no one pays for, uses, or trusts.

---

*Richard J. Young, Ph.D. · richard@deepneuro.ai · Senior AI Research Scientist, UnitedHealth Group · UNLV Lee Business School · ORCID 0000-0002-1109-7552*