

# Provider Misinterpretation, Documentation, and Follow-Up of Definitive Urine Drug Testing Results



Isaac Chua, MD, MPH<sup>1,2</sup>, Athena K. Petrides, PhD<sup>2,3</sup>, Gordon D. Schiff, MD<sup>2,4</sup>,  
Jaime R. Ransohoff, BA<sup>3,4</sup>, Michalis Kantartjis, BS<sup>3,4</sup>, Jocelyn Streid, BS<sup>2,5</sup>,  
Christiana A. Demetriou, PhD<sup>6,7</sup>, and Stacy E. F. Melanson, MD, PhD<sup>2,3</sup>

<sup>1</sup>Department of Psychosocial Oncology and Palliative Care, Dana-Farber Cancer Institute, Boston, MA, USA; <sup>2</sup>Harvard Medical School, Boston, MA, USA; <sup>3</sup>Department of Pathology, Brigham and Women's Hospital, Boston, MA, USA; <sup>4</sup>Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA; <sup>5</sup>Harvard Kennedy School, Boston, MA, USA; <sup>6</sup>Department of Primary Care and Population Health, University of Nicosia Medical School, Nicosia, Cyprus; <sup>7</sup>The Cyprus School of Molecular Medicine, The Cyprus Institute of Neurology and Genetics, Nicosia, Cyprus.

**BACKGROUND:** Urine drug testing (UDT) is an essential tool to monitor opioid misuse among patients on chronic opioid therapy. Inaccurate interpretation of UDT can have deleterious consequences. Providers' ability to accurately interpret and document UDT, particularly definitive liquid chromatography-tandem mass spectrometry (LC-MS/MS) results, has not been widely studied.

**OBJECTIVE:** To examine whether providers correctly interpret, document, and communicate LC-MS/MS UDT results.

**DESIGN:** This is a retrospective chart review of 160 UDT results (80 aberrant; 80 non-aberrant) between August 2017 and February 2018 from 5 ambulatory clinics (3 primary care, 1 oncology, 1 pain management). Aberrant results were classified into one or more of the following categories: illicit drug use, simulated compliance, not taking prescribed medication, and taking a medication not prescribed. Accurate result interpretation was defined as concordance between the provider's documented interpretation and an expert laboratory toxicologist's interpretation. Outcome measures were concordance between provider and laboratory interpretation of UDT results, documentation of UDT results, results acknowledgement in the electronic health record, communication of results to the patient, and rate of prescription refills.

**KEY RESULTS:** Aberrant results were most frequently due to illicit drug use. Overall, only 88 of the 160 (55%) had any documented provider interpretations of which 25/88 (28%) were discordant with the laboratory toxicologist's interpretation. Thirty-six of the 160 (23%) documented communication of the results to the patient. Communicating results was more likely to be documented if the results were aberrant compared with non-aberrant (33/80 [41%] vs. 3/80 [4%],  $p < 0.001$ ). In all cases where provider interpretations were discordant with the laboratory interpretation, prescriptions were refilled.

**CONCLUSIONS:** Erroneous provider interpretation of UDT results, infrequent documentation of interpretation,

lack of communication of results to patients, and prescription refills despite inaccurate interpretations are common. Expert assistance with urine toxicology interpretations may be needed to improve provider accuracy when interpreting toxicology results.

**KEY WORDS:** compliance monitoring; chronic pain; urine drug testing; opioid; liquid chromatography-tandem mass spectrometry; oncology; primary care; substance use disorder; diagnostic error.

## Abbreviations

LC-MS/MS	Liquid chromatography-tandem mass spectrometry
BWH	Brigham and Women's Hospital
EHR	Electronic health records
UDT	Urine drug testing
SUD	Substance abuse disorder
SAMSHA	Substance Abuse and Mental Health Services Administration
COT	Chronic opioid therapy

J Gen Intern Med 35(1):283-90  
DOI: 10.1007/s11606-019-05514-5  
© Society of General Internal Medicine 2019

## INTRODUCTION

Accurately diagnosing opioid misuse is important in the setting of the current opioid epidemic. In 2017, more than 70,200 people in the USA died from drug overdose with a 5.9-fold increase in the total number of deaths involving opioids since 1999.<sup>1</sup> Heroin use is strongly associated with prior history of opioid misuse,<sup>2</sup> and approximately one-quarter of patients on chronic opioid therapy (COT) misuse opioids.<sup>3</sup> Urine drug testing (UDT) is used to monitor patients on COT.<sup>4-14</sup> UDT can assess medication adherence and detect aberrant behavior, such as diversion or non-prescribed drug use. However, the clinical utility and providers' ability to correctly interpret UDT is of growing concern.<sup>12-19</sup>

Most providers have a limited understanding of opioid metabolism and cross-reactivity of immunoassays and are not proficient in the interpretation of UDT results.<sup>20-26</sup> Family physicians and pediatricians performed poorly on surveys designed to assess knowledge of UDT by immunoassay.<sup>22,</sup>

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s11606-019-05514-5>) contains supplementary material, which is available to authorized users.

Received March 12, 2019

Revised August 13, 2019

Accepted October 2, 2019

Published online November 11, 2019

<sup>24</sup> Most emergency physicians were unable to name the drugs tested at their hospital.<sup>23</sup> Two surveys of residents revealed that they were confident but inaccurate when interpreting UDT.<sup>20,26</sup> Physicians with experience in COT fared no better, rarely answering more than half of the questions correctly in a survey assessing interpretive knowledge of toxicology immunoassays.<sup>21</sup> However, in one case, a successful educational intervention was implemented to improve residents' understanding of UDT, indicating that medical professionals would benefit from more education on results interpretation.<sup>27</sup>

Clinical decisions based on inaccurate result interpretations can have serious consequences for patients.<sup>28</sup> Providers could erroneously suspect aberrant behavior, potentially resulting in discontinuation of medications. Moreover, illicit drug use, undisclosed prescription opioid use, or simulated compliance could go undetected if the results are incorrectly interpreted. Some recent guidelines and research in the field of pain management recommend replacing traditional immunoassay screens with more definitive testing such as liquid chromatography-tandem mass spectrometry (LC-MS/MS).<sup>12, 29–34</sup> However, provider ability to interpret UDT by LC-MS/MS results has not been previously studied.

This retrospective study examines provider ability to interpret and document the results of LC-MS/MS UDT for patients on COT, assesses the frequency of communicating UDT results to the patient, and investigates the potential clinical consequences of provider interpretations.

## METHODS

### Study Setting and Sample Selection

This retrospective chart review was approved by the Partners IRB committee and conducted at Brigham and Women's Hospital (BWH), a 749-bed tertiary care hospital in Boston. Eligible patients included adult patients ( $\geq 18$  years old) on COT for whom UDT by LC-MS/MS was ordered in 5 ambulatory clinics between August 2017 and February 2018, which included 3 primary care clinics, 1 oncology/palliative care clinic, and 1 pain management clinic. Among the three primary care locations, one (a community clinic) had a higher prevalence of patients with substance use disorders (SUD), so this location was analyzed separately.

From this pool of UDTs, we reviewed 183 randomly selected results to achieve our study design of 160 cases (80 aberrant and 80 non-aberrant). We excluded 23 cases. Reasons included documented direct communication between the provider and the laboratory for interpretation assistance (7 cases), patients not receiving COT (10 cases), and restricted patient information in the electronic health record (EHR) (2 cases). If multiple UDT specimens were submitted prior to the next follow-up appointment, we randomly selected one and excluded the rest (4 cases). We defined the next appointment as the nearest occurring encounter from the date of the randomly selected UDT.

## Measures

The primary outcome was concordance between provider and laboratory interpretation of UDTs. Secondary outcomes include provider documentation of UDT results, acknowledgment of results in the EHR, communication of results to the patient, and rate of prescription refills during follow-up appointments.

Provider level covariates included age, gender, education/degree, clinical training (residency and/or fellowship), years in practice, and practice location. We calculated age by subtracting the UDT collection date from the provider's date of birth obtained from the EHR. We obtained gender, education/degree, clinical training, and practice location from cross-referencing the provider's BWH online profile, Massachusetts Board of Registration database, and LinkedIn profiles. We calculated years in practice by subtracting the UDT collection date from the year of graduation from residency or fellowship (MDs), nurse practitioner school (NPs), or pharmacy school (Pharm Ds).

Patient characteristics included age, race, SUD history, and type of pain (i.e., cancer, non-cancer, or both).

## Laboratory Information and Result Interpretation

The BWH chemistry laboratory performs UDT by immunoassay as well as LC-MS/MS. However, LC-MS/MS is considered the gold standard due to its higher sensitivity and specificity to assess patient medication compliance and to detect aberrant behavior.<sup>12,29</sup> BWH performs approximately 3500 UDT panels by LC-MS/MS annually. The panel detects opioids, benzodiazepines, and stimulants but not tetrahydrocannabinol. Please see Table A-1 online for a complete list of drugs and metabolites included in the panel.

One of two laboratory directors (A.P., S.M.) interpreted each LC-MS/MS UDT result in conjunction with prescribed medications (as documented in the EHR) at the time of sample collection. Both laboratory directors are experts in clinical chemistry and toxicology with more than 25 combined years of laboratory experience. Results were categorized as either "aberrant" or "non-aberrant," by comparing the metabolites to the medications prescribed. Results were categorized as "aberrant" if they showed evidence of one or more of the following: illicit drug use, simulated compliance, not taking a prescribed drug(s), or taking a drug(s) not prescribed. The interpretation was classified as "non-aberrant" if results were consistent with the prescribed medications. For patients taking opioids on a PRN basis, both the presence and absence of the drug were considered non-aberrant. See Table A-2 online for definitions of aberrant and non-aberrant subcategories.

## Chart Review

Medical records were reviewed by one of three clinicians (I.C., S.M., J.S.) using a standardized review instrument. To assess

reliability of the expert laboratory interpretations and medical record review, 10% of the cases were re-reviewed by a second physician or laboratory director. Please refer to Table A-3 online for the definitions of key terms.

**Statistical Analysis**

Cohen’s kappa statistical test was performed to assess agreement between two reviewers. The Fleiss kappa test, which assesses agreement between more than two reviewers, was used to assess concordance of medical record review.<sup>35</sup>

We investigated associations between categorical variables using Pearson’s chi-squared ( $\chi^2$ ) test and Fisher’s exact test when expected cell counts were < 5. We used Student’s *t* test for 2 group comparisons of normally distributed numeric variables and Mann-Whitney *U* (2 groups) and Kruskal-Wallis (> 2 groups) to compare non-normally distributed outcome variables. The association between number of UDTs ordered and age and years in practice was investigated using Spearman rank correlation. Univariate and multivariate analyses of provider characteristics against categorical outcome variables were performed using logistic regression models. Clinical training was excluded from the multivariate models due to reduced number of observations (i.e., not applicable for NPs and PharmDs) and its strong correlation with location. All calculations were performed using the statistical software package STATA version SE15.

**RESULTS**

**Demographics**

Among the 160 cases included for analysis, there were 136 unique patients and 39 providers who ordered UDTs. Most COT patients were white and were treated for non-cancer pain (Table 1). 42.6% of patients had a diagnosis of SUD. Most providers were men and physicians (Table 2). The most common residency and fellowship completed were internal medicine (59.0%) and pain medicine (12.8%), respectively. The median number of years of clinical experience was 10 (IQR 1.75–18.25). Half of the providers practice in a primary care setting.

**Table 1 Patient Characteristics**

Characteristic	Unique patients (N=136)
Age, mean (SD)	53.3 (12.1)
Female, n (%)	69 (50.7)
Race, n (%)	
White	82 (60.3)
Black	16 (11.8)
Hispanic	16 (11.8)
Other*	9 (6.6)
Missing data	13 (9.6)
Substance use disorder, n (%)	58 (42.6)
Type of pain, n (%)	
Non-cancer	107 (78.7)
Cancer	24 (17.6)
Both	5 (3.7)

\*Includes American Indian, Asian, Multiple Races, or Declined

**Table 2 Provider Characteristics**

Characteristic	Total (n = 39)
Age, median (IQR)	44 (35–53)
Female, n (%)	18 (46.2)
Degree, n (%)	
MD	32 (82.0)
NP	6 (15.4)
Pharm D	1 (2.6)
Residency, n (%)	
Internal medicine	23 (59.0)
Anesthesia	6 (15.4)
Emergency medicine	1 (2.6)
Internal medicine/pediatrics	1 (2.6)
Psychiatry	1 (2.6)
N/A*	7 (17.9)
Fellowship, n (%)	
Single	
Pain medicine	5 (12.8)
General internal medicine	2 (5.1)
Hospice/palliative medicine	2 (5.1)
Hematology oncology	1 (2.6)
Rheumatology	1 (2.6)
Multiple	
Hospice/palliative medicine + pain medicine	2 (5.1)
Hospice/palliative medicine + hematology oncology	2 (5.1)
Addiction medicine + pain medicine + forensic psychiatry	1 (2.6)
None/residency only	15 (38.5)
N/A*	7 (17.9)
Years of clinical practice, median (IQR)	10 (1.75–18.25)
Practice location, n (%)	
Primary care	16 (41.0)
Oncology	10 (25.6)
Pain medicine	9 (23.1)
Community clinic†	4 (10.3)

\*N/A applies to PharmD and NP providers since they do not complete a clinical residency or fellowship

†Community clinic is a primary care clinic that was analyzed separately due to a higher prevalence of SUD

**Urine Drug Testing Orders**

The median number of UDTs ordered per patient/year was 2 (IQR 1–6). The number of UDTs ordered varied significantly based on the provider’s gender, degree/education, clinical training, and practice location (Table 3). Women (4 [IQR 4–10]), pharmacists (9 [IQR 3–13]), physicians who completed internal medicine training (3 [IQR 1–9]), and providers practicing in the community clinic (12 [IQR 6–14]) ordered the most UDTs per patient/year in their respective subcategories.

**Reviewer Agreement**

Laboratory directors achieved perfect agreement for toxicology interpretations. For medical record review, reviewers achieved perfect agreement (*K* = 1.00) on the presence of COT and documentation of result communication with the patient. Reviewers achieved substantial agreement on provider concordance with the laboratory (*K* = 0.601), documentation of an interpretation (*K* = 0.741), history of SUD (*K* = 0.655), and result acknowledgment in the EHR (*K* = 0.737). Reviewers achieved fair agreement on prescription refills (*K* = 0.256).

**Table 3** Number of UDTs Ordered per Patient/Year by Provider Characteristics

Characteristic	Median (IQR)	<i>p</i> value
Age	3 (1–7)	0.97
Sex		
Male	2 (1–5)	0.003
Female	4 (4–10)	
Degree		
MD	3 (1–6)	0.002
NP	2 (1–7)	
Pharm D	9 (3–13)	
Residency		
Internal medicine*	3 (1–9)	< 0.001
Anesthesia	1 (1–2)	
Emergency medicine	1 (1–1)	
Psychiatry	1 (1–1)	
Years in practice	3 (1–7)	0.12
Practice location		
Primary care	3 (2–4)	< 0.001
Oncology	3 (1–7)	
Pain medicine	1 (1–2)	
Community clinic <sup>†</sup>	12 (6–14)	

\*Internal medicine/pediatrics was combined with internal medicine

<sup>†</sup>Community clinic is a primary care clinic that was analyzed separately due to a higher prevalence of SUD

## Laboratory Test Result Interpretations

Among aberrant cases ( $n = 80$ ), 37 (46%) contained illicit substances, of which the most common were cocaine (70%), fentanyl (27%), and morphine suggestive of heroin use (16%). Twenty-two (28%) were interpreted as simulated compliance, of which 19 (86%) had urine creatinine < 20 mg/dL (suggestive of urine dilution) and 3 (14%) had abnormally high levels of the parent drug (suggestive of dropping the drug into the urine sample). Eighteen (23%) samples lacked evidence that patients were taking their prescribed medication(s). Thirty-four (43%) were positive for medication(s) that patients were not prescribed. Among clinic locations, oncology (52 (65%)) and community clinic (42 (53%)) have the highest percentage of aberrant results (Fig. 1).

## Provider Interpretation, Documentation, and Acknowledgement of Results

Provider interpretations of UDT were documented for 88/160 (55%) cases. Among the documented interpretations ( $n = 88$ ), 63 (72%) provider interpretations were concordant with laboratory interpretations and 25 (28%) were discordant. Results were more likely to be aberrant when interpretations were discordant compared with when interpretations were concordant (23/25 [92%] vs. 27/63 [43%],  $p < 0.001$ ). Simulated compliance was more frequently seen in discordant cases than concordant cases (28% vs. 6%,  $p = 0.006$ ).

Providers electronically acknowledged 146/160 (91%) results in the EHR. Of those not acknowledged ( $n = 14$ ), 6 (43%) were aberrant and 8 (57%) were non-aberrant. Thirty-six of the 160 (23%) cases had documented communication of the results to the patient. Providers documented communication of results more often when results were aberrant versus non-

aberrant (33/80 [41%] vs. 3/80 [4%],  $p < 0.001$ ). When there was no documentation of result communication to the patient ( $n = 124$ ), 47 (38%) were aberrant and 77 (62%) were non-aberrant.

## Follow-up and Prescription Refills

One hundred fifty-four of the 160 (96%) cases had a documented follow-up visit in the EHR. Among non-aberrant cases with documented follow-up, 74/76 (97%) opioid prescriptions were refilled. Among aberrant cases with documented follow-up, 67/78 (86%) prescriptions were refilled. Although not statistically significant ( $p = 0.51$ ), 57/63 (90%) prescriptions were refilled for concordant interpretations and 25/25 (100%) prescriptions were refilled for discordant interpretations. If there was no documentation of an interpretation, 59/72 (82%) prescriptions were refilled.

## Association between Outcomes and Provider Characteristics

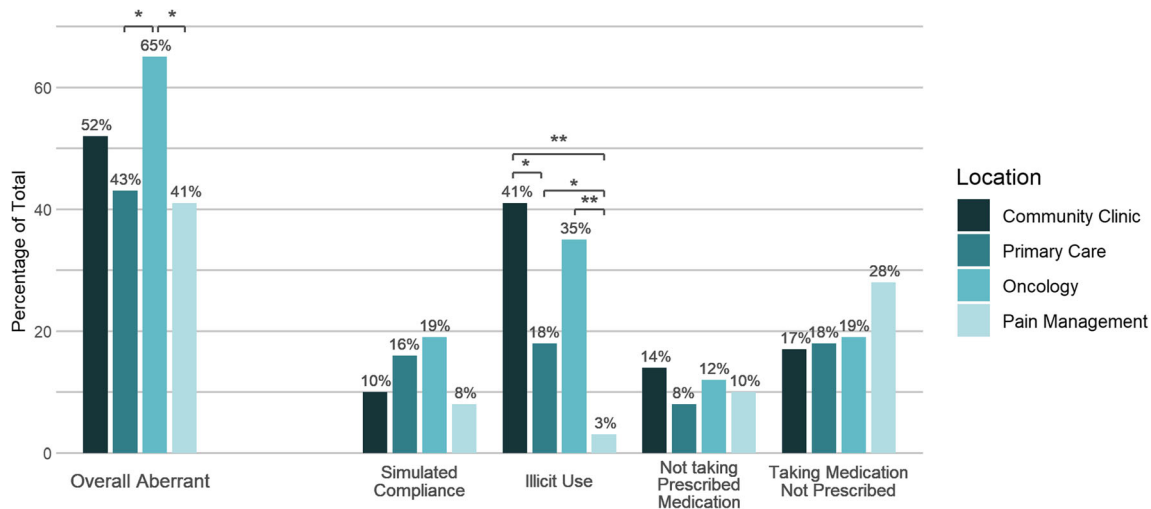
Interpretation concordance and prescription refills were not significantly associated with provider characteristics in both the univariate (Table A-4 online) and multivariate analyses (Table A-5 online). In the multivariate analyses (Table 4), odds of documenting UDT results increase for each year out of training (OR 1.15 [95% CI 1.01–1.31]) but decrease for each year of provider age (OR 0.84 [95% CI 0.71–0.99]). Odds of documentation increase for providers practicing in the oncology clinic (OR 18.61 [95% CI 4.54–76.30]) and community clinic (OR 8.22 [95% CI 1.52–44.16]). Compared with male providers, female providers are less likely to document UDT results (OR 0.16 [95% CI 0.04–0.66]) and communicate results to patients (OR 0.13 [95% CI 0.03–0.58]).

## DISCUSSION

Our study demonstrated that providers across medical specialties have difficulty on correctly interpreting UDT by LC-MS/MS. They also infrequently document their interpretation and rarely document communication of results with patients. Furthermore, prescription refills frequently occurred when no documentation of interpretation was present and when provider interpretation was discordant with laboratory interpretation.

Strengths of this study include examining provider UDT interpretation accuracy in a live practice environment, using the new gold standard for opioid compliance monitoring (i.e., LC-MS/MS), and defining four non-mutually exclusive categories of aberrant UDT results similar to categories previously described in the literature.<sup>4,8,13</sup> Compared with testing clinicians on simulated cases, studying provider behavior in a live environment allows us to capture actual behavior following an interpretation (e.g., prescription refills, results communication). The high sensitivity and specificity of LC-MS/MS allows the laboratory and clinicians to interpret the results with





**Figure 1 Breakdown of aberrant results by location. The percentage of aberrant cases for community clinic (dark blue bars), primary care (medium dark blue bars), oncology (medium light blue bars), and pain management (light blue bars) is depicted. The overall aberrant cases are broken down by location into 4 categories (simulated compliance, illicit use, not taking prescribed medication, taking medication not prescribed) and the percentage of cases in each category is depicted. Some cases may have results in more than one category. Results that differ significantly by location are shown with a \* ( $p < 0.05$ ) and \*\* ( $p < 0.001$ ).**

fewer concerns of false-positive or false-negative results.<sup>12,29–33</sup> Moreover, using well-described categories of aberrancy allowed us to compare our findings with preexisting literature. Our prevalence of aberrancy subtypes resembled prior literature with some differences (e.g., similarly high rates of illicit drug use but our study had higher rates of simulated compliance).<sup>8,13,36</sup>

Providers incorrectly interpreted UDT approximately one-quarter of the time across all disciplines. The prevalence of diagnostic error of opioid misuse and compliance is unknown, but diagnostic error in the ambulatory setting is estimated to occur in 5% of encounters.<sup>37</sup> We found that providers had the most inaccuracies interpreting aberrant results, which is not surprising given the complexity of the opioid metabolic pathway and ability to recognize more nuanced aspects of interpretation (e.g., urine

dilution). The number of factors that influence drug metabolism (e.g., pharmacogenetics) and the time window of detection of drugs and/or metabolites in urine can also complicate result interpretation.<sup>29</sup>

Providers were significantly more likely to misinterpret cases of simulated compliance. Fourteen percent of aberrant results were consistent with simulated compliance, of which 86% was due to concerns for dilution. The Substance Abuse and Mental Health Services Administration (SAMSHA) considers a creatinine  $< 20$  mg/dL suspicious for urine dilution,<sup>38</sup> which can occur either by ingesting large amounts of liquid prior to providing the specimen or adding water directly to their specimen to dilute the urine to obscure illicit drug use. Only 3 patients had evidence of simulated compliance by dropping a drug directly into their urine. Although aberrancy by simulated compliance occurs less frequently than illicit

**Table 4 Multivariate Analysis of Acknowledgement, Documentation, and Communication with Patient of UDT Results by Provider Characteristics**

Characteristic (n)	Acknowledgement		Documentation		Communication	
	OR (95% CI)	p value	OR (95% CI)	p value	OR (95% CI)	p value
Age (n = 39)	0.76 (0.56–1.03)	0.07	0.84 (0.71–0.99)	0.04	0.91 (0.77–1.08)	0.27
Sex						
Male (n = 21)	Ref		Ref		Ref	
Female (n = 18)	0.09 (0.00–5.77)	0.26	0.16 (0.04–0.66)	0.01	0.13 (0.03–0.58)	0.008
Degree						
MD (n = 32)	Ref		Ref		Ref	
NP* (n = 6)	–	–	1.86 (0.37–9.30)	0.45	0.67 (0.15–3.06)	0.61
Pharm D (n = 1)	0.05 (0.00–2.75)	0.14	4.39 (0.83–23.11)	0.08	2.37 (0.37–15.06)	0.36
Years in practice (n = 39)	1.48 (1.08–2.01)	0.01	1.15 (1.01–1.31)	0.03	1.02 (0.88–1.19)	0.75
Practice location						
Primary care (n = 16)	Ref		Ref		Ref	
Oncology* (n = 10)	–	–	18.61 (4.54–76.30)	$< 0.001$	4.25 (1.12–16.15)	0.03
Pain medicine (n = 9)	0.01 (0.00–1.53)	0.08	3.02 (0.77–11.75)	0.11	0.13 (0.01–1.25)	0.08
Community clinic† (n = 4)	0.10 (0.00–14.32)	0.36	8.22 (1.52–44.16)	0.01	2.39 (0.40–14.45)	0.34

\*Empty cells in this row reflect the model's inability to generate an effect estimate for this particular predictor because the outcomes of the subcategory were all the same (e.g., either all yes or all no)

†Community clinic is a primary care clinic that was analyzed separately due to a higher prevalence of SUD

drug use, this type of aberrancy is often overlooked by providers, thus warrants increased vigilance.

Documentation of UDT interpretations was infrequent, which is consistent with prior studies.<sup>25,39</sup> Although almost all providers acknowledged the results in the EHR by clicking “marked as reviewed,” documentation of UDT interpretations occurred only 55% of the time. This suggests that providers may be acknowledging results to clear their queues without critically reviewing them, or they are unsure how to interpret the results and thus avoid documenting their interpretations. Furthermore, only 23% of providers documented communication of results to the patient. Although providers may have verbally communicated results to the patient without documenting the discussion, this practice exposes providers to liability since opioid prescribing is considered a high-risk endeavor.<sup>40</sup>

The number of UDT ordered per patient/year varied by provider characteristics and clinic location. Although some variation may be explained by differences among each clinic’s patient population (e.g., SUD prevalence), variations were also associated with the provider’s training background. This is not surprising since consensus recommendations are clinical society-based.<sup>41–43</sup> Guidelines vary from no specific recommendations about UDT ordering frequency to repeating UDT at least 8 times a year.<sup>6,12,14,29,44,45,46</sup>

Inappropriate prescribing in the setting of UDT misinterpretation is seldom recognized as a potential cause of harm and risks contributing to the opioid epidemic. Nonetheless, the decision to change opioid prescribing practices is complex and hinges upon multiple patient variables, even among aberrant UDT results.<sup>39,47</sup> In our study, 85% of aberrant cases and 100% of discordant interpretations had a refill at the subsequent follow-up appointment. Without additional information, we cannot definitively conclude whether these refills were truly inappropriate. However, our findings do suggest that there is a potential for inappropriate prescribing with such misinterpretations.

The lack of association between provider characteristics and interpretation concordance or prescription refills is likely due to lack of power. The association of practice location with increased odds of communication of results and documentation may be due to higher prevalence of illicit drug use in the oncology and community clinics. Although this finding was statistically significant, the wide confidence intervals were due to small sample sizes in each clinic. The decreased odds of documenting UDT results among female providers and the inverse relationship between provider age and years of practice are harder to explain. Future studies examining provider attitudes towards UDT documentation and result communication would help clarify this finding.

Providers might benefit from expert assistance when interpreting UDT results. There is evidence to suggest a collective intelligence approach is associated with higher diagnostic accuracy.<sup>48</sup> One study implemented a pharmacist e-consult UDT interpretation service that could help guide

provider follow-up actions.<sup>49</sup> Other strategies may include involving pathologists and laboratory services as part of the diagnostic team.<sup>50–52</sup>

This study has several limitations. First, inter-observer variability during chart reviews may introduce information bias. However, a standardized template was utilized for both medical record review and laboratory interpretations and the Cohen and Fleiss Kappa statistics suggest that interpretations and reviews were consistent among reviewers. Second, by randomly selecting cases, providers who routinely ordered UDT more frequently or ordered additional UDT specifically to address aberrancy were more likely to be included. Because these providers may be more knowledgeable at interpreting results, inclusion of these providers conservatively biases results towards greater interpretation knowledge when compared with those who order UDT less frequently, thus strengthening the generalizability of our conclusions. Third, we were unable to exclude cases when providers contacted the lab and no documentation of such communication was charted. Such communication occurs infrequently and would also make our results more conservative, underestimating the overall problem of provider misinterpretations. Fourth, laboratory directors based their interpretations on the results and medications listed in the EHR. They did not have access to the Massachusetts Prescription Monitoring Program fill data because review of this database is not permitted for research studies. Such access could have provided a more comprehensive account of prescribed and dispensed controlled substances. Fifth, we assumed that patients with PRN prescriptions were non-aberrant if drugs and/or metabolites were not detected. However, providers may have been expecting these patients to have the drug(s) and/or metabolite(s) in the urine, leading to erroneous classification as non-aberrant. Finally, providers may have copied and pasted previous notes, resulting in an overestimation of documentation.

## CONCLUSION

This is the first study to evaluate provider interpretation and documentation of definitive UDT by LC-MS/MS. Erroneous provider interpretation of UDT results, infrequent documentation of result interpretation, lack of documented communication of results to patients, and prescription refills despite discordant interpretation are common. Expert assistance with urine toxicology interpretations may be needed to improve provider accuracy when interpreting toxicology results.

---

**Acknowledgments:** Dr. Schiff acknowledges support from the Gordon and Betty Moore Foundation and Controlled Risk Insurance Company (CRICO) for the diagnostic error and improvement of the work.

**Corresponding Author:** Stacy E. F. Melanson, MD, PhD; Department of Pathology Brigham and Women's Hospital, Boston, MA, USA (e-mail: semelanson@bwh.harvard.edu).

#### Compliance with Ethical Standards:

This retrospective chart review was approved by the Partners IRB committee and conducted at Brigham and Women's Hospital (BWH).

**Conflict of Interest:** The authors declare that they do not have a conflict of interest.

## REFERENCES

- National Institute on Drug Abuse (NIDA). Overdose Death Rates. Available at: <https://www.drugabuse.gov/related-topics/trends-statistics/overdose-death-rates>. Accessed 23 Sept 2019.
- Jones CM. Heroin use and heroin use risk behaviors among nonmedical users of prescription opioid pain relievers – United States, 2002–2004 and 2008–2010. *Drug Alcohol Depend*. 2013;132(1–2):95–100.
- Vowles KE, McEntee ML, Julnes PS, Frohe T, Ney JP, van der Goes DN. Rates of opioid misuse, abuse, and addiction in chronic pain. *Pain*. 2015;156(4):569–576.
- Christo PJ, Manchikanti L, Ruan X, et al. Urine drug testing in chronic pain. *Pain Physician*. 14(2):123–143.
- Manchikanti L, Manchukonda R, Pampati V, et al. Does random urine drug testing reduce illicit drug use in chronic pain patients receiving opioids? *Pain Physician*. 2006;9(2):123–129.
- Trescott AM, Helm S, Hansen H, et al. Opioids in the management of chronic non-cancer pain: an update of American Society of the Interventional Pain Physicians' (ASIPP) Guidelines. *Pain Physician*. 2008;11(2 Suppl):S5–S62.
- Melanson SEF, Kredlow MI, Jarolim P. Analysis and interpretation of drug testing results from patients on chronic pain therapy: a clinical laboratory perspective. *Clin Chem Lab Med*. 2009;47(8).
- Michna E, Jamison RN, Pham L-D, et al. Urine toxicology screening among chronic pain patients on opioid therapy: frequency and predictability of abnormal findings. *Clin J Pain*. 2007;23(2):173–179.
- Starrels JL, Becker WC, Alford DP, Kapoor A, Williams AR, Turner BJ. Systematic review: treatment agreements and urine drug testing to reduce opioid misuse in patients with chronic pain. *Ann Intern Med* 2010;152(11):712.
- Katz NP, Sherburne S, Beach M, et al. Behavioral monitoring and urine toxicology testing in patients receiving long-term opioid therapy. *Anesth Analg* 2003;97(4):1097–1102.
- Heit HA, Gourlay DL. Urine drug testing in pain medicine. *J Pain Symptom Manag* 2004;27(3):260–267.
- Argoff CE, Alford DP, Fudin J, et al. Rational urine drug monitoring in patients receiving opioids for chronic pain: consensus recommendations. *Pain Med* 2018;19(1):97–117.
- Turner JA, Saunders K, Shortreed SM, et al. Chronic opioid therapy urine drug testing in primary care: prevalence and predictors of aberrant results. *J Gen Intern Med* 2014;29(12):1663–1671.
- Manchikanti L, Atluri S, Trescott AM, Giordano J. Monitoring opioid adherence in chronic pain patients: tools, techniques, and utility. *Pain Physician* 2008;11(2 Suppl):S155–80.
- Reisfield GM, Maschke KJ. Urine drug testing in long-term opioid therapy: ethical considerations. *Clin J Pain* 2014;30(8):679–684.
- Moeller KE, Kissack JC, Atayee RS, Lee KC. Clinical interpretation of urine drug tests: what clinicians need to know about urine drug screens. *Mayo Clin Proc* 2017;92(5):774–796.
- Ceasar R, Chang J, Zamora K, et al. Primary care providers' experiences with urine toxicology tests to manage prescription opioid misuse and substance use among chronic noncancer pain patients in safety net health care settings. *Subst Abus* 2016;37(1):154–160.
- Pesce A, West C, Egan City K, Strickland J. Interpretation of urine drug testing in pain patients. *Pain Med* 2012;13(7):868–885.
- Nafziger AN, Bertino JS. Utility and application of urine drug testing in chronic pain management with opioids. *Clin J Pain* 2009;25(1):73–79.
- Starrels JL, Fox AD, Kumins H V, Cunningham CO. They don't know what they don't know: internal medicine residents' knowledge and confidence in urine drug test interpretation for patients with chronic pain. *J Gen Intern Med* 2012;27(11):1521–1527.
- Reisfield GM, Bertholf R, Barkin RL, Webb F, Wilson G. Urine drug test interpretation: what do physicians know? *J Opioid Manag* 2007;3(2):80–86.
- Reisfield GM, Webb FJ, Bertholf RL, Sloan PA, Wilson GR. Family physicians' proficiency in urine drug test interpretation. *J Opioid Manag*. 3(6):333–337.
- Durback LF, Scharman EJ, Brown BS. Emergency physicians perceptions of drug screens at their own hospitals. *Vet Hum Toxicol* 1998;40(4):234–237.
- Levy S, Harris SK, Sherritt L, Angulo M, Knight JR. Drug testing of adolescents in ambulatory medicine. *Arch Pediatr Adolesc Med* 2006;160(2):146.
- Passik SD, Schreiber J, Kirsh KL, Portenoy RK. A chart review of the ordering and documentation of urine toxicology screens in a cancer center: do they influence patient management? *J Pain Symptom Manag* 2000;19(1):40–44.
- Suzuki J, Garayalde S, Dodoo M, Rodriguez C. Psychiatry residents' and fellows' confidence and knowledge in interpreting urine drug testing results related to opioids. *Subst Abus* 2018;39(4):518–521.
- Aultman W, Fett J, Lauster C, Muench S, Halalau A. Urine drug test interpretation: an educational program's impact on resident knowledge and comfort level. *MedEdPORTAL*. 2018;14(14).
- Alford DP, Cohen ML, Reynolds EE. How would you manage opioid use in these three patients?: grand rounds discussion from Beth Israel Deaconess Medical Center. *Ann Intern Med* 2017;166(7):506–513.
- Jannetto PJ, Bratanow NC, Clark WA, et al. Executive summary: American Association of Clinical Chemistry Laboratory Medicine Practice Guideline—using clinical laboratory tests to monitor drug therapy in pain management patients. *JALM*. 2018;2(4):489–526.
- Cross TG, Hornshaw MP. Can LC and LC-MS ever replace immunoassays? *J Appl Bioanal* 2016;2(4):108–116.
- Pesce A, Rosenthal M, West R, et al. An evaluation of the diagnostic accuracy of liquid chromatography-tandem mass spectrometry versus immunoassay drug testing in pain patients. *Pain Physician*. 13(3):273–281.
- Petrides AK, Melanson SE. LC-MS/MS testing for monitoring compliance in pain management: the cost and technical benefits of a phased approach. *Clinical Laboratory News*. Available at: <https://www.aacc.org/publications/cln/articles/2016/february/lc-ms-ms-testing-for-monitoring-compliance-in-pain-management>. Accessed 23 Sept 2019.
- Snyder ML, Fantz CR, Melanson S. Immunoassay-based drug tests are inadequately sensitive for medication compliance monitoring in patients treated for chronic pain. *Pain Physician* 2017;20(2S):SE1–SE9.
- Darragh A, Snyder ML, Ptolemy AS, Melanson S. KIMS, CEDIA, and HS-CEDIA immunoassays are inadequately sensitive for detection of benzodiazepines in urine from patients treated for chronic pain. *Pain Physician*. 17(4):359–366.
- Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics* 1977;33(2):363–374.
- Turner JA, Saunders K, Shortreed SM, et al. Chronic opioid therapy risk reduction initiative: impact on urine drug testing rates and results. *J Gen Intern Med* 2014;29(2):305–311.
- Singh H, Meyer AND, Thomas EJ. The frequency of diagnostic errors in outpatient care: estimations from three large observational studies involving US adult populations. *BMJ Qual Saf* 2014;23(9):727–731.
- Substance Abuse and Mental Health Services Administration. Federal Guidelines for Opioid Treatment Programs.; 2015. Available at: <https://store.samhsa.gov/product/Federal-Guidelines-for-Opioid-Treatment-Programs/PEP15-FEDGUIDEOTP>. Accessed 23 Sept 2019.
- Morasco BJ, Krebs EE, Cavanagh R, Hyde S, Crain A, Dobscha SK. Treatment changes following aberrant urine drug test results for patients prescribed chronic opioid therapy. *J Opioid Manag* 2015;11(1):45–51.
- Hanscom R, Small M, Lambrecht A. A dose of insight: a data-driven review of the state of medication-related errors and liability in american healthcare. Available at: <https://www.coverys.com/Knowledge-Center/A-Dose-of-Insight-Medication-Related-Errors>. Accessed 23 Sept 2019.
- Owen GT, Burton AW, Schade CM, Passik S. Urine drug testing: current recommendations and best practices. *Pain Physician* 2012;15(3 Suppl):ES119–33.
- Benzon HT, Kendall MC, Katz JA, et al. Prescription patterns of pain medicine physicians. *Pain Pract* 2013;13(6):440–450.
- Gilbert JW, Wheeler GR, Mick GE, et al. Urine drug testing in the treatment of chronic noncancer pain in a Kentucky private neuroscience practice: the potential effect of Medicare benefit changes in Kentucky. *Pain Physician*. 13(2):187–194.

44. FSMB Federation of State Medical Board. Guidelines for the Chronic Use of Opioid Analgesics 2017. Available at: [http://www.fsmb.org/siteassets/advocacy/policies/opioid\\_guidelines\\_as\\_adopted\\_april-2017\\_final.pdf](http://www.fsmb.org/siteassets/advocacy/policies/opioid_guidelines_as_adopted_april-2017_final.pdf) Accessed 23 Sept 2019.
45. Center for Independent Healthcare Education. Urine drug testing in clinical practice: the art and science of patient care. 2015. Available at: [https://www.remitigate.com/wp-content/uploads/2015/11/Urine-Drug-Testing-in-Clinical-Practice-Ed6\\_2015-08.pdf](https://www.remitigate.com/wp-content/uploads/2015/11/Urine-Drug-Testing-in-Clinical-Practice-Ed6_2015-08.pdf). Accessed 23 Sept 2019.
46. California Academy of Family Physicians. Urine drug testing in clinical practice: the art and science of patient care. 2010. Available at: <https://www.healthcare.uiowa.edu/familymedicine/fpinfo/Docs/Drug%20screens%202015.pdf>. Accessed 23 Sept 2019.
47. **Morasco BJ, Krebs EE, Adams MH, Hyde S, Zamudio J, Dobscha SK.** Clinician response to aberrant urine drug test results of patients prescribed opioid therapy for chronic pain. *Clin J Pain* 2019;35(1):1–6.
48. **Barnett ML, Boddupalli D, Nundy S, Bates DW.** Comparative accuracy of diagnosis by collective intelligence of multiple physicians vs individual physicians. *JAMA Netw Open* 2019;2(3):e190096.
49. **Stammet MM, Spradley SS.** Evaluation of treatment changes following electronic consultation to a pharmacist-run urine drug testing service in a veterans healthcare system. *J Opioid Manag* 2016;12(6):389–395.
50. **Schiff GD, Klass D, Peterson J, Shah G, Bates DW.** Linking laboratory and pharmacy: opportunities for reducing errors and improving care. *Arch Intern Med* 2003;163(8):893–900.
51. **Graber ML, Rusz D, Jones ML,** et al. The new diagnostic team. *Diagnosis (Berl)*. 2017;4(4):225–238
52. **Verna R, Velazquez AB, Laposata M.** Reducing diagnostic errors worldwide through diagnostic management teams. *Ann Lab Med* 2019;39(2):121.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.