How Courts May Interpret Data-Driven Healthcare Fraud Suits

By Jaime Jones, Andrée-Anne Fournier and Atang Gilika

Law360 (October 9, 2025, 5:35 PM EDT)


Jaime Jones


Andrée-Anne Fournier


Atang Gilika

On Aug. 19, the Centers for Medicare & Medicaid Services announced the launch of the Crushing Fraud Chili Cook-Off Competition, an initiative that CMS described

> as a market-based research challenge aimed at harnessing explainable artificial intelligence (AI), specifically machine learning (ML) models, to detect anomalies and trends in Medicare claims data that can be translated into novel indicators of fraud.[1]

This initiative follows two other recent steps by the U.S. Department of Justice to strengthen data-driven enforcement in healthcare: the June 30 creation of the Health Care Fraud Data Fusion Center — a multi-agency initiative involving the DOJ, the U.S. Department of Health and Human Services' Office of Inspector General, the Federal Bureau of Investigation and others[2] — and the July 2 renewal of the DOJ-HHS False Claims Act Working Group, with an updated focus on cross-agency collaboration and data mining.[3]

As data analytics tools become increasingly used in enforcement actions, courts have to grapple with whether allegations of fraud based only on data and statistical extrapolations can state viable claims under the FCA and how to assess evidence that arises from data mining techniques.

While there is some dispute among courts, most reject attempts to assert FCA violations that rely merely on purported statistical analyses of data to suggest that fraud likely occurred, particularly where such allegations are advanced by individuals or entities with no inside knowledge of the defendant's operations that allegedly caused the submission of false claims.

And the DOJ has exercised its statutory authority to dismiss actions brought by a professional relator formed for the express purpose of leveraging CMS' publicly available Medicare claims data to seek whistleblower bounties.

But when cases survive challenges to the pleadings and liability for alleged fraud and falsity of claims can be established, courts have allowed damages to be determined based on statistical extrapolation.

Because of the draconian damages and penalties imposed under the FCA, the incentive of the whistleblower bounty, the availability of increasingly robust data sources, and the government's own increasing focus on using data to detect fraud, whistleblowers and government lawyers will likely bring more cases built on data analyses and signals of fraud.

The mere filing of those cases could impose significant costs on defendants, and as data analytics tools are sharpened and combined with facts collected from employees, customers or competitors, some percentage of those cases may survive dismissal efforts, expanding the case law in this area.

The government's recent initiatives underscore the growing role of a data-driven approach to detecting fraud and abuse in healthcare. However, without a foundation in sound theory, clinical expertise and institutional context, such an approach could generate misleading or inaccurate conclusions.

A Potential Data-Driven Approach to Fraud and Abuse Detection

Broadly speaking, data mining involves identifying patterns in large datasets, such as healthcare claims data, using statistical and ML tools. These techniques typically fall into two categories.

Supervised Approaches

Two common examples of supervised approaches include hypothesis-driven testing and classification models. Hypothesis-driven testing starts with a specific concern, e.g., potential overbilling by a provider, and formulates a testable hypothesis.

Statistical methods are then used to test whether the observed data deviate meaningfully from what would be expected under normal conditions. In contrast, classification models use ML to distinguish between categories such as fraudulent and nonfraudulent claims. Trained on historical data, these models can then classify new claims based on their similarity to past patterns of concern.

Unsupervised Approaches

Unsupervised methods search for patterns without preidentified hypotheses or trained models. Such methods include techniques such as clustering (grouping data based on similarity) or anomaly detection (identifying outliers), which identify claims that deviate from established norms without specifying what constitutes potential fraud.

These tools can uncover previously unknown schemes, but without careful interpretation, they may conflate unusual circumstances with unlawful activities.

The Risk of Atheoretical Data Mining

Whether a supervised or an unsupervised approach is used, data mining could lead to improper conclusions if it lacks theoretical or clinical grounding. Suspicious patterns must be validated with outside information.

In supervised approaches, running numerous tests without clear hypotheses increases the risk of false positives — random patterns that appear meaningful but do not indicate fraud. Classification models may also be misleading if trained on unrepresentative data, making them unreliable when applied to real-world data.

In unsupervised approaches, results can be difficult to interpret. Without clear outcomes, it can be unclear why something was flagged as unusual. Models may rely on variables that merely correlate with unusual behavior rather than cause it, which could lead to improper conclusions.

Here are a few examples of how data-driven approaches can lead to incorrect conclusions.

Outlier does not mean fraudulent.

Unusual billing patterns alone are not evidence of fraud. If patients were randomly assigned to physicians, then a provider whose billing patterns looked very different from their peers' might raise concerns. In reality, though, patients often choose physicians based on their specific needs, leading to natural variation.

For example, a highly regarded specialist may see more complex cases, order more tests or prescribe a broader range of treatments. While their billing may differ significantly from that of local peers, it could reflect their specialized patient population rather than fraud.

Substitution patterns can be misinterpreted.

In healthcare, different treatments can serve as substitutes. For example, patients with coronary artery disease may receive either a cardiac stent (a less invasive procedure by a cardiologist) or bypass surgery (a more invasive procedure by a cardiothoracic surgeon). If access to one option decreases, patients may shift to the other.

A sudden rise in cardiac stent billing at a rural hospital might raise concerns about unnecessary procedures, especially if it cannot be explained by changes in patient

volume or staffing. However, if the area's only cardiothoracic surgeon has left, the increase may simply reflect a shift in treatment patterns due to reduced access to bypass surgery, not improper care.

Billing clusters can be legitimate.

Billing clusters, or groups of services that are frequently submitted together on insurance claims, often reflect standard medical practices, such as combinations of procedures, tests or treatments for a specific diagnosis or visit.

While unusual billing clusters might raise concerns, they do not necessarily indicate fraud. For example, if certain lab tests that were previously billed separately begin appearing together, it could be flagged as suspicious. However, if updated clinical guidelines recommend ordering these tests together, the new pattern could reflect sound medical practice instead of improper billing.

Why Data Analytics Grounded in Theory Is Critical

As the examples above illustrate, grounding data analytics in sound theory is critical to avoid reaching incorrect conclusions.

A principled approach to detecting potential fraud and abuse through data analytics should begin with careful consideration of what patterns one should expect before examining the data.
Examples of questions that can help orient this understanding include:

- What behavior should be expected given patient needs, payment incentives and clinical norms?
- What policy changes or institutional constraints might explain a pattern?
- What would a baseline of typical or legitimate activity look like for this type of provider, service or patient population?

A careful approach should also include:

- Thorough preprocessing of the data, e.g., ensuring that differences in reporting metrics across healthcare providers or other anomalies are addressed up front;
- Rigorous sanity checks and sensitivity analyses, particularly when using unsupervised methods; and
- Validation of findings generated through data analytics using external or independent sources of evidence.

How Courts Approach Healthcare Fraud and Abuse Detected Through Data Mining

Courts have had a mixed reception to complaints relying on information gathered through data mining to allege FCA liability.

Some courts have rejected these data mining complaints at the motion-to-dismiss stage when there was an obvious nonfraudulent explanation for the data.

For example, the U.S. Courts of Appeals for the Fifth Circuit and Ninth Circuit granted motions to dismiss in two similar cases, U.S. ex rel. Integra Med Analytics LLC v. Baylor Scott & White Health in 2020, and Integra Med Analytics LLC v. Providence Health & Services in 2021.

Integra Med Analytics is a company that specializes in using statistical analysis to identify alleged fraud. In both cases, Integra analyzed CMS data and determined the respective defendants submitted proportionally more claims with higher-paying diagnosis codes than comparable institutions.

Both courts found that Integra failed to state a plausible claim for violation of the FCA because its allegations did not rule out a clear legal alternative explanation for the data results: that the defendants were simply ahead of the healthcare industry in effectively coding for Medicare reimbursement.

Although Integra also relied on statements from several former Baylor medical coders to support its allegations in the Baylor Scott case, those statements also were consistent with the alternative explanation.

Despite granting the motions to dismiss, neither court wholesale rejected the notion that statistical data could be used to meet the pleading standards of the Federal Rules of Civil Procedure 8 and 9 in an FCA case.

Indeed, not all cases have failed at the motion to dismiss stage. In United States v. Mariner Health Care Inc., decided by the U.S. District Court for the Northern District of California in 2021, the court denied Mariner's motion to dismiss in another case where Integra analyzed Medicare claims data to identify alleged fraud.

In grappling with whether a statistical analysis could be sufficient on its own to allege fraud with particularity to state an FCA claim, the court discussed that Rule 9(b) "is designed to ensure that factual allegations are 'specific enough to give defendants notice of the particular misconduct which is alleged to constitute the fraud charged so that they can defend against the charge."

Further, like the Fifth and Ninth Circuits, the court said there is not a categorical exclusion against using statistical analysis to allege an FCA claim. Here, the court said Integra "pled with specificity allegations which support the empirical reliability and probative value of its statistical study" based on its utilization of a large sample size of 13 million samples and the quality of its statistical analysis.

The court considered alternative, nonfraudulent explanations put forth by Mariner, but unlike the Fifth and Ninth Circuits, did not find them so convincing as to render Integra's explanation implausible.

The court based its decision to deny the motion to dismiss on the sufficiency of the statistical analysis alone. It noted, however, that "even if an FCA claim requires specific evidence of unlawful conduct in addition to statistical evidence," Integra also alleged testimonials from employees and family members of former patients to supplement its statistical analysis.

Data analytics are powerful tools that can be useful in identifying potential healthcare fraud and abuse, but their value depends on how thoughtfully they are applied.

Without theoretical grounding, clinical insight and institutional awareness, the use of these tools could lead to misclassification of legitimate behavior as fraudulent.

As the DOJ and other agencies increasingly turn to data mining as an enforcement tool, courts will have to determine how far data alone can take a fraud case. A principled, context-aware approach will be essential to avoid the trap of letting the data decide in isolation.

Jaime L.M. Jones is a partner and co-leader of the global healthcare practice at Sidley Austin LLP.

Andrée-Anne Fournier is a managing principal at Analysis Group Inc.

Atang Gilika is a vice president at Analysis Group Inc.

Endnotes

1 https://www.cms.gov/priorities/crushing-fraud-waste-abuse/overview/crushing-fraud-chili-cook-competition.

2 https://www.justice.gov/opa/pr/national-health-care-fraud-takedown-results-324-defendants-charged-connection-over-146.

3 https://www.justice.gov/opa/pr/doj-hhs-false-claims-act-working-group.