

Evaluating Workplace Drug Testing Effectiveness: A System Sensitivity Framework for Urine and Oral Fluid Testing

White Paper on Detection Effectiveness, System Sensitivity, and Structural Barriers to Adoption

Jim Tedrow, MT(ASCP)

The Compliance Resource Group, Inc.

March 2026

This paper examines differences in real-world drug detection outcomes between urine and oral fluid testing using a **system sensitivity framework**. Despite growing evidence that oral fluid testing can more effectively identify illicit drug use than traditional urine testing, adoption across workplace testing remains limited. By combining program-level datasets with industry-scale reporting, this analysis explores the structural, economic, and behavioral factors that have slowed adoption and considers mechanisms that may help align testing practices with detection effectiveness.

Summary

Workplace drug testing programs have traditionally evaluated specimen types through an analytical framework focused on laboratory cutoffs, assay performance, and detection windows. Under this lens, urine and oral fluid testing are often viewed as functionally comparable options within modern laboratory environments.

However, analytical performance represents only one component of a testing system. The effectiveness of a drug testing program ultimately depends on its ability to obtain an authentic specimen from the intended donor and detect drug use under real-world behavioral conditions. This broader concept—referred to in this paper as system sensitivity—captures the combined influence of specimen collection design, donor incentives, and laboratory detection.

This paper examines specimen effectiveness through a system sensitivity framework using three complementary sources of evidence:

- **Program-level comparative datasets** in which employers transitioned from urine to oral fluid testing within the same operational environments.
- **Contemporaneous paired testing data** in which DOT urine and non-DOT oral fluid specimens were collected from the same donors during the same testing events.
- **National industry reporting** from the Quest Diagnostics Drug Testing Index (DTI), encompassing millions of workplace drug tests.

Across the datasets examined in this paper, a consistent pattern emerges: oral fluid testing identifies substantially more illicit drug use than urine testing in comparable workplace populations. In program comparisons where employers transitioned from urine to oral fluid testing, oral fluid testing frequently identified **~2.6x – 3.3x more positive results** than urine testing within the same operational environments. Paired testing datasets where oral fluid and urine samples were collected on the same donors at the same time produced similar outcomes with oral fluid identifying **3.5x more positive results** and **6x more discordant**

positive results when compared with urine, indicating that the difference cannot be explained by population variation alone.

Analysis of national data from the Quest Drug Testing Index (DTI) further reinforces this pattern. When evaluated using relative system sensitivity—a comparison of oral fluid and urine positivity rates within the same drug category—oral fluid testing demonstrates markedly higher detection rates across all illicit drug classes analyzed. In categories such as marijuana, cocaine, PCP, and methamphetamine, oral fluid testing identifies **~3x – 5x more positive results** than urine testing.

Notably, this divergence is not uniform across all drug classes. Drug categories associated primarily with illicit use show the largest detection gaps, while drug classes frequently associated with legitimate prescription medications show substantially smaller differences—with higher positivity in urine testing. This pattern closely mirrors behavioral incentives within workplace testing programs. Individuals using illicit substances face stronger incentives to evade detection, increasing the likelihood of circumvention attempts where collection systems permit them.

Viewed through this behavioral lens, the observed differences between specimen types are not primarily a function of laboratory analytical capability. Instead, they reflect how specimen collection design interacts with donor incentives within real-world testing environments. Testing systems that reduce opportunities for circumvention more reliably obtain authentic specimens from the intended donor, resulting in a more accurate assessment of drug use.

Oral fluid testing volumes reported in the Quest DTI from 2016 to 2023 fluctuate modestly year by year but exhibit no discernible growth trend. In fact, reported volumes in 2023 are approximately the same as those observed in 2016. Yet the persistence of observed drug detection patterns across transitional program comparisons, contemporaneous paired collections, and large-scale national datasets raises an important question: *if oral fluid testing consistently identifies illicit drug use more effectively, why has its adoption remained limited across workplace testing programs?*

The answer appears to lie not in scientific uncertainty, but in a combination of **structural, operational, and economic factors** embedded within the current drug testing ecosystem. Collection supply logistics, cost structures, and long-standing program conventions have historically been optimized around urine-based testing. At the same time, the absence of a widely adopted framework for evaluating system sensitivity has created a form of industry blind spot: many stakeholders see the reported outcomes of urine testing but lack visibility into drug use that may go undetected within the system. Without a clear view of this detection gap, employers and program administrators may overestimate the effectiveness of urine testing and underestimate the practical value of alternative specimen types. As a result, the underlying incentives to confront the structural, operational, and economic barriers to oral fluid adoption remain weak.

This paper therefore examines not only the behavioral dynamics that produce differences in system sensitivity, but also the practical barriers that have slowed the broader adoption of oral fluid testing across the industry. These barriers include collection device inventory risk, cost allocation challenges within testing programs, the absence of scalable operational frameworks designed specifically for oral fluid testing, and the overall higher price point of oral fluid testing relative to urine testing.

Addressing these constraints does not require changes to laboratory science. Instead, it requires rethinking how oral fluid testing programs are structured, supplied, and supported operationally. The final sections of

this paper explore these structural challenges and outline several mechanisms that could help remove existing barriers, enabling testing programs to capture the detection advantages demonstrated throughout this analysis.

Table of Contents

Summary

1. Introduction: A Familiar Debate at an Inflection Point

2. The Industry's Analytical Assumption

3. Analytical Sensitivity vs. System Sensitivity

- 3.1 Analytical Sensitivity
- 3.2 System Sensitivity
- 3.3 Why Specimen Type Matters at the System Level
- 3.4 Reframing Performance Comparisons
- 3.5 Implications for Data Interpretation

4. Comparative Program Analysis: When Testing Conditions Are Similar

- 4.1 Importance of Program-Level Comparisons
- 4.2 Observational Program Transitions
- 4.3 Contemporaneous DOT Comparisons
- 4.4 System Sensitivity as the Primary Differentiator

5. Quest Drug Testing Index Through a Relative System Sensitivity Lens

- 5.1 Relative System Sensitivity
- 5.2 What the Quest DTI Measures Well
- 5.3 Limitations of Aggregate Datasets
- 5.4 Drug Class Patterns and Behavioral Incentives
- 5.5 Population Signals Reinforcing Comparison Findings

6. Why Higher Detection Persists: Donor Behavior and Collection Design

- 6.1 Drug Testing as a Behavioral System
- 6.2 Privacy, Control, and Opportunity in Urine Collection
- 6.3 How Oral Fluid Collection Alters Donor Incentives
- 6.4 Why the Effect Persists Across Drug Classes
- 6.5 Implications for Program Effectiveness

7. If the Data Is Clear, Why Hasn't Adoption Followed?

- 7.1 An Information Gap Across the Testing Ecosystem
- 7.2 Adoption Lags Are Not Evidence Gaps
- 7.3 Inventory and Capital Friction at the Collection Site Level
- 7.4 Misalignment Between Cost Timing and Revenue Events
- 7.5 Limited Visibility Across the Device Lifecycle

- 7.6 Risk Concentration in the Wrong Places
- 7.7 Why Structural Barriers Persist
- 7.8 The Nature of the Problem Defines the Nature of the Solution

8. What a Scalable Oral Fluid System Requires

- 8.1 Device Lifecycle Visibility
- 8.2 Alignment of Economic Timing
- 8.3 Risk Distribution Proportional to Control
- 8.4 Operational Simplicity
- 8.5 Preservation of Clinical and Compliance Integrity
- 8.6 A Coordinating Layer, Not a Replacement Layer

9. Implications for Stakeholders

- 9.1 Employers
- 9.2 Collection Sites
- 9.3 Laboratories
- 9.4 Manufacturers
- 9.5 Third-Party Administrators (TPAs)
- 9.6 Regulators, Insurers, and Risk Stakeholders

10. Conclusion: From Evidence to Action

Acknowledgements

Appendix A – Methodology

Appendix B – CRG Dataset Description

Appendix C – Quest DTI Data

Sources

1. Introduction: A Familiar Debate at an Inflection Point

Over the past two decades, oral fluid drug testing has transitioned from emerging alternative to established methodology. Regulatory frameworks have evolved, laboratory technologies have matured, and large-scale datasets now span multiple years of comparative performance.

Yet despite this maturation, adoption patterns across the industry remain uneven. In many testing programs, urine remains the default specimen type—even in contexts where comparative data suggests that oral fluid testing may identify illicit drug use at materially higher rates.

This divergence between demonstrated performance and operational adoption warrants examination.

Historically, specimen-type debates centered on analytical concerns and matrix characteristics. These questions were appropriate during early adoption phases. Today, however, the evidence base is broader and more longitudinal. The discussion must evolve accordingly.

The question is no longer whether oral fluid testing can be performed reliably. It can.

The question is whether the industry is evaluating specimen effectiveness using the right framework—and whether structural factors, rather than scientific uncertainty, are shaping adoption decisions.

This paper examines specimen performance through a systems-based lens, integrating observational comparative analysis, population-level data, and operational realities across the testing ecosystem. The objective is not to advocate for a particular stakeholder or disrupt established practices. It is to clarify how specimen type interacts with system design—and why that interaction matters.

2. The Industry's Analytical Assumption

Drug testing effectiveness is often evaluated as a laboratory question. In this framework, specimen types are compared primarily on analytical performance: sensitivity, specificity, assay validation, and regulatory compliance.

Under this analytical perspective, the performance evaluation effectively begins at the point the specimen arrives at the laboratory.

Implicitly, it assumes that:

- The specimen originated from the intended donor
- That valid specimen collection has not been circumvented or otherwise compromised (e.g. substituted or adulterated)
- Detection outcomes are determined primarily by laboratory analysis

When viewed strictly through this lens, differences between urine and oral fluid testing may appear incremental or context dependent. Modern laboratories are highly capable across both matrices, and analytical technologies have largely converged in sophistication.

However, this framework does not fully account for what occurs prior to laboratory receipt—where donor behavior, collection design, and opportunities for specimen manipulation can materially influence what is ultimately analyzed.

By focusing narrowly on analytical sensitivity, the industry has at times overlooked the broader concept of system sensitivity—the end-to-end effectiveness of a testing program in identifying drug use within real-world behavioral contexts.

As long as specimen types are evaluated solely through an analytical lens, persistent differences in detection outcomes will remain incompletely explained. A more comprehensive framework is required.

The following section introduces that framework.

3. Analytical Sensitivity vs. System Sensitivity

Discussions comparing urine and oral fluid drug testing are often framed as analytical debates—focused on laboratory cutoffs, assay performance, and detection windows. While these considerations are important, they address only one portion of the overall testing process and do not fully explain the performance differences observed between specimen types in real-world testing programs.

To meaningfully evaluate testing effectiveness, it is necessary to distinguish between analytical sensitivity and system sensitivity.

3.1 Analytical Sensitivity

Analytical sensitivity refers to the ability of an assay to detect a specific substance once a valid specimen has been received and processed. This includes factors such as:

- Assay design and performance
- Cutoff levels
- Instrumentation

Modern laboratories performing urine and oral fluid testing operate under rigorous regulatory and accreditation standards. When a valid specimen is properly collected, labeled, transported, and analyzed, there is little evidence to suggest meaningful analytical performance differences between urine and oral fluid testing in contemporary laboratory environments.

Importantly, analytical sensitivity assumes that:

- The specimen originated from the intended donor, and
- The specimen has not been substituted, adulterated, or otherwise compromised prior to laboratory receipt.

These assumptions are often implicit in discussions of test performance—but they are not guaranteed by all collection methods.

While urine and oral fluid testing are both supported by highly capable analytical methodologies, it is important to acknowledge that specimen types differ in their underlying biological characteristics, including detection windows and drug incorporation patterns. These differences are driven by matrix-specific pharmacokinetic factors rather than limitations in laboratory assay performance.

For the purposes of this paper, analytical sensitivity is defined as the ability of a laboratory method to detect a drug when it is present in a valid specimen. In contrast, factors that influence whether drug is present in the specimen at the time of collection—including detection windows, pharmacokinetics, and donor behavior—are considered components of system sensitivity.

3.2 System Sensitivity

System sensitivity describes the end-to-end ability of a drug testing system to identify drug use—from donor presentation through final laboratory reporting. It encompasses not only laboratory analysis, but also the collection environment, donor behavior, and opportunities for specimen manipulation prior to analysis.

System sensitivity is influenced by factors including:

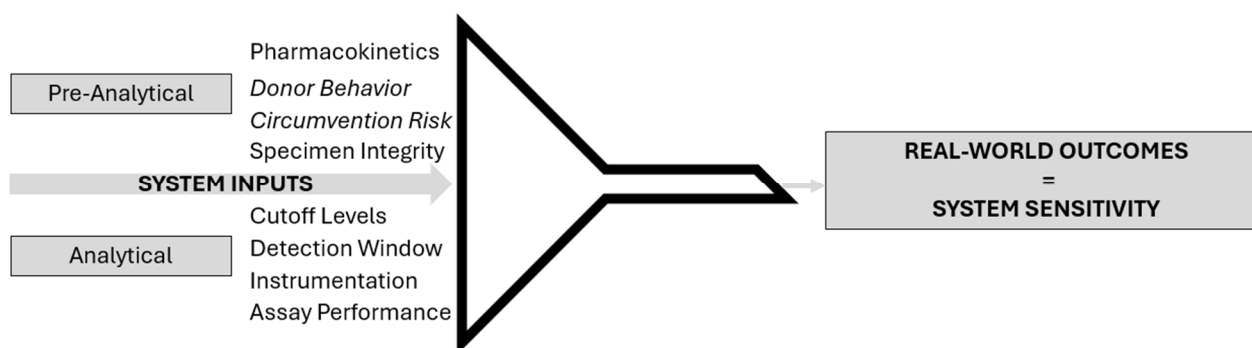
- Pharmacokinetics
- Detection window
- Whether the collection is directly observed
- The degree of donor control over the specimen
- Opportunities for circumvention (e.g. substitution or adulteration)
- Timing and predictability of the testing process

In this context, the specimen type itself plays a central role—not because of laboratory performance, but because of how it shapes the collection system.

The interaction of these factors is best understood as a combined system in which pre-analytical and analytical inputs collectively shape real-world detection outcomes, as illustrated in Figure 1.

Figure 1 – System Sensitivity Framework

Conceptual model illustrating how pre-analytical and analytical factors combine to influence real-world detection outcomes



When populations are comparable, system sensitivity reflects the combined impact of upstream factors influencing real-world detection outcomes.

3.3 Why Specimen Type Matters at the System Level

Urine collection, by design, introduces elements that may reduce system sensitivity. Even when conducted according to established procedures, urine collection typically involves:

- Privacy accommodations that limit direct observation
- Physical separation between the collector and the donor during specimen production
- Opportunities for pre-collection planning or substitution

Over time, these characteristics have influenced donor expectations and behaviors. In many testing populations, donors understand that valid prescription explanations, circumvention techniques, or timing strategies may reduce the likelihood of a reportable positive result.

It is important to note that directly observed urine collections are typically limited to specific regulated circumstances (e.g., DOT return-to-duty or follow-up testing) and represent a relatively small proportion of overall testing. In most routine workplace testing programs urine collections are conducted without direct observation, preserving donor privacy but also introducing potential opportunities for specimen manipulation.

Oral fluid collection alters this dynamic. Observed collections, direct donor identification, and continuous line-of-sight during specimen acquisition materially reduce opportunities for specimen manipulation. As a result, the testing system more consistently obtains a sample from the individual intended to be tested, at the moment the test is administered.

These differences are consistent with behavioral and collection-related factors that may influence specimen integrity, including the potential for circumvention in less controlled collection environments.

3.4 Reframing Performance Comparisons

When oral fluid testing demonstrates higher positive rates than urine testing in comparable populations, the difference should not be interpreted as a laboratory advantage. Rather, it reflects a system-level improvement in the ability to:

- Obtain an authentic specimen from the intended donor
- Minimize circumvention as a risk factor
- Align donor behavior with testing intent

For this reason, comparisons limited to analytical sensitivity alone are insufficient to explain observed outcomes. A system sensitivity framework provides a more accurate and actionable lens for understanding why oral fluid testing often identifies illicit drug use at higher rates than urine in real-world settings.

3.5 Implications for Data Interpretation

This distinction has important implications for how industry data is evaluated. Population-level datasets comparing urine and oral fluid positivity rates are not merely reflecting laboratory performance differences; they are capturing the cumulative effects of specimen integrity, donor behavior, and collection design.

Accordingly, analyses that incorporate system sensitivity offer greater explanatory power when assessing:

- Differences in positivity rates by specimen type
- Drug class-specific detection patterns
- Real-world program effectiveness

The sections that follow apply this framework first to comparative datasets and then to large-scale industry reporting.

4. Comparative Program Analysis: When Testing Conditions Are Similar

Population-level datasets can provide important directional insight, but they cannot isolate causality. Differences observed across large testing populations may reflect variations in employer practices, donor demographics, or testing protocols rather than the inherent characteristics of the specimen type itself.

To evaluate specimen effectiveness more directly, it is useful to examine datasets in which key operational variables are more tightly aligned. When employer populations, testing conditions, and laboratory standards are broadly comparable, differences in detection outcomes can be more reasonably attributed to the collection system itself.

4.1 Importance of Program-Level Comparisons

Meaningful evaluation of testing effectiveness requires more than comparing aggregate positivity rates across unrelated populations. To isolate system sensitivity effects, it is beneficial for analyses to minimize variation across factors such as:

- Employer type and industry
- Testing reason (e.g., pre-employment, random, post-accident)
- Geographic distribution
- Testing timeframes
- Laboratory analytical standards

When these variables are reasonably aligned, specimen-type differences become more interpretable and less likely to reflect population bias.

4.2 Observational Program Transitions

In the testing programs included in this portion of the analysis, employers transitioned from urine-based testing to oral fluid testing within the same operational environment. Because these programs involved the same employers, job categories, and testing populations across sequential time periods, they provide an opportunity to observe how specimen type influences detection outcomes in practice.

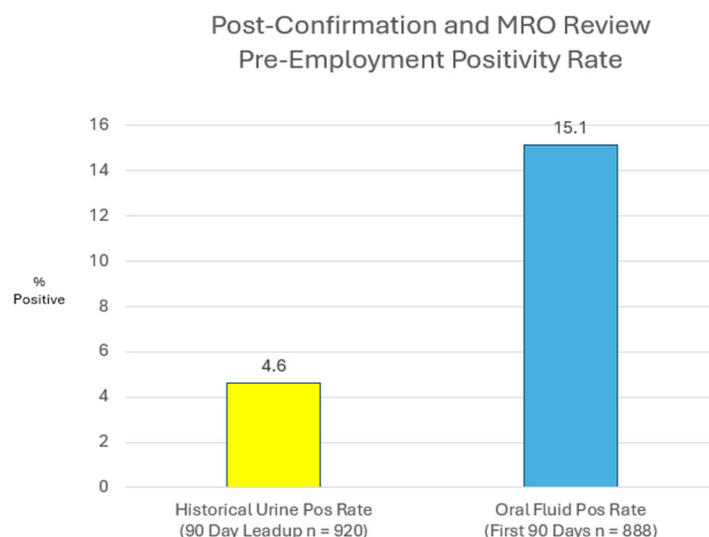
Across these transitions, a consistent pattern emerged: oral fluid testing identified drug use at materially higher rates than urine testing within the same employer populations.

This pattern appeared across distinct industries and operational settings, including warehouse operations and healthcare environments. While observational comparisons cannot fully eliminate all potential confounding

variables, the consistency of the observed effect across independent programs suggests that specimen type plays a meaningful role in detection outcomes.

These differences are illustrated in Figures 2 and 3, which compare urine positivity during pre-conversion periods with oral fluid positivity following program transitions within the same employer populations.

Figure 2 – Warehouse Program Transition: Urine vs Oral Fluid Positivity



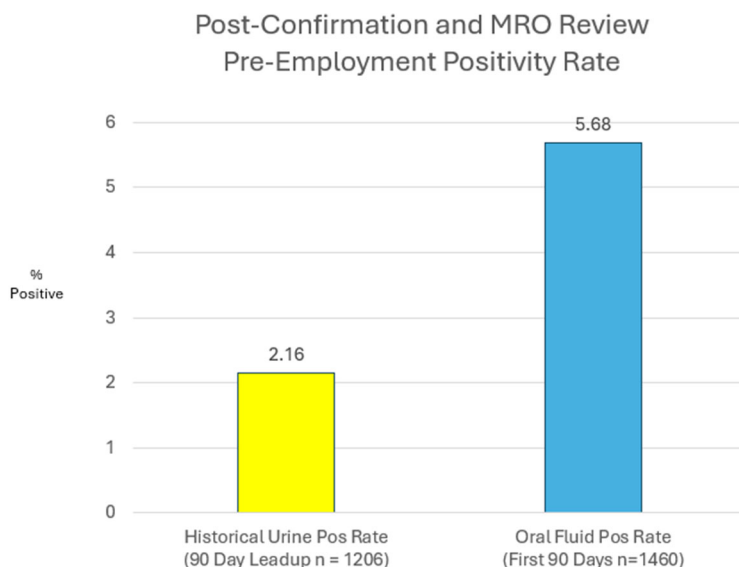
Interpretation

During the 90-day period prior to program conversion, urine testing identified positive results in **4.6%** of applicants—fewer than **5 positive results per 100 pre-employment tests**.

During the first 90 days following conversion to oral fluid testing, positivity increased to **15.1%**, representing **more than 15 positive results per 100 tests** within the same employer population.

In practical terms, oral fluid testing identified **more than 3x as many drug-using applicants** as urine testing under otherwise comparable program conditions.

Figure 3 – Hospital Program Transition: Urine vs Oral Fluid Positivity



Interpretation

Across multiple hospital programs, historical urine testing identified positive results in **2.16%** of tests, or slightly more than **2 positives per 100 tests**. Following conversion to oral fluid testing, positivity increased to **5.68%**, corresponding to nearly **6 positives per 100 tests**.

While the absolute positivity rates are lower than those observed in the warehouse setting, oral fluid testing still identified over **2.5x as many positive results** within comparable testing populations.

4.3 Contemporaneous DOT Comparisons

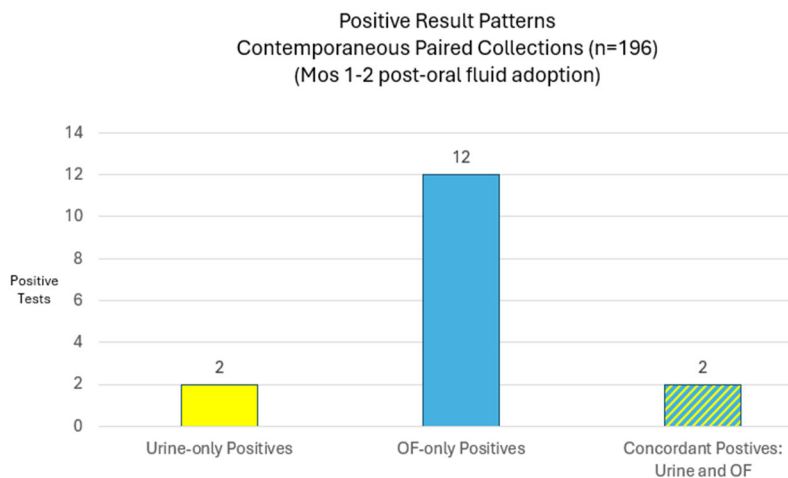
Additional insight can be obtained from datasets in which urine and oral fluid specimens were collected contemporaneously from the same donor population during the same testing events. In these cases, the timing of testing, employer context, and donor population are effectively held constant.

Under these conditions, the analytical component of testing is largely controlled. Modern laboratories operating under accreditation standards perform both urine and oral fluid testing with comparable analytical rigor. When the same laboratory infrastructure processes both specimen types, analytical sensitivity differences are unlikely to explain meaningful divergence in results.

Yet even under these conditions, oral fluid testing has been observed to identify drug use at materially higher rates than urine testing. Because the same donors, testing events, and laboratory standards are involved, the remaining differentiating factor is the specimen collection system itself.

As illustrated in Figure 4, comparison of paired DOT urine and non-DOT oral fluid specimens collected from the same donors during the same testing events shows oral fluid testing identifying substantially more positive results despite identical donor populations, timing, and laboratory conditions.

Figure 4 - Contemporaneous DOT Urine and Oral Fluid Testing Comparison



Interpretation:

In this paired dataset, urine and oral fluid specimens were collected from 196 donors during the same pre-employment testing events, allowing a direct comparison of detection outcomes under identical testing conditions.

DOT urine testing identified positive results in **2.04%** of tests (4 positives among 196 paired samples), while oral fluid testing identified positives in **7.14%** of tests (14 positives among the same 196 donors) – **a ratio of 3.5 to 1**.

Examination of discordant results further illustrates this difference. Among the paired specimens, **12 cases were identified by oral fluid testing alone**, compared with **2 cases identified by urine testing alone**, while **2 cases were positive in both specimen types**. Oral fluid testing produced **6x as many discordant positive results as urine testing (12 vs. 2)** within the same donor population. Because both specimens were collected at the same time from the same donors, these discordant results cannot be attributed to population variation or testing conditions. Instead, they highlight how specimen collection systems differ in their ability to capture authentic, contemporaneous drug use.

4.4 System Sensitivity as the Primary Differentiator

When testing conditions are aligned and laboratory performance is effectively controlled, the primary remaining determinant of detection effectiveness becomes the integrity of the specimen itself.

Oral fluid testing increases system sensitivity by improving the reliability of specimen collection. Observed collections, continuous line-of-sight during specimen acquisition, and direct donor identification reduce opportunities for specimen circumvention efforts such as substitution or adulteration.

The result is a testing system that more consistently obtains a specimen originating from the intended donor at the moment the test is administered.

Across these program comparisons, observed differences are consistent with system-level factors influencing specimen integrity, including reduced opportunities for circumvention in observed oral fluid collections compared to less controlled urine collection environments.

These findings raise an important question: *if oral fluid testing consistently identifies drug use at higher rates in comparable testing environments, do similar patterns appear across broader industry datasets?*

5. Interpreting the Quest Drug Testing Index Through a Relative System Sensitivity Lens

The Quest Diagnostics Drug Testing Index (DTI) is one of the most widely cited datasets in workplace drug testing, reporting annual positivity trends across specimen types, drug categories, test reason, and employee classifications. In recent reporting years, the DTI reflects results from approximately 5-6 million urine tests and 1-2 million oral fluid tests annually across U.S. workplace testing programs.

To provide a stable basis for comparison, the analysis presented in this paper aggregates DTI reporting of General U.S. Workforce (non-Federally Mandated) test data across a five-year period (2020–2024), representing approximately 25-30 million urine tests and 6-8 million oral fluid tests as published in the 2025 DTI.

While observational comparisons of similar employee populations help isolate system-level effects within specific programs, the DTI offers a broader view of testing outcomes across millions of workplace drug tests conducted annually in the United States.

When interpreted through a **Relative System Sensitivity** framework, the patterns observed in the DTI provide an important form of external validation.

5.1 Relative System Sensitivity

Because the Quest Drug Testing Index does not report fully aligned specimen-type comparisons across identical testing populations, direct interpretation of raw positivity rates can be challenging. To provide a consistent basis for comparison, this analysis evaluates specimen performance using relative system sensitivity. Relative system sensitivity expresses oral fluid detection performance relative to urine testing within the same drug category. Urine is used as a proxy benchmark because it remains the dominant specimen type across most workplace testing programs.

Relative system sensitivity is calculated as:

$$\text{Relative System Sensitivity} = \text{Oral Fluid Positivity Rate} \div \text{Urine Positivity Rate}$$

In this framework, a value of **1.0** indicates equivalent detection rates between specimen types. Values greater than 1.0 indicate higher detection rates in oral fluid testing relative to urine testing, values less than 1.0 indicate lower detection rates in oral fluid relative to urine testing.

This approach does not imply that urine represents a true baseline prevalence of drug use. Rather, it provides a consistent reference point for evaluating how specimen type influences observed detection outcomes across drug categories.

5.2 What the Quest DTI Measures Well

The DTI excels at identifying directional signals across large and diverse testing populations. Its strengths include:

- Multi-year trend analysis across specimen types
- Consistent national-scale methodology
- Stratification by test reason (e.g., pre-employment, random)
- Reporting across individual drug classes

Because of its scale, the DTI provides valuable insight into how specimen types perform across the broader U.S. workplace testing landscape.

5.3 Limitations of Aggregate Datasets

Despite its value, the DTI is not designed to function as a controlled comparison between specimen types. Several factors remain outside its analytical scope, including:

- Whether the same employers use both specimen types
- Whether donor populations are directly comparable
- Whether tests are paired or randomized within programs
- Variation in collection site procedures
- Differences in donor expectations and behaviors across testing environments

For these reasons, the DTI should not be interpreted as establishing causality. Instead, it functions as a large-scale observational dataset that highlights persistent detection differences across specimen types.

5.4 Drug Class Patterns and Behavioral Incentives

One of the most informative aspects of the DTI dataset is the variation in detection differences across drug classes. The divergence between urine and oral fluid positivity rates is not uniform; instead, it follows a pattern that closely mirrors donor behavioral incentives within workplace testing programs.

From a purely analytical perspective, this outcome is somewhat counterintuitive. Urine testing typically has longer detection windows for many substances, which might be expected to produce higher positivity rates if analytical sensitivity alone determined outcomes.

The persistence of higher oral fluid detection rates therefore suggests that factors outside the laboratory are influencing results. Under a system sensitivity framework, this pattern becomes easier to interpret.

Drug classes associated primarily with illicit use—such as marijuana, cocaine, PCP and methamphetamine—show the largest divergence between specimen types. In these categories, oral fluid testing consistently identifies substantially higher positivity rates than urine testing. Because these substances are rarely associated with legitimate prescription use, individuals testing positive for them face a high likelihood of adverse employment consequences. As a result, the incentive to evade detection is strongest.

Under these conditions, weaknesses in the collection system become most consequential. Urine testing environments that permit opportunities for circumvention techniques described earlier may allow significant drug use to go undetected. Observed oral fluid collection, by contrast, substantially reduces these opportunities, increasing the likelihood that the specimen analyzed originates from the intended donor at the moment the test is administered.

In contrast, drug classes frequently associated with legitimate prescription use tend to exhibit much smaller divergence between specimen types with urine testing demonstrating higher positivity rates in the drug classes included in this analysis – opioids and amphetamine. This outcome is consistent with the structure of workplace drug testing programs, where laboratory positives for prescription medications are routinely reconciled through medical review processes. Because valid prescriptions can explain positive laboratory results, donors using legitimately prescribed medications have little incentive to evade detection.

When the incentive to circumvent the testing process is reduced or eliminated, differences in system sensitivity between specimen types become far less pronounced. Under these circumstances, the longer detection window associated with urine testing may produce equal or greater positivity rates for certain drug classes.

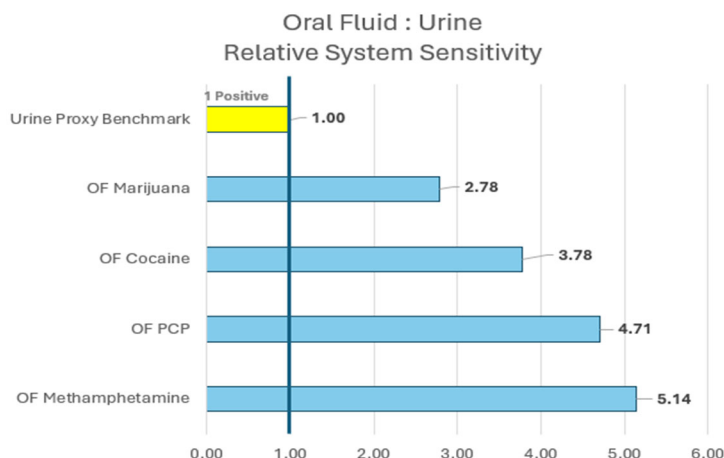
Taken together, these patterns reinforce the central premise of the system sensitivity framework. Differences in detection outcomes between urine and oral fluid testing are unlikely to be driven by analytical performance. Instead, they reflect how donor incentives interact with the design of the collection system.

Drug classes associated with illicit use create strong incentives for circumvention, and in those environments specimen integrity becomes the dominant factor influencing detection effectiveness. Drug classes associated with legitimate prescription use create far weaker incentives for evasion, and the differences between specimen types narrow accordingly.

This behavioral alignment helps explain why oral fluid testing consistently demonstrates higher detection rates in illicit drug categories across both observed program comparisons and large-scale industry datasets.

The drug-class patterns described above are illustrated in Figures 5 and 6, which compare relative system sensitivity between oral fluid and urine testing across selected drug categories reported in the Quest Drug Testing Index.

Figure 5 – Illicit Drug Classes

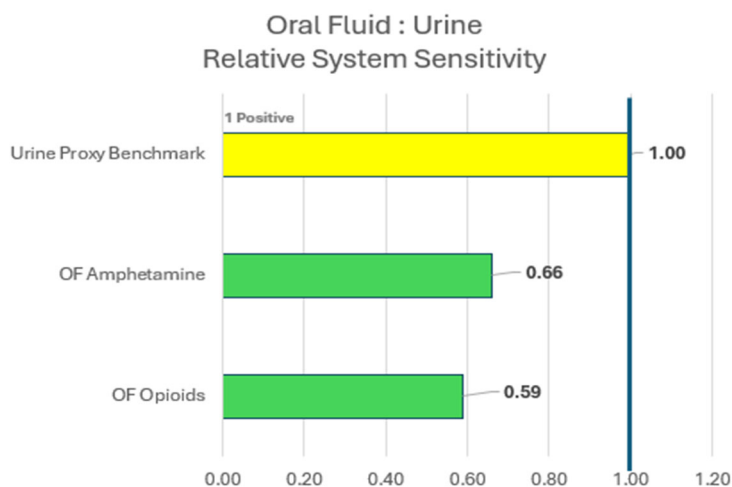


Interpretation

Across illicit drug categories reported in the Quest Drug Testing Index, oral fluid testing demonstrates substantially higher relative system sensitivity than urine testing. In illicit drug categories included in this analysis, oral fluid detection rates range from **2.78x – 5.14x higher** than those observed with urine testing.

When viewed through a relative system sensitivity framework, this pattern suggests that a meaningful proportion of illicit drug use may go undetected in urine-based testing programs. In practical terms, these ratios imply that **urine testing may be missing significantly more illicit drug use than it identifies** within certain drug categories.

Figure 6 – Prescription-Associated Drug Categories



Interpretation

Drug categories frequently associated with legitimate prescription use show **substantially smaller divergence** between specimen types.

Because valid prescriptions can explain laboratory positives through medical review processes, **donors have less incentive to evade detection through circumvention** of the specimen collection process.

Under these conditions, urine’s longer detection window may produce equal or higher positivity rates.

While specimen types reflect different detection profiles, the evidence presented here suggests that in typical workplace testing environments, oral fluid testing provides a more reliable measure of real-world illicit drug use. In settings where opportunities for circumvention are inherently limited—such as directly observed collections or certain unannounced testing scenarios—the longer detection window associated with urine testing may offer situational advantages. However, these contexts are not representative of most testing environments and may introduce operational tradeoffs, including increased collection time and logistical complexity. As such, differences between specimen types are best understood as context-dependent rather than balanced tradeoffs across typical workplace testing conditions.

5.5 Population Signals Reinforcing Comparison Findings

When viewed alongside the transitional program comparisons and contemporaneous paired collection data described in the previous sections, the DTI serves as an important validation layer.

Comparison datasets demonstrate that oral fluid testing identifies drug use at substantially higher rates when testing conditions are held constant. The DTI then shows that similar patterns appear across millions of workplace tests conducted across diverse industries and employer populations.

Taken together, these datasets suggest that the observed detection differences are not isolated anomalies but rather the predictable outcome of how specimen type interacts with collection system design.

However, population-level datasets alone cannot fully explain *why* these differences occur. Understanding the mechanisms behind these patterns requires examining how donor behavior, collection environments, and incentives interact within real-world testing systems.

The following section explores these behavioral dynamics and explains why collection design plays such a central role in determining system sensitivity.

6. Why Higher Detection Persists: Donor Behavior and Collection Design

The persistent detection advantage observed in oral fluid testing cannot be explained by analytical performance alone. Instead, it reflects how collection design interacts with donor behavior within real-world testing programs. Understanding this interaction is essential to explaining why oral fluid testing continues to identify illicit drug use at higher rates despite a narrower window of detection compared to urine.

6.1 Drug Testing as a Behavioral System

Drug testing programs do not operate in a vacuum. Donors are aware they are being tested, often understand the stakes involved, and may adjust their behavior accordingly. Over time, predictable patterns emerge as donors learn which testing systems are more forgiving, more predictable, or easier to circumvent.

Accordingly, the effectiveness of a testing program is shaped not only by laboratory performance, but by how donors perceive and respond to the collection process itself.

6.2 Privacy, Control, and Opportunity in Urine Collection

Urine collection protocols necessarily balance privacy with verification. Even when conducted in compliance with established procedures, urine collections typically involve:

- Physical separation between donor and collector during specimen production
- Limited or indirect observation
- Donor control over the specimen prior to handoff

These design elements introduce opportunities—both real and perceived—for circumvention techniques to be employed. Over time, they also shape donor expectations. In many testing populations, donors understand that:

- Valid prescription explanations may negate consequences
- Pre-collection planning can reduce risk
- The probability of undetected circumvention (including substitution) is non-zero

The result is not widespread misconduct, but systemic vulnerability—particularly among donors with strong incentives to avoid detection.

While the extent to which these behaviors occur is difficult to quantify directly, the observed differences in detection outcomes across specimen types are consistent with the influence of these system-level vulnerabilities.

6.3 How Oral Fluid Collection Alters Donor Incentives

Oral fluid collection fundamentally changes the collection environment by ensuring:

- Direct observation throughout specimen acquisition
- Continuous line-of-sight between donor and collector
- Immediate association between donor identity and specimen

These elements substantially reduce donor control over the specimen and compress the opportunity window for manipulation. Just as importantly, they alter donor perception.

When donors believe that evasion is unlikely, behavior changes. The testing system becomes not only more secure, but more credible.

6.4 Why the Effect Persists Across Drug Classes

The persistence of higher detection rates across multiple illicit drug classes further supports a behavioral explanation.

Where legitimate prescription use is common and consequences are mitigated through medical review, divergence between specimen types is often modest. Where use is primarily illicit and consequences are more inevitable, divergence is more pronounced.

This pattern is consistent with rational donor behavior interacting with collection design. System sensitivity differences are most visible where motivation to evade detection is highest.

6.5 Implications for Program Effectiveness

The evidence suggests that oral fluid testing improves detection not by raising analytical sensitivity, but by reducing behavioral escape routes within the testing system. By more reliably obtaining authentic specimens from the intended donor at the time of testing, oral fluid collection strengthens the integrity of the entire program.

This distinction matters. It explains why higher detection rates persist despite narrow detection windows. It also clarifies why further analytical optimization alone is unlikely to eliminate observed performance gaps.

The remaining question, therefore, is not whether oral fluid testing is effective, but why—despite clear advantages—its adoption has not scaled proportionally across the industry.

The next section addresses the structural and economic factors that have constrained broader adoption.

7. If the Data Is Clear, Why Hasn't Adoption Followed?

Given the persistent detection advantages demonstrated by oral fluid testing, the uneven pace of adoption may appear counterintuitive. This disconnect is not the result of skepticism about the data, it instead reflects structural, economic, and informational frictions that have constrained scale despite demonstrated effectiveness.

Understanding these constraints is essential to identifying a viable path forward.

7.1 An Information Gap Across the Testing Ecosystem

While large laboratories and national data aggregators have long had visibility into broad specimen-type performance trends, that visibility has not been evenly distributed across the drug testing ecosystem. Collection sites, third-party administrators (TPAs), and most significantly employers—who ultimately influence or make testing decisions—often lack access to comparative data that meaningfully illustrates system-level performance differences between urine and oral fluid testing.

In many cases, these stakeholders rely on secondary guidance, legacy assumptions, or incomplete value metrics rather than real-world performance comparisons. As a result, oral fluid testing is frequently evaluated through outdated frameworks or perceived as a niche alternative sold at a higher price point rather than a demonstrably more effective system in certain testing contexts.

This imbalance is not the result of deliberate withholding, but of structural incentives. Laboratories, which are closest to large-scale testing data, face natural disincentives to aggressively promote oral fluid adoption. These include the operational friction associated with change, collection device logistics challenges, and the reality that oral fluid testing often carries higher per-test costs and margin implications compared to urine testing. In such an environment, there is limited economic motivation to proactively challenge entrenched testing paradigms.

Consequently, the stakeholders most directly affected by testing effectiveness—employers—may never see the data necessary to fully evaluate oral fluid testing on its merits. Without visibility into system-level performance outcomes, adoption decisions default to familiarity and lower cost rather than effectiveness.

This paper seeks to address that gap by bringing empirical evidence and system-based analysis directly to the broader stakeholder community. Doing so enables informed evaluation of oral fluid testing not as an emerging technology, but as a mature system whose performance characteristics compare favorably with urine testing.

7.2 Adoption Lags Are Not Evidence Gaps

In established industries, adoption of new technology rarely stalls due to a lack of evidence—it stalls when implementation introduces friction for key stakeholders. Oral fluid testing presents a clear example of this dynamic.

Most stakeholders in the drug testing industry with an understanding of both urine and oral fluid test system performance characteristics acknowledge the advantages of oral fluid testing. Yet agreement on outcomes does not automatically translate into operational feasibility, particularly when costs, risks, and responsibilities are unevenly distributed.

7.3 Inventory and Capital Friction at the Collection Site Level

Unlike urine testing where collection supplies have no defined shelf and supply costs are bundled into lab test fees, oral fluid testing introduces a physical collection device that must be pre-purchased, stocked, managed, and replaced. For collection sites, this creates challenges that urine testing largely avoids:

- Devices represent upfront capital rather than no cost consumables
- Inventory may sit idle for unpredictable periods
- Expired devices create unrecoverable cost
- Demand variability complicates ordering and forecasting
- Efforts to minimize capital carry by maintaining minimal device inventory increase the risk of stock-outs, leaving collection sites unable to perform oral fluid collections when donors arrive

Even when per-test economics are favorable, these factors introduce perceived risk—particularly for smaller or lower-volume collection sites.

7.4 Misalignment Between Cost Timing and Revenue Events

In many oral fluid programs, device costs are incurred well before a corresponding billable test event occurs. This disconnect creates friction across the value chain:

- Collection sites absorb inventory risk without direct linkage to reimbursement
- Laboratories incur increased capital carry and additional device-related logistics complexity, all within a supply chain that deviates from legacy urine models in which collection supplies are bundled into the laboratory test fee.
- Employers face price increases that are difficult to justify in the absence of well understood system performance characteristics

When cost and revenue timing are misaligned and value metrics are unclear, adoption slows—particularly when underlying performance differences are not clearly understood by stakeholders.

7.5 Limited Visibility Across the Device Lifecycle

Once oral fluid collection devices leave a warehouse, visibility into their status is often limited. Stakeholders may not know whether a device is:

- Sitting on a shelf
- Expired
- Collected but not yet tested
- Lost or discarded

This lack of transparency complicates forecasting, reconciliation, and accountability. In contrast, urine testing relies on widely available, low-cost supplies that minimize the consequences of lost visibility.

7.6 Risk Concentration in the Wrong Places

Perhaps most importantly, the risks associated with oral fluid testing adoption are often concentrated among stakeholders least able to absorb them.

Collection sites are asked to assume inventory risk without guaranteed throughput. Laboratories may struggle with collection supply logistics and increased capitol carry. Employers may face pricing models that appear unappealing if the underlying value proposition is not clearly recognized.

None of these challenges reflect deficiencies in oral fluid testing itself. They reflect the absence of a system designed to coordinate responsibility, visibility, and timing across the testing ecosystem.

7.7 Why Structural Barriers Persist

Many commonly cited barriers to oral fluid adoption—such as device logistics, higher end user cost, training requirements, or stakeholder familiarity—are often considered in isolation. While each may influence adoption, addressing them individually does not resolve the underlying structural misalignment.

The core challenge is not any single factor, but the lack of coordination between device distribution, verified utilization, cost-revenue event timing, and clearly understood value metrics.

Without a mechanism to align these elements, oral fluid testing remains more difficult to scale operationally than urine testing—even when it delivers superior outcomes.

7.8 The Nature of the Problem Defines the Nature of the Solution

The barriers to oral fluid adoption are not primarily scientific or regulatory. They arise from a combination of information gaps and structural misalignment within the testing ecosystem.

As a result, meaningful progress will require a solution that:

- Preserves the performance advantages of oral fluid testing
- Promotes responsible behavior within the testing ecosystem
- Reduces operational burden
- Effectively aligns cost and revenue events
- Improves visibility without increasing complexity

The next section outlines the characteristics such a system must possess—without prescribing a specific implementation—focusing instead on the principles required to support scalable adoption.

8. What a Scalable Oral Fluid System Requires

The uneven adoption of oral fluid testing does not reflect limitations in the technology itself, but rather how the current testing ecosystem manages device flow, financial responsibility, and operational visibility, leaving many employers with little insight into detection effectiveness and reinforcing the assumption that long-standing urine-based programs are performing as intended.

Providing employers and program decision-makers with clear system-level performance visibility is the first step toward generating market demand—but demand alone cannot resolve adoption friction if the underlying system remains misaligned. A scalable oral fluid operating model requires a framework that addresses visibility, accountability, financial timing, and incentive alignment across all stakeholders.

The objective is not disruption—it is progression: education to drive awareness, awareness to generate demand, and demand being met with coordinated system alignment.

8.1 Device Lifecycle Visibility

A scalable system must treat oral fluid collection devices as traceable assets rather than anonymous consumables.

This includes the ability to:

- Monitor device distribution from manufacturer shipment through laboratory testing
- Distinguish between deployed, unused, expired, and utilized devices
- Reconcile device flow with confirmed laboratory testing events
- Reduce uncertainty surrounding stranded or idle inventory

Improved lifecycle visibility reduces forecasting friction, minimizes expiration risk, and enables more accurate operational planning across stakeholders.

8.2 Alignment of Economic Timing

One of the most persistent sources of friction in oral fluid testing is the misalignment between when costs are incurred and when revenue is realized.

A scalable model must better align:

- Device cost allocation

- Proper device usage resulting in a confirmed specimen collection
- Laboratory reporting
- Billing and reimbursement events

When economic responsibility more closely tracks verified utilization, risk becomes distributed more proportionally. This reduces hesitation at the collection site level and improves predictability for manufacturers and laboratories alike, while creating a pathway to incorporate collection supply costs within laboratory test fees, consistent with legacy urine testing models.

8.3 Risk Distribution Proportional to Control

In the current environment, inventory risk, expiration risk, and utilization uncertainty are often concentrated at points in the system with limited visibility or leverage.

A more resilient framework distributes risk according to:

- Operational control
- Access to information
- Ability to influence outcomes

Such alignment improves confidence in adoption decisions and reduces the need for defensive ordering behavior.

8.4 Operational Simplicity

Any scalable solution must reduce complexity, not introduce it.

Collection sites should not be required to adopt burdensome administrative processes to mitigate structural risk. Similarly, laboratories and manufacturers should not rely on manual reconciliation to maintain visibility.

Automation, standardized tracking, and integrated data flow are essential components of a scalable system. The goal is seamless coordination that operates in the background without increasing frontline workload.

8.5 Preservation of Clinical and Compliance Integrity

Importantly, improvements in lifecycle management and economic alignment must not compromise:

- Chain-of-custody integrity
- Regulatory compliance
- Laboratory independence
- Employer oversight

A coordinating framework should strengthen—not dilute—the safeguards already embedded within the testing process.

8.6 A Coordinating Layer, Not a Replacement Layer

The future of oral fluid adoption does not require displacing manufacturers, laboratories, collection sites, or TPAs. It requires a coordinating layer capable of synchronizing the flow of devices, data, and economic accountability.

Such a layer would:

- Enhance transparency without centralizing control
- Reward responsible management and usage of collection device assets
- Align incentives without dictating outcomes
- Reduce friction without disrupting established relationships

By addressing visibility, economic timing, and risk distribution in a coordinated manner, the industry can transition from debating specimen type to optimizing system performance.

The opportunity is not incremental improvement but coordinated system alignment—enabling oral fluid testing to scale as a cohesive, high-performing testing model.

9. Implications for Stakeholders

The evidence and structural analysis presented in this paper suggest that oral fluid testing effectiveness is not solely an analytical question, but a system-design question. Recognizing this distinction carries practical implications for each stakeholder in the testing ecosystem.

9.1 Employers

For employers, the central objective of any drug testing program is risk mitigation. The analysis presented herein suggests that specimen type materially influences real-world detection effectiveness due to differences in system sensitivity.

Implications for employers include:

- Re-evaluating specimen choice in light of detection effectiveness rather than historical precedent
- Considering whether current testing modalities align with organizational risk tolerance
- Evaluating whether lower-cost testing modalities deliver the level of detection effectiveness assumed in program design and budgeting—are current programs achieving the outcomes they are intended to produce?
- Recognizing that identifying drug-using applicants during the pre-employment screening process may materially reduce workforce risk exposure by limiting the entry of individuals with active drug use into the workforce

Employers are often downstream of system design decisions. Increased visibility into how specimen type influences detection outcomes enables more informed procurement and policy decisions.

9.2 Collection Sites

Collection sites operate at the front line of the testing process and often absorb disproportionate operational and financial risk.

Benefits and Implications include:

- Operational efficiency gains associated with observed oral fluid collections, including reduced restroom dependency, elimination of shy bladder delays, and improved donor throughput
- Greater scheduling flexibility when collections can be performed in standard office space rather than dedicated restroom facilities
- Reduced staffing requirements as same gender observed collections are eliminated
- Acknowledgement that collection devices must be managed and used responsibly
- Recognition that expiration and stock-out risk stem from structural misalignment
- Opportunity to advocate for models that better align device cost with confirmed utilization
- Potential to reduce defensive ordering behavior if lifecycle visibility improves

Collection sites are not barriers to adoption; they are stakeholders whose operational realities must be reflected in system design.

9.3 Laboratories

Laboratories play a central role in shaping industry practice. They possess both analytical capability and broad visibility into testing patterns.

Implications include:

- Recognition that specimen-type performance differences reflect system effects, not laboratory shortcomings
- Opportunity to support transparent, data-driven dialogue across stakeholders
- Consideration of models that align specimen flow with economic accountability
- Long-term strategic positioning in an environment where effectiveness may increasingly influence demand
- Development of oral fluid pricing structures designed to maintain, not increase, margins relative to urine testing

Laboratories are uniquely positioned to help bridge information gaps and contribute to system-level optimization.

9.4 Manufacturers

Manufacturers bear upstream financial exposure associated with device production, distribution, and credit extension.

Implications include:

- Reduced expiration and stranded inventory risk through improved lifecycle visibility
- Enhanced forecasting accuracy with better linkage between distribution and utilization
- Greater confidence in scaling production when downstream friction is mitigated
- Increased demand for oral fluid testing reagents as constraints are removed and adoption increases
- Improved device traceability, enabling faster identification of quality issues, targeted field communications, and more efficient recall management when necessary

For manufacturers, improved coordination can reduce volatility and increase demand without altering core production operations.

9.5 Third-Party Administrators (TPAs)

TPAs serve as intermediaries between employers, collection sites, and laboratories. Their influence on specimen selection and program configuration is significant.

Implications include:

- Increased ability to advise clients using system-level performance data
- Opportunity to differentiate through data-informed policy guidance
- Reduced administrative friction when inventory and billing reconciliation improve
- Enhanced credibility when specimen recommendations align with measurable effectiveness

TPAs are often positioned to translate operational realities into employer decisions. Greater transparency supports more strategic advisory roles.

9.6 Regulators, Insurers, and Risk Stakeholders

Regulatory frameworks have historically prioritized analytical validity, standardization, donor rights, and legal defensibility—objectives that have been well served by urine-based testing protocols. However, the findings presented in this paper suggest that analytical equivalence does not necessarily translate into equivalent real-world effectiveness. As oral fluid testing becomes more widely adopted and its system-level advantages better understood, regulators may face an opportunity to reassess how program effectiveness is defined and measured. Incorporating concepts such as system sensitivity—alongside traditional analytical criteria—could support more comprehensive policy frameworks that better align testing protocols with the goal of identifying drug use in safety-sensitive populations to better protect workers and the public at large.

For insurers and other risk-focused entities, drug testing effectiveness intersects with claims exposure and workforce safety.

Implications include:

- Recognition that specimen selection may influence actual risk mitigation outcomes
- Potential to incorporate testing modality considerations into underwriting frameworks
- Opportunity to incentivize systems that demonstrably reduce undetected illicit drug use

While not traditionally central to specimen decisions, risk stakeholders may play an increasing role in shaping adoption incentives.

Closing Perspective

Across stakeholder groups, a common theme emerges: the performance advantages of oral fluid testing are increasingly clear, but structural misalignment has limited scale.

The path forward does not require uniform mandates or abrupt transition. It requires transparent and well understood performance data, aligned incentives, and system-level coordination that promotes confident decision-making.

When effectiveness and economics are aligned, adoption becomes a rational progression rather than a contested debate.

10. Conclusion: From Evidence to Action

Over the past two decades, oral fluid testing has moved from emerging alternative to validated methodology. The analytical capability exists. Regulatory frameworks are established. Large-scale datasets demonstrate persistent differences in real-world detection outcomes.

The central insight of this paper is not that oral fluid testing is analytically superior, but that it is systemically distinct. When evaluated through a relative system sensitivity framework, its higher detection rates are coherent, predictable, and persistent.

The remaining barriers to broader adoption are not rooted in uncertainty about performance. They are structural - stemming from information asymmetry, economic timing misalignment, inventory risk concentration, and misunderstood value metrics.

These are solvable problems.

Progress does not require disruption of existing stakeholders, nor does it require abandoning established testing programs. It requires alignment—between specimen flow and economic accountability, between operational control and risk exposure, and between data visibility and decision authority.

The industry now faces a choice. It can continue to evaluate specimen types through legacy assumptions and fragmented incentives, or it can acknowledge that testing effectiveness is inseparable from system design.

As the data becomes more widely understood, the question shifts from “Does oral fluid work?” to “How do we make it work at scale?”

That question invites collaboration.

Manufacturers, laboratories, collection sites, employers, TPAs, regulators, and risk stakeholders each operate within constraints shaped by history and economics. Yet all share a common objective: reliable detection, operational stability, and sustainable program design.

The next phase of oral fluid adoption will not be driven by persuasion alone. It will be driven by coordinated systems that reduce friction while preserving effectiveness.

The evidence presented in this paper suggests that the industry may be operating with an incomplete view of testing effectiveness—one that assumes analytical capability alone is sufficient to ensure real-world performance. However, when evaluated at the system level, differences in specimen design, collection dynamics, and donor behavior produce materially different outcomes.

This does not imply that the intent of existing programs is flawed, but it does raise an important question: *are legacy testing models capturing drug use to the extent we assume they are?*

Addressing that question from an economic perspective introduces practical considerations. Oral fluid testing is often perceived as incrementally more expensive than urine on a per-test basis. However, when evaluated in the context of system sensitivity, that comparison becomes incomplete. If one testing approach consistently identifies materially more illicit drug use under real-world conditions, the relevant question is not simply cost per test, but cost per effective detection. In that context, even a modest increase in per-test cost may represent a favorable trade-off if it results in substantially higher identification rates and an accompanying reduction in risk. Employers routinely make similar decisions in other areas of risk management—accepting incremental cost in exchange for improved reliability and reduced exposure. Drug testing programs should be evaluated through the same lens.

Quantifying the economic implications of these detection differences represents an important area for future analysis, particularly in evaluating the downstream impact of undetected drug use on workplace safety, productivity, and risk exposure.

The path forward is not speculative. The data, the technology, and the regulatory foundation are already in place. What is needed is the recognition that better outcomes are achievable—and the collective willingness to rethink long-standing assumptions and work together to realize them.

Acknowledgements

The author would like to acknowledge several individuals for their valuable input, technical insight, and discussion during the development of this analysis. Their perspectives helped refine the methodology and interpretation of the data presented.

The views and conclusions expressed in this paper are those of the author and do not necessarily reflect the views of the individuals acknowledged.

Eran Bendavid, MD, MS, Stanford Medicine

Jason Hudson, PhD, Scientific Director and Laboratory Director, Quest Toxicology

David Kuntz, PhD, Executive Director of Analytical Toxicology/Regulatory Affairs, CRL

Thomas Kupiec, PhD, CEO, ARL Bio Pharma and DNA Solutions, Inc.

James K. Lee, MD, C-MRO, Compliance Resource Group, Inc.

Gerardo Ramos, PhD, Director, CLO; Forensic Toxicologist, qVector Laboratory Resources

Mark Villoria, PhD, Scientific Affairs, Abbott Toxicology

APPENDIX

Appendix A – Methodology

A.1 System Sensitivity Framework

System sensitivity describes how effectively a drug-testing program identifies true drug use under real-world conditions. Unlike analytical sensitivity, which evaluates laboratory detection once a specimen is received, system sensitivity incorporates factors influencing specimen integrity prior to laboratory analysis, including donor behavior, specimen collection design, and opportunities for circumvention.

Analytical Sensitivity

Drug present in specimen → Positive laboratory result

Factors influencing analytical sensitivity include:

- Laboratory cutoff concentrations
- Instrument and assay performance
- Confirmation chemistry

System Sensitivity

Drug use occurred → Positive laboratory result

System sensitivity, while impacted by analytical sensitivity, incorporates factors that influence whether drug use is ultimately detected within the testing system, including:

- Detection window
- Donor behavior
- Risk of circumvention efforts
 - substitution
 - adulteration
- Specimen integrity

Conceptual Framework

Drug Use Occurs → Specimen Collection → Specimen Integrity → Laboratory Detection → Reported Result →
System Sensitivity

Two testing approaches with identical analytical performance characteristics may produce **substantially different real-world detection outcomes at the system level.**

A.2 Relative System Sensitivity Calculation

Relative System Sensitivity

To compare specimen types at the program level, this analysis introduces **Relative System Sensitivity**, using urine (the legacy testing matrix) as a proxy benchmark.

Calculated as:

$$\text{Relative System Sensitivity} = \text{Oral Fluid Positive Rate} \div \text{Urine Positive Rate}$$

Values greater than **1.0** indicate higher real-world detection effectiveness in oral fluid relative to urine testing, values less than **1.0** indicate lower real-world detection effectiveness in oral fluid relative to urine.

This framework allows specimen types to be compared not only based on laboratory performance, but on their ability to identify drug use within the full testing system.

A.3 Quest DTI Drug Class Grouping Methodology

For purposes of comparative analysis, individual analytes were grouped into broader drug classes to align with common reporting conventions in workplace drug testing programs.

The opioid category includes positive rates for the following opiate sub-categories: codeine, morphine, hydrocodone, hydromorphone, oxycodone, oxymorphone, and 6-acetylmorhpine. Fentanyl is not included in DTI oral fluid reporting and is therefore excluded from this analysis for both specimen types.

Separation of amphetamine and methamphetamine into sub-category positive rates within the amphetamines category was performed through aggregation of industry specific positive rates for each sub-category as published in the “Industry Insights” section on pages 42-43 of the DTI.

These groupings were designed to reflect how results are typically interpreted at the program level while preserving the ability to evaluate meaningful differences in detection performance across specimen types.

Appendix B – CRG Dataset Description

B.1 Purpose of the CRG Comparative Datasets

In addition to aggregate industry data, this paper draws on three comparative datasets developed from the Compliance Resource Group (CRG) testing activity to evaluate real-world differences between urine and oral fluid testing. These datasets were selected because they provide program-level visibility into specimen-type performance under real-world workplace testing conditions.

Two of the datasets are derived from non-DOT observational testing populations, while one dataset is derived from a contemporaneous DOT testing population. Taken together, these comparisons illustrate how detection outcomes differ between specimen types across distinct testing environments.

These datasets are not intended to function as nationally representative prevalence studies. Rather, they are used to examine how specimen type influences real-world detection performance within defined testing populations.

All reported positive results reflect laboratory-confirmed findings, including confirmatory testing in accordance with standard workplace drug-testing protocol.

B.2 Non-DOT Observational Dataset 1

Non-DOT Observational Dataset 1 compares urine and oral fluid testing outcomes within a non-DOT workplace testing environment using observational program data surrounding a program transition from urine to oral fluid testing. The purpose of this dataset is to assess whether specimen-type differences observed at the industry level are also evident within a real-world operational setting.

Because this dataset is observational, it should not be interpreted as a randomized or paired-sample comparison. Its value lies in its ability to reflect specimen-type performance under actual program conditions, where donor behavior, collection conditions, and specimen integrity all contribute to final detection outcomes.

This dataset is used in the paper to evaluate relative positivity differences between urine and oral fluid testing in a non-DOT context

Table B1 – Non-DOT Observational Dataset 1 (Warehouse Program Transition)

Industry	Day 1–90 OF Tests	Day 1–90 OF Positives	OF Positivity (%)	90-Day Pre-Conversion Urine Tests	90-Day Pre-Conversion Urine Positives	Urine Positivity (%)	Relative System Sensitivity OF: Urine
Warehouse	888	134	15.09%	920	42	4.57%	3.30x
Total	888	134	15.09%	920	42	4.57%	3.30x

As shown in Table B1, oral fluid testing identified substantially higher positive rates during the first 90 days following program conversion compared with historical urine testing during the preceding 90-day period.

B.3 Non-DOT Observational Dataset 2

Non-DOT Observational Dataset 2 provides a second real-world non-DOT comparison of testing outcomes surrounding a program transition from urine to oral fluid testing. It serves as an independent observational dataset used to evaluate whether the directional findings observed in Dataset 1 persist across a separate non-DOT testing population.

As with Dataset 1, this comparison is observational and reflects actual program performance rather than controlled laboratory equivalence. Its inclusion strengthens the analysis by reducing the likelihood that specimen-type differences observed in one non-DOT population are isolated or incidental.

Used together, the two non-DOT observational datasets provide consistent real-world context for evaluating specimen-type performance outside the DOT environment.

Table B2 – Non-DOT Observational Dataset 2 (Hospital Program Transitions)

Industry	Day 1–90 OF Tests	Day 1–90 OF Positives	OF Positivity (%)	90-Day Pre-Conversion Urine Tests	90-Day Pre-Conversion Urine Positives	Urine Positivity (%)	Relative System Sensitivity OF: Urine
Hospital A	415	27	6.51%	348	8	2.30%	2.83x
Hospital B	99	8	8.08%	64	1	1.56%	5.18x
Hospital C	72	3	4.17%	82	2	2.44%	1.71x
Hospital D	875	45	5.14%	712	15	2.11%	2.44x
Total	1,461	83	5.68%	1,206	26	2.16%	2.63x

As shown in Table B2, oral fluid testing identified substantially higher positive rates during the first 90 days following program conversion compared with historical urine testing during the preceding 90-day period.

B.4 Contemporaneous DOT Dataset

The contemporaneous DOT dataset provides a comparison within a DOT-regulated testing environment, allowing specimen-type performance to be evaluated within a structured and time-aligned workplace testing population.

This dataset is described as contemporaneous because Federally Regulated (DOT) urine and company mandated non-DOT oral fluid specimens were collected **from the same donors at the same time**, allowing a more direct comparison of detection outcomes than would be possible using separated historical periods.

The inclusion of a DOT dataset provides an important contrast to the non-DOT observational populations and helps assess whether specimen-type performance differences remain visible within a regulated testing context.

All oral fluid positives fell within drug classes included in the DOT panel and were subject to confirmation and MRO review.

Table B3 – Contemporaneous DOT Dataset (Trucking Program Comparison)

Industry	Day 1–60 OF Tests	Day 1–60 OF Positives	OF Positivity (%)	60-Day DOT Urine Tests	60-Day DOT Urine Positives	DOT Urine Positivity (%)	Relative System Sensitivity OF: Urine
Carrier A	132	8	6.06%	132	2	1.52%	3.99x
Carrier B	64	6	9.38%	64	2	3.13%	3.00x
Total	196	14	7.14%	196	4	2.04%	3.50x

As shown in Table B3, oral fluid testing produced higher positive rates than contemporaneous DOT (Federally Mandated) urine testing within the same donor population.

B.5 Role of the CRG Comparative Datasets in This Paper

These three CRG comparative datasets supplement broader industry-level findings by providing program-level evidence of how urine and oral fluid testing perform in actual workplace testing environments.

Their role in this paper is not to establish national prevalence, but to:

- evaluate whether specimen-type differences persist across multiple real-world datasets
- compare findings across non-DOT and DOT populations
- support interpretation of detection differences through the **system sensitivity framework described in Appendix A**

Detailed results derived from these datasets are presented in the main body of the paper and expanded in subsequent appendices.

Appendix C – Quest Drug Testing Index (DTI) Data

C.1 Purpose of the DTI Comparison

In addition to the program-level datasets described in Appendix B, this analysis incorporates data from the Quest Diagnostics Drug Testing Index (DTI) to provide broader industry context. The DTI is one of the largest publicly reported summaries of workplace drug testing activity in the United States and provides annual reporting on drug test positivity trends.

The purpose of incorporating DTI data in this analysis is to determine whether patterns observed in the CRG program datasets are also reflected in a large national testing dataset.

C.2 Data Source and Scope

The Quest Diagnostics Drug Testing Index compiles results from millions of workplace drug tests conducted annually across a wide range of industries and employer programs. The DTI reports aggregate positivity trends by specimen type, drug category, test reason, and employee classification.

However, the DTI reporting structure does not allow for a direct comparison of overall positivity rates between urine and oral fluid testing. In particular, marijuana positivity is reported separately for oral fluid testing and excluded from the overall oral fluid positivity metric.

Because of this reporting structure, direct comparisons of overall specimen-type positivity rates can be misleading. For this reason, the analysis presented in this paper focuses on **drug-category comparisons**, which allow specimen types to be evaluated on a more consistent basis.

C.3 Drug Class Detection Comparisons

DTI reporting provides positivity rates for several individual drug categories, allowing comparisons between urine and oral fluid testing within specific analyte groups.

Examining drug-category trends helps illuminate differences in detection outcomes that may be masked when evaluating aggregate positivity rates alone. Differences in positivity between specimen types reflect a combination of pharmacokinetics, detection windows, and specimen collection characteristics.

Table C1 – Quest Drug Testing Index Drug Category Comparisons

Drug Category	Oral Fluid Aggregate Positivity (%)	Urine Aggregate Positivity (%)	Relative System Sensitivity (Oral Fluid: Urine)
Opioids*	.45	.76	.59x
Amphetamine	.858	1.305	.66x
Marijuana	11.56	4.16	2.78x
Cocaine	.87	.23	3.78x
PCP	.0358	.0076	4.71x
Methamphetamine	.606	.118	5.14x

**Opioids include combined Quest DTI categories for codeine/morphine, hydrocodone/hydromorphone, oxycodone/oxymorphone, and 6-acetylmorphine. Fentanyl is not included in DTI oral fluid reporting and is therefore excluded from this analysis for both specimen types.*

As shown in Table C1, oral fluid testing demonstrates higher relative system sensitivity across several drug categories commonly associated with illicit drug use. This pattern is consistent with behavioral incentives within workplace testing programs, where individuals seeking to avoid detection have historically focused on circumventing urine-based testing methods.

C.4 Role of the DTI Data in This Paper

The DTI dataset serves as an external industry reference for the program-level findings described in Appendix B. While the CRG datasets provide detailed comparisons within specific testing programs, the DTI provides a broader view of positivity trends across the U.S. workplace testing landscape.

Taken together, the CRG datasets and DTI reporting provide complementary perspectives on how specimen type influences real-world drug detection outcomes

Sources

1. CRG Laboratories / The Compliance Resource Group

Comparative workplace drug testing datasets used in this analysis, including warehouse, healthcare, and DOT transportation program comparisons. Data derived from workplace testing programs administered by The Compliance Resource Group and analyzed by CRG Laboratories.

2. Quest Diagnostics

Drug Testing Index™: Trends in Workforce Drug Use.

Quest Diagnostics Employer Solutions.

<https://www.questdiagnostics.com/content/dam/corporate/restricted/documents/drug-testing-index/2025-Drug-Testing-Index-report.pdf>

