# The CMS autism toolkit is a build spec in disguise: where the 421 percent ABA spending curve creates real openings in utilization management, credentialing, documentation, and payment integrity

AUG 07, 2026



## 🎧 Part I Podcast free on Spotify.

Thoughts on Healthcare Markets & Technology

🎧 **Part I: The CMS autism toolkit is a build spec in disguise: where the 421 percent ABA spending curve creates real openings in UM, credentialing, documentation, and payment integrity**

ABA spending in Medicaid grew 421% from 2017 to 2022. Children served grew 66%. That gap is hours and billing. And it is a software problem…

Listen now

13 hours ago · Thoughts on Healthcare

## 🎧 Part II Podcast episode for paid subscribers only. Also available on Spotify.

Thoughts on Healthcare Markets & Technology

🎧 **Part II: The CMS autism toolkit is a build spec in disguise: where the 421 percent ABA spending curve creates real openings in UM, credentialing, documentation, and payment integrity**

ABA spending in Medicaid grew 421% from 2017 to 2022. Children served grew 66%. That gap is hours and billing. And it is a software problem...

Listen now

13 hours ago · Thoughts on Healthcare

*To listen to paid episodes in Apple or Spotify, li*
*settings on those platforms (instructions inside th*
*Podcast).*



# Table of Contents



# Abstract


CMS released the State Medicaid and CHIP Applied Behavior Analysis Toolkit in early August 2026. Headline data point: ABA spend in Medicaid and CHIP grew 421% from 2017 to 2022 against 66% growth in children served. The toolkit is explicitly not a new federal requirement and explicitly does not touch EPSDT. What it does is hand more than 50 state agencies and territories a shared vocabulary for medical necessity, treatment intensity, provider qualification, and fraud detection in a benefit that went from rounding error to line item in about six years.

Key takeaways for builders and investors:

- The spend/utilization gap (421 vs 66) is mostly hours per child and rate creep, not more children. That is a UM problem, and UM problems are software problems.

- The rendering provider gap is the structural flaw: technicians deliver the care, BCBAs own the NPI on the claim. Every integrity product eventually points here.
- EVV for in-home ABA is the obvious next mandate. Nobody has built the good version yet.
- State agencies can buy program integrity tooling at 90/10 and 75/25 federal match. That changes the ROI math on a slow sales cycle.
- CMS-0057 prior auth APIs land Jan 2027. ABA is the highest-friction PA category nobody built for.
- Evidence base is thin enough that outcomes measurement is the only defensible position once rates compress.



## What CMS actually shipped and why the 421 percent number matters

Read the press release twice and the second read is more interesting than the first. CMS says multiple times, in different ways, that the toolkit is not a restriction, does not create federal requirements, does not reduce EPSDT obligations, and does not endorse a particular flavor of ABA. That is the sound of an agency lawyer standing behind the policy shop with a hand on their shoulder. EPSDT is a statutory entitlement, states have lost in court repeatedly when they tried to hard cap medically necessary services for children, and CMS knows it. So the toolkit is guidance, checklists, and suggested practices, which in Medicaid-speak means states can adopt any of it, all of it, or none of it, and the ones with budget problems will adopt most of it by Q2 next year.

The 421 percent against 66 percent is the whole ballgame. Spend grew roughly six times faster than the population being served. There are only three ways that happens: rates went up, hours per child went up, or the billing got creative. Rates did go up in a handful of states but nowhere near six-fold, and several states cut. So it is hours and it is billing. A child authorized at 12 hours a week and a child authorized at 38 hours a week generate wildly different annual costs, roughly $30k versus close to $100k depending on the state rate, and the clinical literature supporting the difference

between those two is thinner than anyone in the industry likes to say out loud in a room with an actuary.

That gap is where the companies get built. Not in the politics of it, which are exhausting and will be litigated for years, but in the plumbing that lets a state or a plan tell the difference between the two children without generating 400 appeals a month.

## The ABA market as it exists today, warts included

Autism prevalence in the CDC's surveillance network is now about 1 in 36 eight-year-olds in the most recent published cycle, up from 1 in 44 two cycles prior and roughly 1 in 150 twenty years ago. Roughly half of those children touch Medicaid at some point. Every state has to cover medically necessary ABA for children under EPSDT, a posture CMS made explicit in a 2014 informational bulletin, and every state also has a commercial autism mandate on the books. Demand is not the constraint. Supply is.

The workforce is a two-tier structure that nobody would design on purpose. A little over 65,000 Board Certified Behavior Analysts nationally, and something in the neighborhood of 200,000 Registered Behavior Technicians. The BCBA writes the treatment plan, does the assessment, modifies protocols, and supervises. The RBT does the actual hours, sitting on the floor with a child running discrete trials, usually for $20 to $24 an hour with turnover that clinics quietly admit runs 65 to 75 percent a year. A staffing model with three-quarters annual churn in the delivery role is not a clinical model, it is a call center with better intentions.

The billing runs through a small set of CPT codes that most people outside the space cannot keep straight. 97151 for assessment, 97153 for the technician-delivered direct treatment in 15 minute units, 97155 when the behavior analyst modifies protocol, 97156 for caregiver guidance, plus the group codes. 97153 is the volume driver and it is where the entire economic argument lives. State Medicaid rates on it typically run $12 to $16 a unit, call it $48 to $64 an hour, with real spread from the low $30s to north of $80 depending on the state and whether there is a managed care markup.

The default treatment intensity is where things get uncomfortable. The 40-hour-a-week comprehensive model traces back to a 1987 study with 19 children in the treatment arm. That study did something important for the field, and it also became a load-bearing wall for a multibillion dollar benefit in a way its author probably did not anticipate. The Council of Autism Service Providers now distinguishes focused treatment at roughly 10 to 25 hours from comprehensive at 26 to 40, which is a

reasonable framework, and which a lot of providers treat as a menu where every child orders the tasting menu.

## Where the money leaks and who gets paid to plug it

CMS mentions investigations, prosecutions, and convictions related to kickbacks and harm to children. That is not abstract. The pattern in the enforcement actions is consistent enough to be a product spec. Recruiters paid per head to bring children into clinics. Diagnoses generated by clinicians with financial relationships to the treatment provider. Hours billed for sessions that did not occur or occurred with a family member watching TV nearby. Telehealth sessions billed as in-home direct treatment at the higher rate. Technicians whose credentials lapsed months ago still generating claims. And the classic: a single BCBA NPI on claims totaling more than 24 hours in a calendar day, which is either fraud or a genuinely remarkable person.

That last one is the structural flaw and it deserves its own paragraph. In most states the RBT cannot enroll as a Medicaid provider, so the claim carries the supervising BCBA's NPI as the rendering provider even though the BCBA was not in the room. The claim line literally does not know who delivered the care. Every downstream integrity question, every supervision ratio check, every no-show detection scheme, every attribution model for outcomes, runs into this wall. Some states have started requiring technician-level identifiers, either as an ordering/referring relationship or through a modifier convention, and it is a mess of state-by-state variation.

Building the resolution layer for that is a real company. Not a fraud dashboard, a registry and attribution service that maps technician identity, active credential status, supervision relationship, and session-level presence to the claim line, and does it across the credential registry, the state enrollment file, the practice management system, and the payer's claim history. Boring. Necessary. Very hard to rip out once it is in.

## Utilization management that clinicians will not hate

The toolkit's language about treatment intensity reflecting each child's clinical needs rather than standardized service models is a polite way of saying stop authorizing 40 hours by default. The states that have tried the blunt version have gotten mixed results. Indiana is the cautionary tale everyone cites. It watched ABA spend go from a rounding error in the late 2010s to a few hundred million a year, then in 2024 cut the

97153 rate to roughly $68 an hour and layered on a 30-hour weekly cap with a multi-year limit on comprehensive treatment. Providers left, families sued, and the state ended up in a fight it partly created by not having a defensible clinical process in the first place.

The lesson for builders is that hard caps are legally fragile under EPSDT and politically radioactive. Individualized medical necessity determination is neither. It is just expensive to do well, which is the definition of a software opportunity.

What the product looks like: an authorization engine that ingests the assessment instruments, the treatment plan, the goal set, and the prior authorization period's actual progress data, and produces a defensible hour recommendation with the clinical rationale attached. Tiering logic that distinguishes focused from comprehensive on something other than the provider's request. Automatic step-down triggers when goals are mastered. Escalation paths that route the 8 percent of genuinely complex cases to a human reviewer instead of the 92 percent that are template renewals.

The timing hook is CMS-0057. Impacted payers, including Medicaid fee-for-service and managed care plans, have to stand up FHIR-based prior authorization APIs by Jan 1, 2027, with decision timelines of 72 hours expedited and seven calendar days standard already phasing in. Everybody built for radiology and ortho. Almost nobody built for behavioral health, and ABA is arguably the highest-friction PA category in Medicaid: long authorization periods, huge dollar values, clinical documentation that lives in a specialty practice management system that does not speak FHIR, and a renewal cadence that hits every six months forever. Whoever ships the ABA-native DTR and PAS implementation gets to be the default.

## Credentialing and supervision is a data problem wearing a compliance costume

CMS calls out helping states ensure ABA providers are qualified and properly supervised. Two separate problems hiding in one sentence.

Qualification is an enrollment and screening question. States already have authority under the provider screening regulations to assign risk categories, require fingerprint-based background checks, and unannounced site visits. Several states have moved or are moving ABA providers into the high risk category, which triggers the full screening battery. That creates work: revalidation cycles, ownership disclosure, site visit logistics, license and certification monitoring against the certifying board's registry, and the tedious reality that a mid-size ABA group with 40 clinics across six

states is managing thousands of individual credential records with a shared inbox and a spreadsheet named final_v4.

Supervision is the more interesting problem because it is continuous rather than periodic. The certifying board requires that a percentage of technician hours be supervised, and states layer their own requirements on top. Nobody verifies this at the claim level in anything close to real time. A supervision monitoring product that reconciles technician direct-service hours against documented supervision contacts, flags ratio breaches before they become recoupments, and produces the audit packet on demand would sell to providers as risk insurance and to plans as an integrity control. Same underlying data model, two buyers, very different pitches.

The adjacent play is audit defense as a service. When states get aggressive with post-payment review, and they will, the recoupment demands land on providers who kept documentation in a system that was designed for scheduling, not for surviving a statistical extrapolation. There is a services business and eventually a software business in extrapolation challenges, sampling methodology disputes, and appeal workflow. Unglamorous, high retention, and it grows exactly when the market gets scary.

## Documentation, session capture, and the AI note trap

Here is a prediction that seems safe: electronic visit verification comes for in-home ABA within three years. The federal EVV mandate under the 21st Century Cures Act already covers personal care and home health services, states built the infrastructure, and the same fraud pattern that motivated it (services delivered in a home with no independent witness, billed in 15 minute increments) applies to a benefit that just grew 421 percent. The toolkit does not say EVV. The toolkit does not have to say EVV. Some state Medicaid director reading the fraud section is going to connect those dots by lunch.

The version that works captures the session data the clinician was going to collect anyway, trial-by-trial, and treats the compliance artifact as a byproduct rather than the point. Cameraless, no mandatory selfie every 15 minutes, GPS optional, timestamp attestation via passive network signal. The bar to clear is that the clinician's workflow is no worse than the paper data sheet they were using before. Everyone building "clinician-facing AI" for behavioral health should tape that on their monitor.

The AI note trap is adjacent and worse. There is a real product in ambient documentation for the BCBA sessions where narrative notes actually matter (parent

training, assessments, protocol modifications). There is a fraud accelerant in generating narrative notes for the tech-delivered 97153 hour, because the notes are what payers audit against and a plausible generated note is not the same thing as a session that happened. Any tool that autocompletes 97153 notes without linkage to session-level trial data is going to get somebody indicted eventually. Product decision: build the trial-level capture first, then let the LLM summarize what actually happened. Reversing that order is a lawsuit.

## Payment integrity, SIU tooling, and the very slow, very rich state buyer

State Medicaid Program Integrity units and their MCO SIU counterparts are the buyers who most obviously need this, and are the most painful to sell to. Sales cycles run 12 to 24 months, procurements go through the state's MMIS vendor of record, and the incumbent gets first look at every RFP because they already have the claims feed. The offset is that the buyer has real money and the federal match is the highest in health tech: 90/10 for design, development, and installation of MMIS-related systems, 75/25 for ongoing operations. A $3M enterprise deal effectively costs the state $300K in year one. That changes what "expensive" means.

The winning wedge is not "we detect fraud." Every incumbent claims that. The winning wedge is a specific measurable recoupment or avoided-payment number tied to a specific pattern that the incumbent misses, delivered as a pilot with a shared-savings option and a plain-English report a Medicaid director can hand to a legislator. Start with the technician-attribution and supervision-ratio patterns, because they are quantifiable, defensible in a hearing, and specific to ABA in a way generic FWA tools cannot match without a rebuild. Then land the enterprise MMIS integration on the second contract.

## Diagnosis, wait lists, and the front door of the funnel

Every conversation about ABA UM eventually collides with the diagnosis bottleneck. Waits for a developmental pediatrician or a full ADOS-2 evaluation run six to eighteen months in most metros and worse in rural counties. That wait is where the darker corners of the market operate: fast-turnaround "eval mills" that produce a diagnosis in a single 45-minute visit, sometimes co-owned with the ABA provider that receives the referral. The kickback risk here is not theoretical, it is in several of the DOJ press releases the toolkit implicitly gestures at.

Two product categories fall out. One, a legitimate telehealth diagnostic pathway that shortens the wait without shortening the assessment: async instrument administration, structured caregiver interview, video-observation review by a licensed clinician, with disclosure of any downstream referral relationship and a firm rule against ownership overlap with the ABA provider. Two, a screening triage layer that puts children into the right lane earlier — many kids waiting on an ABA eval need speech, OT, or a broader neurodevelopmental workup instead of, or in addition to, ABA. Payers would fund the second one directly if it can show reduced spend by routing the 20 percent of misdirected referrals somewhere else. It cannot be branded as gatekeeping. It has to be branded as care navigation.

## What this does to platform economics and the PE roll-ups

The private equity thesis in ABA since roughly 2018 was straightforward: fragmented single-clinic market, buy at 6 to 8x EBITDA, roll up to 15 to 25 clinics at 10 to 12x on exit, extract synergies from centralized RCM and credentialing. Rate stability was the underwriting assumption. That assumption is now visibly wrong. Indiana, Colorado, and a handful of others have cut rates or capped hours; every state Medicaid director on a budget deadline has the toolkit on their desk this quarter.

What happens to the roll-ups depends on their utilization mix. Groups whose average authorized hours were closer to focused-treatment norms (15 to 25 hours) have room to absorb a 10 to 15 percent rate cut without becoming unprofitable. Groups that built their unit economics on 30 to 40 hour authorizations across the entire caseload are the ones that go quiet in 2026 and 2027. Expect a wave of platform-to-platform sales at revised multiples, some tuck-in acquisitions that were quietly written down, and a handful of high-profile bankruptcies among the more aggressive comprehensive-only providers.

For a founder, the interesting reading is not that the roll-ups are struggling. It is that the exit market for a genuinely differentiated services business — outcome-measured, EVV-native, low-turnover workforce, clean payer relationships — is about to widen materially, because the strategic buyers need something to point at when their next fund closes and “we have 200 clinics” no longer clears the bar.

## Outcomes measurement as the only moat that survives a rate cut

When rates compress, the only providers who keep negotiating leverage are the ones who can prove they get children to specific, payer-relevant milestones faster or cheaper than the market average. That claim requires a measurement infrastructure the field mostly does not have. Standardized instruments exist — VB-MAPP, ABLLS-R, Vineland-3, PDDBI — and are used inconsistently, at inconsistent intervals, and almost never rolled up in a way a payer analytics team can benchmark across a network.

The product is not another assessment tool. There are already too many. The product is the layer that ingests whichever instrument the clinician uses, normalizes to a common outcomes ontology, benchmarks the individual child's trajectory against a matched cohort, and produces a report that survives contact with a health plan's medical director. Bonus points for tying the outcomes to specific goal domains payers actually care about (communication requests, toileting independence, dangerous behaviors reduced) rather than the field's more academic constructs.

This is also the underwriting layer for any value-based ABA contract, and there will be value-based ABA contracts within the next 24 months because the fee-for-service model is politically untenable at current growth rates. First mover has a real advantage: whoever defines the ontology tends to keep defining it.

## Underwriting these bets without getting flattened by the next policy swing

A few filters for anyone evaluating a company in this space over the next twelve months:

Reimbursement diversification. Anything with more than 70 percent revenue from a single state Medicaid program is one legislative session away from a bad quarter. Multi-state, mixed-payer, some commercial exposure is the durable posture.

Utilization posture. Ask for authorized-hours distribution across the caseload. A bell curve centered in the 15 to 25 hour range with a long right tail for genuinely complex cases is defensible. A bimodal or right-skewed distribution centered at 30 to 40 is a policy risk masquerading as a growth story.

Workforce structure. Annual RBT turnover under 45 percent is a sign the operator is running a real workforce program. Above 65 percent means the unit economics require constant churn and every incremental rate cut compounds directly onto retention costs.

Rendering-provider data hygiene. If the company cannot tell you, per claim, which specific technician delivered which session and which BCBA supervised it during that authorization period, they will not survive the integrity cycle coming in 2026-2028.

Outcomes measurement maturity. Not “we use VB-MAPP.” Ask when it was last administered on the median child in the caseload, what the cadence policy is, and whether they can produce a payer-facing outcomes report on demand. Most cannot.

PA-readiness for CMS-0057. Do they have FHIR endpoints. Are they participating in a payer’s PAS/DTR pilot. If the answer is a blank stare, they are going to spend 2026 in a very expensive integration project instead of growing.

The unlock in this toolkit is not the toolkit itself. It is that CMS has now put a coordination point on the table for 50-plus fragmented state programs, and every builder who has been waiting for a shared vocabulary to sell against just got one. The companies that ship the boring plumbing — attribution, supervision monitoring, PA automation, outcomes normalization — become the picks-and-shovels layer under whatever the states decide to enforce. The companies still selling a comprehensive-hour-based clinical model without any of that infrastructure are going to spend 2026 explaining themselves

.



14 Likes • 1 Restack


← Previous

Next →

## Discussion about this post

Comments Restacks

Write a comment...