#ForensicPerspectives

Sampling as a strategic tool in litigation and investigations: Precision and power for complex matters

In high-stakes litigation, investigations, and periods of heightened scrutiny, reviewing every document, transaction, or claim is rarely feasible. Sampling can transform overwhelming data volumes into clear, defensible insights—supporting legal strategy while managing costs and compressing time.

We advise clients and counsel to design statistically defensible sampling methodologies that deliver rigor without overreach. Our approach helps clients:

- ✓ Validate findings with statistical rigor and credibility
- ✓ Facilitate critical data is not overlooked
- ✓ Refine review scope—avoiding over- or under-inclusiveness
- ✓ Achieve defensible outcomes in litigation and regulatory proceedings
- ✓ Cost-effectively bolster confidence in case strategy
- ✓ Confirm completeness and accuracy of findings
- ✓ Define clear, consistent units of analysis in support of reliable conclusions

Whether responding to a government inquiry, defending a class action, conducting an internal investigation or compliance review, or assessing the impact of a cyber breach, sampling offers a strategic advantage—enabling efficient, defensible decision-making.

How sampling accelerates analysis

Sampling does more than reduce data volumes, it improves the quality and defensibility of findings.

eDiscovery search term validation

Use sampling to test whether search terms are hitting the mark to capture what matters and excluding what doesn't. The result is statistically grounded decisions that reduce disputes and streamline discovery.

AML alert testing

To support both compliance and regulator confidence, sampling allows financial institutions to evaluate whether transaction monitoring systems are tuned to capture high-risk behaviors without generating excessive noise.

Review completeness

Statistical sampling helps to validate that the investigation covers the right ground, bolstering credibility when findings are scrutinized.

Cyber breach assessments

In exfiltration scenarios, a valid sample of affected files can determine whether PII, PHI, or trade secrets were truly compromised to inform legal exposure, notification obligations, and response strategy.

Defining units of analysis

Sampling enforces structure—by vendor, geography, transaction type—so your team can deliver targeted, actionable findings.

CRA Charles River Associates

Technical considerations

Below are key technical elements to address:

Define the sampling unit

Decide what is being sampled, this decision shapes the entire sampling frame and analysis.

Design a sampling frame

The sampling frame is the entire list of units from which the sample will be drawn; it should be comprehensive, current, and deduplicated, as gaps or errors in the frame can introduce bias from the start.

Plan for missing or incomplete records

Some units may have missing data or incomplete details, at which point its important to understand regulatory expectations around how these gaps should be handled, through extrapolation, substitution, or accepting a defined level of error.

Choose your sampling method

The method should match the goals, resources, and structure of the population.

Examples of successful sampling applications

Healthcare system under regulatory review: In a federal inquiry into alleged Medicare upcoding, counsel faced thousands of patient records—far too many for a full review. A statistically valid sample allowed the legal team to identify potential overpayments, coding anomalies, and documentation gaps, while confirming that no major patient populations or procedures were overlooked. The sampling results supported focused remediation, informed negotiation with regulators, and helped frame a credible response strategy.

eDiscovery search term validation: In a large-scale litigation involving millions of documents across multiple custodians, both sides agreed that the search terms were adequate but not excessive. By drawing a statistically valid sample of the documents that were not flagged by the search terms, we were able to confirm the precision and effectiveness of the terms.

Cybersecurity data breach assessment: When a client experienced a massive data exfiltration event from across a variety of data stores, counsel wanted to quickly determine whether regulated information had been resident in each data store. By using statistical sampling, our team was able to defensibly hive off exfiltrated data sets from certain data stores, thereby reducing the review population from 3.3 million to 1.3 million documents, reducing costs accordingly and substantially accelerating time to notification.

Money laundering alert system review: For a financial services client subject to a Consent Order, it was critical to demonstrate to the regulator that the newly deployed transaction monitoring algorithms were effectively alerting all transactions that required additional review, while not overburdening the client's AML compliance team by over alerting and generating excessive false positives. By sampling transactions that had not been alerted, we were able to defensibly demonstrate and confirm that the algorithms were working effectively as designed, enabling us to expeditiously conduct a Look Back Review that encompassed over 1 billion transactions and 28 million customers.

Kristofer Swanson, CPA/CFF, CFE, CAMS
Vice President and Practice Leader, Forensic Services
+1-312-619-3313 | kswanson@crai.com

Patricia Peláez, CPA/CFF, CFE, CPC, CAMS, CMS
Principal, Forensic Services
+1-312-577-4180 | ppelaez@crai.com

CRA's Forensic Services Practice – including our digital forensics, eDiscovery, and cyber incident response lab – is certified under ISO 27001:2022 standards. The Practice has been recognized by National Law Journal, Global Investigations Review, and ranked by Chambers. CRA's clients over the past two years included 98% of the AmLaw 100 law firms, and 85% of the Fortune 100 companies.