Most of your coding finishes itself.

CodeSight reads the visit note and assigns the billing codes — accurately, in seconds, at a scale no coding team can match.

10.9s

to code one chart, start to finish

92.9%

of condition codes finished with **no human review**

98.1%

matched the **visit level the practice billed**

Measured on 945 real office visits from one de-identified outpatient specialty practice, graded against the codes insurers actually paid, with 376 held back as a fair test. The 92.9% counts the codes that say what's wrong with the patient; counting every code, including administrative ones like "history of," it is 87.9%. The automated share stayed above 95% accurate either way.

WHAT THIS MEANS FOR YOUR PRACTICE

Most charts never reach a coder.

The ones that do arrive already coded, flagged with the reason — a review queue, not a workload.

Volume becomes a setting, not a hiring plan.

About 300 charts an hour per lane, and lanes can be added. A coder works through 20–30.

Nothing gets billed on a guess.

When CodeSight isn't confident, the chart goes to a certified coder instead.

Full report — methods, results, errors, and limits: **medmio.com/case-studies/codesight-accuracy-benchmark**

WHAT WE TESTED

CHARTS
945 real office visits

PRACTICE
One de-identified outpatient specialty practice

THE ANSWER KEY
The codes actually billed, and actually paid

CUSTOM SETUP
None — it had never seen this practice

THE FAIR TEST
376 charts set aside before we started, never used to build the system

THE FOUR QUESTIONS TO ASK ANY CODING VENDOR

QUESTION	MEDMIO	TYPICAL VENDOR
How many charts did you test on?	**945 real charts, billed and paid**	rarely says
Did you set charts aside for a fair test?	**Yes — 376, locked before scoring**	rarely says
How much does it finish with no human review?	**92.9% of condition codes, with no human review**	accuracy only
Which way do the mistakes lean?	**4 too high, 3 too low, out of 376**	rarely disclosed

12
DISAGREEMENTS

And when it disagreed, it was often right.
That is every disagreement across all 945 charts — and in 6 of them the documentation supported CodeSight, not the bill. Revenue the practice had already earned and never claimed.

Scope of this test. These results cover routine office visits, where the data is deep enough to state a number with confidence; the highest-complexity visits are rarer in this practice's mix and are reported separately. Every disagreement with the billed record is scored against CodeSight — a deliberately hard standard — and goes to an independent panel of certified coders.

Get the full benchmark report.

Six pages of method, results, and confidence intervals — every figure traceable to the test that produced it. Free, no email required.

Read the report →