Cohort Study

Laboratory-Generated Urine Toxicology Interpretations: A Mixed Methods Study

Isaac S. Chua, MD[1-3], Jaime R. Ransohoff, BA[1,4], Olga Ehrlich, PhD, RN[5], Ethan Katznelson, BA[3], Zain M. Virk, BS[3], Christiana A. Demetriou, PhD[6,7], Athena K. Petrides, PhD[3,4], Endel J. Orav, PhD[1,3], Gordon D. Schiff, MD[1,3], and Stacy E.F. Melanson, MD, PhD[3,4]

From: [1]Division of General Internal Medicine and Primary Care, Department of Medicine, Brigham and Women's Hospital, Boston, MA, USA; [2]Department of Psychosocial Oncology and Palliative Care, Dana-Farber Cancer Institute, Boston, MA, USA; [3]Harvard Medical School, Boston, MA, USA; [4]Department of Pathology, Brigham and Women's Hospital, Boston, MA, USA; [5]Phyllis Cantor Center for Research in Nursing and Patient Care Services, Dana Farber Cancer Institute, Boston, MA, USA; [6]Department of Primary Care and Population Health, University of Nicosia Medical School, Nicosia, Cyprus; [7]The Cyprus School of Molecular Medicine, The Cyprus Institute of Neurology and Genetics, Nicosia, Cyprus

Address Correspondence:
Stacy E.F. Melanson, MD, PhD
Brigham and Women's Hospital
75 Francis Street, Amory 2
Boston, MA 02115, USA
E-mail:
semelanson@bwh.harvard.edu

Disclaimer: There was no external funding in the preparation of this manuscript.

Conflict of interest: Each author certifies that he or she, or a member of his or her immediate family, has no commercial association (i.e., consultancies, stock ownership, equity interest, patent/licensing arrangements, etc.) that might pose a conflict of interest in connection with the submitted manuscript.

Manuscript received: 05-19-2020
Accepted for publication: 09-01-2020

Free full manuscript:
www.painphysicianjournal.com

Background: Clinicians frequently order urine drug testing (UDT) for patients on chronic opioid therapy (COT), yet often have difficulty interpreting test results accurately.

Objectives: To evaluate the implementation and effectiveness of a laboratory-generated urine toxicology interpretation service for clinicians prescribing COT.

Study Design: Type II hybrid–convergent mixed methods design (implementation) and pre–post prospective cohort study with matched controls (effectiveness).

Setting: Four ambulatory sites (2 primary care, 1 pain management, 1 palliative care) within 2 US academic medical institutions.

Methods: Interpretative reports were generated by the clinical chemistry laboratory and were provided to UDT ordering providers via inbox message in the electronic health record (EHR). The Partners Institutional Review Board approved this study.

Participants were primary care, pain management, and palliative care clinicians who ordered liquid chromatography-mass spectrometry UDT for COT patients in clinic. Intervention was a laboratory-generated interpretation service that provided an individualized interpretive report of UDT results based on the patient's prescribed medications and toxicology metabolites for clinicians who received the intervention (n = 8) versus matched controls (n = 18).

Implementation results included focus group and survey feedback on the interpretation service's usability and its impact on workflow, clinical decision making, clinician-patient relationships, and interdisciplinary teamwork. Effectiveness outcomes included UDT interpretation concordance between the clinician and laboratory, documentation frequency of UDT results interpretation and communication of results to patients, and clinician prescribing behavior at follow-up.

Results: Among the 8 intervention clinicians (median age 58 [IQR 16.5] years; 2 women [25%]) on a Likert scale from 1 ("strongly disagree") to 5 ("strongly agree"), 7 clinicians reported at 6 months postintervention that the interpretation service was easy to use (mean 5 [standard deviation {SD}, 0]); improved results comprehension (mean 5 [SD, 0]); and helped them interpret results more accurately (mean 5 [SD, 0]), quickly (mean 4.67 [SD, 0.52]), and confidently (mean 4.83 [SD, 0.41]). Although there were no statistically significant differences in outcomes between cohorts, clinician-laboratory interpretation concordance trended toward improvement (intervention 22/32 [68.8%] to 29/33 [87.9%] vs. control 21/25 [84%] to 23/30 [76.7%], $P = 0.07$) among cases with documented interpretations.

Limitations: This study has a low sample size and was conducted at 2 large academic medical institutions and may not be generalizable to community settings.

Conclusions: Interpretations were well received by clinicians but did not significantly improve laboratory-clinician interpretation concordance, interpretation documentation frequency, or opioid-prescribing behavior.

Key words: Compliance monitoring, chronic pain, urine drug testing, opioid, liquid chromatography-tandem mass spectrometry, palliative care, primary care, substance use disorder, diagnostic error, clinical decision support

Pain Physician 2021: 24:E191-E201

Opioids were responsible for the majority of the 70,200 drug overdose deaths in 2017 (1). For patients on chronic opioid therapy (COT), urine drug testing (UDT) is considered an important component of universal precautions to assess risk of misuse and for detecting aberrant behavior (2-11). The Centers for Disease Control and Prevention (CDC) recommends ordering UDT before starting opioid therapy and then at least annually for ongoing prescribing (12). However, several studies have demonstrated that clinicians frequently misinterpret UDT results (13-19). Consequently, a major hurdle to utilizing UDT effectively is the lack of clinician competence when ordering and interpreting these tests (20).

The 2 most common methodologies for UDT are immunoassay and liquid chromatography-tandem mass spectrometry (LC-MS/MS). Although immunoassay is often used as the first-line screening methodology, some recommend replacing immunoassay with more definitive testing by LC-MS/MS in the pain management population because of its greater sensitivity and specificity (11,21-26). However, interpreting LC-MS/MS results requires a sophisticated understanding of drug metabolic pathways because results are often reported as a complex array of positive and negative metabolites. A recent study showed that clinicians rarely documented interpretations of LC-MS/MS UDT, and 28% of documented interpretations were discordant with expert toxicology interpretations (19).

Efforts to improve clinician UDT interpretation accuracy have focused on clinician education or providing a pharmacist-based interpretation service to the primary team (27-29). One educational intervention improved residents' knowledge and comfort with UDT (27). A pharmacist-run UDT electronic consultation service recommended immediate action in 50% of cases when unexpected results were identified (28). Informal consultations with the laboratory are encouraged (20), but to our knowledge, the effectiveness of formal laboratory-generated urine toxicology interpretation has not been studied. The delivery of laboratory-generated UDT interpretations to ordering clinicians might overcome shortcomings of prior methods (i.e., relying on individual clinician effort to close the knowledge gap on UDT interpretations and lacking standardized, expert guidance from toxicologists).

The aim of this study was to evaluate the implementation and effectiveness of an expert laboratory-generated LC-MS/MS UDT interpretation service for clinicians managing patients on COT. We hypothesized that clinicians would find the laboratory-generated UDT interpretations useful and that the interpretations would improve clinician-laboratory interpretation concordance, documentation of interpretations and results communication with patients, and modify prescribing patterns in the setting of aberrant toxicology results.

Methods

Study Design and Participants

The study design was a type II hybrid that combined a convergent mixed methods design (implementation evaluation) with a pre–post prospective cohort study with a matched control group (effectiveness evaluation). The study period lasted from May 2018 to May 2019, which was divided into the pretest period (May 2018 to October 2018) and posttest period (November 2018 to May 2019). Partners Institutional Review Board approved this study.

Participating clinics included 3 ambulatory sites (primary care, community clinic, pain management) at one academic, tertiary care hospital and 1 palliative care clinic at a comprehensive cancer center in Boston, Massachusetts. The community clinic is a primary care clinic with a higher prevalence of patients with substance use disorders (SUD). Any clinician practicing at one of the participating clinics who ordered a UDT by LC-MS/MS for patients on COT during the study period was eligible for inclusion in the study. When selecting the intervention cohort, we performed a convenience sampling of the 2 highest UDT ordering clinicians at each site. After signing informed consent, intervention group clinicians received laboratory-generated interpretations for LC-MS/MS UDT ordered during the postintervention period. The interpretation was sent to the intervention clinician via an in-basket message several days after the toxicology results had finalized. Results were characterized as either "aberrant" or "nonaberrant" by comparing results to the medications prescribed. See Appendix 1 for definitions of aberrant and nonaberrant results. We selected the control cohort by retrospectively identifying clinicians who ordered LC-MS/MS UDT during the study period. The control cohort were different in the pre- and posttest periods but were closely matched to the intervention clinicians according to practice location, education, age, and gender. See Appendix 2 for description of the matching algorithm.

Laboratory Information

The tertiary care hospital performs approximately 3,500 UDT panels by LC-MS/MS annually. The panel detects opioids, benzodiazepines, and stimulants but

not tetrahydrocannabinol (THC). See Appendix 3 for a complete list of drugs and metabolites included in the panel. Once the UDT results were available, 1 of 2 toxicology experts (S.M. and A.P.) interpreted the results in conjunction with prescribed medications—as documented in the electronic medical record (EMR)—at the time of urine sample collection.

Implementation Evaluation

The intervention group participants completed an adapted version of the Technology Acceptance Model Questionnaire (TAM) (30,31) before the intervention was implemented and at 3 and 6 months postintervention. The TAM has been used in similar studies and is validated for use in evaluating health care information and communication technology (32). The adapted TAM was used to evaluate usefulness and ease-of-use of the laboratory-generated interpretations. Clinicians rated each survey item on a Likert scale from 1 ("strongly disagree") to 5 ("strongly agree").

After the posttest period ended, 2 authors (I.C. and O.E.) co-conducted focus groups from May to June 2019 with intervention group clinicians at each site using a semistructured interview guide. O.E. is a palliative care nurse and an experienced qualitative researcher; I.C. is a palliative care clinician who routinely orders UDT. The questions explored clinician views regarding the intervention's usefulness, barriers to implementation, impact on workflow, clinical decision making, clinician-patient relationships, and interdisciplinary teamwork (see Appendix 4).

Qualitative Analysis

All focus group interviews were audio recorded, transcribed verbatim, and subjected to deductive and inductive thematic analysis. Two authors (I.C. and O.E.) coded the interviews. Initially, each author coded transcripts separately according to a priori themes derived from interview guide questions (i.e., helpfulness, barriers, communication, workflow, patient management and communication) but also allowed themes to emerge from the data. Subsequently, they jointly reviewed the coded transcripts to reconcile codes and themes until both were in full agreement.

Effectiveness Outcome Measures and Data Collection

The primary outcome measure was concordance between clinician and laboratory interpretation of UDT results. Secondary outcome measures included documentation of UDT results interpretation, documentation of UDT results communication to the patient, and clinician prescribing behavior at the subsequent follow-up visit. To assess these outcomes, 2 reviewers (Z.V. and E.K.) retrospectively reviewed patient charts among eligible clinicians during the study period using a standardized instrument. Both reviewers were blinded to clinicians who received the intervention. Each reviewer reexamined 10% of each other's cases to assess interrater reliability. We performed stratified random sampling of patient charts by clinic location to reflect differences in UDT ordering volume per location. Charts were excluded if there were repeat patients, patients not on COT, or circumstances that precluded the ability to evaluate the intervention's effectiveness (e.g., unclear medication information, no documented follow-up, clinician who saw the patient in follow-up or who prescribed the opioid refill differed from the UDT ordering clinician).

Statistical Methods

We investigated associations between categorical variables using the Pearson χ^2 test and the Fisher exact test when cell counts were < 5. For 2 group comparisons, we used Wilcoxon rank-sum for nonnormally distributed continuous variables. A Cohen kappa statistical test was performed to assess agreement between 2 reviewers. A logistic regression model using generalized estimating equations was used to assess differences in change pre–post between the intervention and control groups. A minimum of 50 patient charts was needed in each group (i.e., pretest control, pretest intervention, posttest control, and posttest intervention) in the retrospective chart review to detect 20% change with 80% power in documented clinician-laboratory interpretation concordance, interpretation and communication documentation frequency, and clinician prescribing behavior. All analyses were performed using SAS version 9.4 (SAS Institute, Cary, NC).

Results

Clinician Characteristics

Overall, 8 clinicians participated in the intervention group, and 18 clinicians were included in the control group. The clinicians in the intervention group were older than those in the control group [57.1 vs. 46.3 years, $P = 0.02$]. Otherwise, there was no significant difference in demographics or clinician characteristics between the 2 groups. In both groups, the majority of participants

were men, physicians, completed residency in internal medicine, and completed fellowship training (Table 1).

Chart Review Characteristics

We reviewed 377 charts and excluded 176 for the following reasons: patients were not on COT (n = 89), no follow-up (n = 27), follow-up with a different clinician (n = 37), opioid refill by a different clinician (n = 4), repeat patients (n = 17), unclear medication information (n = 17), and greater than one UDT, ordered making it unclear which result the clinician interpreted (n = 1). Except for race and SUD history, patient characteristics and UDT results did not differ pre–post within each cohort (Table 2). In total, 3 cases contained a laboratory interpretation error: (1) laboratory was unaware of the patient's intrathecal pump, (2) laboratory was unaware that the patient self-discontinued opioids as a result of side effects, and (3) typographical error listed the incorrect opioid in the laboratory interpretation.

Table 1. *Clinician characteristics*.

	Intervention (n = 8)	Control (n = 18)	*P* Value
Age, median (IQR)	58 (16.5)	44 (16)	0.02
Female, no. (%)	2 (25)	4 (22.2)	1.0
Degree, no. (%)			1.0
MD	7 (87.5)	16 (88.9)	
NP	1 (12.5)	2 (11.1)	
Residency training, no. (%)			0.95
Anesthesia	2 (25)	6 (33.3)	
Internal medicine	5 (62.5)	7 (38.9)	
Internal medicine and anesthesia	0 (0)	1 (5.6)	
Emergency medicine	0 (0)	1 (5.6)	
Psychiatry	0 (0)	1 (5.6)	
N/A[a]	1 (12.5)	2 (11.1)	
Fellowship training, no. (%)			0.64
None	3 (37.5)	5 (27.8)	
Pain medicine	1 (12.5)	7 (38.9)	
Hospice and palliative medicine	1 (12.5)	0 (0)	
Pain medicine and hospice	1 (12.5)	1 (5.6)	
Rheumatology	1 (12.5)	1 (5.6)	
Hematology/ oncology	0 (0)	1 (5.6)	
Pain medicine and addiction	0 (0)	1 (5.6)	
N/A[a]	1 (12.5)	2 (11.1)	
Provider location, no. (%)			0.62
Pain management	2 (25)	9 (50)	
Community clinic	2 (25)	3 (16.7)	
Primary care	2 (25)	2 (11.1)	
Palliative care	2 (25)	4 (22.2)	

[a]N/A applies to NP providers because they do not complete a clinical residency or fellowship.

Reviewer Agreement

Reviewers achieved substantial agreement on clinician interpretation concordance with the laboratory (K = 0.92), presence of a documented interpretation (K = 0.78), and determining follow-up visit date (K = 0.79). Reviewers achieved moderate agreement on determining clinician action at follow-up appointments (K = 0.46), and whether or not clinicians communicated UDT results with patients (K = 0.50).

Interpretation Concordance, Documentation, Communication

There were no significant differences pre–post between intervention and control groups regarding clinician-laboratory interpretation concordance, frequency of documentation interpretation, or results communication (Table 3). However, when cases of no documentation are excluded, clinician-laboratory concordance trends toward improvement in the intervention group (intervention 22/32 [68.8%] to 29/33 [87.9%] vs. control 21/25 [84%] to 23/30 [76.7%], P = 0.07). Further adjusting for clinician age did not significantly affect clinician-laboratory concordance (P = 0.08). When toxicology results were aberrant, there were no significant differences in prescribing opioid refills at follow-up between both cohorts (intervention 4/14 [28.6%] to 2/8 [25%] vs. control 4/13 [30.8%] to 7/15 [46.7%], P = 0.56].

Survey Data

Among the 8 intervention clinicians, 7 completed the survey at baseline, 3 months, and 6 months. At baseline, respondents had a mean score of 3.3 (standard deviation [SD], 0.76) for "easy to understand urine toxicology results," 2.7 (SD, 0.76) for "interpret toxicology results quickly," 3.1 (SD, 1.07) for "confident in their interpretation of urine toxicology results," and 4 (SD, 0.58) for "interpret the urine toxicology results accurately." At 3 and 6 months,

Table 2. *Patient characteristics and toxicology results.*

	Intervention			Control		
	Pretest (n = 51)	**Posttest (n = 50)**	***P* Value**	**Pretest (n = 50)**	**Posttest (n = 50)**	***P* Value**
Patient characteristics						
Age, median (IQR)	56 (14)	57.5 (13.8)	0.49	56.5 (18.5)	53.5 (20.8)	0.33
Female, no. (%)	33 (65)	28 (56)	0.37	28 (56)	27 (54)	0.84
Race, no. (%)						
White	38 (75)	40 (80)	0.002	37 (74)	42 (84)	< 0.001
Black	8 (16)	4 (8)		6 (12)	0 (0)	
Hispanic	2 (4)	0 (0)		3 (6)	4 (8)	
Other	2 (4)	4 (8)		4 (8)	3 (6)	
Missing data	0 (0)	2 (4)		0 (0)	1 (2)	
Pain type, no. (%)						
Non-cancer-related pain	48 (94.1)	49 (98)	0.62	48 (96)	49 (98)	1.00
Cancer-related pain	3 (5.9)	1 (2)		2 (4)	1 (2)	
Psychiatric history, no. (%)						
SUD	12 (23.5)	13 (26)	0.77	7 (14)	17 (34)	0.02
Alcohol use disorder	1 (2)	0 (0)	1.00	2 (4)	2 (4)	1.00
Schizophrenia	0 (0)	0 (0)	---	0 (0)	1 (2)	1.00
Bipolar disorder	1 (2)	0 (0)	1.00	3 (6)	1 (2)	0.62
ADHD	0 (0)	1 (2)	0.50	1 (2)	1 (2)	1.00
Depression	6 (11.8)	11 (22)	0.17	13 (26)	14 (28)	0.82
Anxiety	11 (21.6)	8 (16)	0.47	11 (22)	14 (28)	0.49
Clinic location, no. (%)						
Pain clinic	41 (80.4)	40 (80)	1.00	41 (82)	39 (78)	0.83
Community clinic	6 (11.8)	6 (12)		5 (10)	7 (14)	
Primary care	2 (3.9)	2 (4)		1 (2)	2 (4)	
Palliative care	2 (3.9)	2 (4)		3 (6)	2 (4)	
House staff documentation, no. (%)	25 (49)	20 (40)	0.36	22 (44)	19 (38)[a]	0.66
Toxicology results, no. (%)						
Nonaberrant	37 (72.6)	42 (84)	0.16	37 (74)	35 (70)	0.66
Aberrant[b]	14 (27.6)	8 (16)		13 (26)	15 (30)	
Illicit use	4 (7.8)	3 (6)	1.00	4 (8)	6 (12)	0.74
Not taking prescribed medication	6 (11.8)	1 (2)	0.11	2 (4)	4 (8)	0.68
Simulated compliance	3 (5.9)	0 (0)	0.24	2 (4)	1 (2)	1.00
Taking medication not prescribed	6 (11.8)	5 (10)	0.78	6 (12)	7 (14)	0.77
Laboratory interpretation error	1 (2)	1 (2)	0.20	0 (0)	1 (2)	0.36

[a]Missing data (2 entries).
[b]Subcategories of aberrant will not add up to 100% because more than one subcategory can be present in each aberrant sample.
Abbreviations: ADHD, attention deficit hyperactivity disorder; IQR, interquartile range.

respondents "agreed" to "strongly agreed" that the urine toxicology interpretation service made it easier to understand results; allowed them to interpret results more accurately, quickly, and confidently; was useful when managing patients who were prescribed a controlled substance; was easy to learn how to use; and the interpretations were clear and understandable (Fig. 1).

Table 3. *Clinician-laboratory interpretation concordance, results interpretation documentation, and results communication documentation.*

	Intervention Group		Control Group		*P* Value[a]
	Pretest (n = 51)	Posttest (n = 50)	Pretest (n = 50)	Posttest (n = 50)	
Physician interpretation agrees with laboratory interpretation, no. (%) [b,c]	22 (44)	29 (59.2)	21 (42)	23 (47)	0.55
Physician interpretation is present, no. (%)	33 (64.7)	34 (68)	26 (52)	31 (62)	0.71
Discussion of results with patient is documented, no. (%)	8 (15.7)	6 (12)	4 (8)	4 (8)	0.77

[a]Generalized estimating equations logistic regression.
[b]Three patients were excluded because the laboratory interpretation was erroneous and therefore could not be compared to the provider interpretation.
[c]Success is counted as agreement. Failure is counted as either disagreement or no physician documentation.

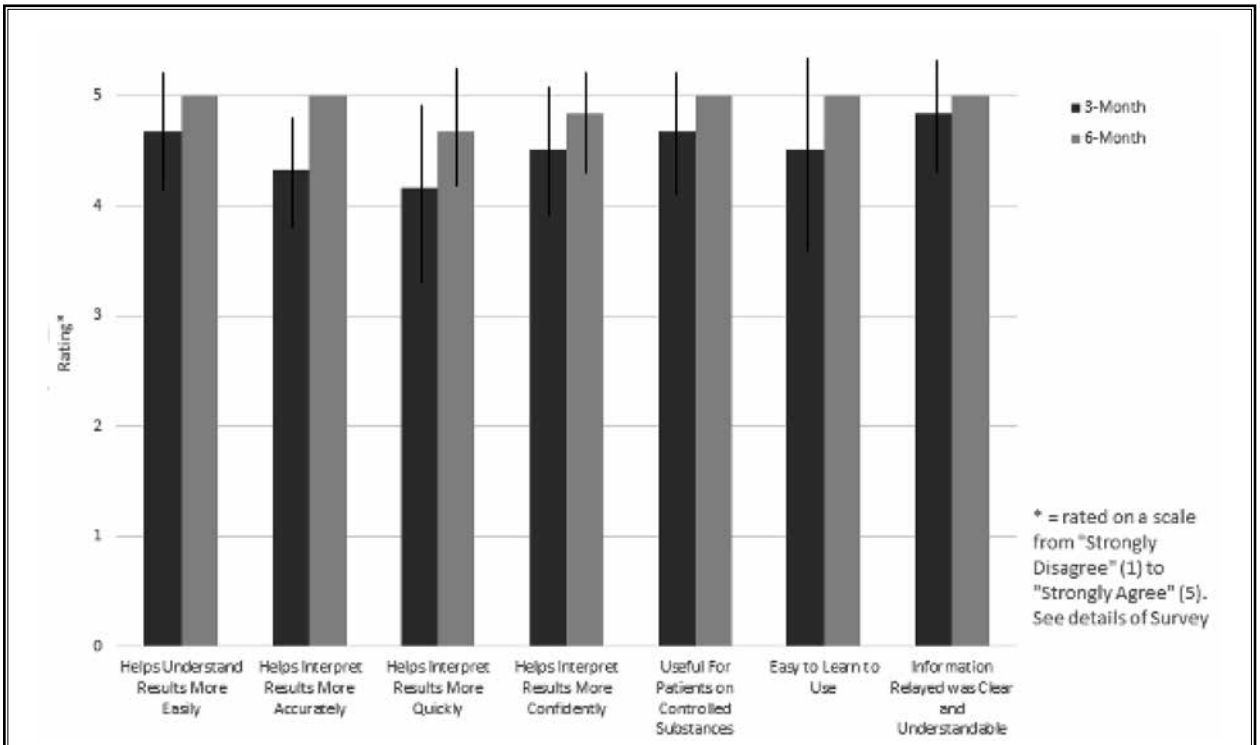

Fig. 1. *Perceived usefulness and ease of use of urine toxicology interpretation service. Intervention group clinicians "agreed" to "strongly agreed" at 3 and 6 months posttest that the interpretation service helped them understand results more easily; interpret results more accurately, quickly, and confidently; was easy to learn to use; was useful for managing patients on controlled substances; and relayed information clearly and understandably. Clinicians' perceptions of the usefulness and ease of use of the interpretation service increased with time.*

Focus Groups

Five major themes emerged from the focus groups. Supporting quotations are cited in Table 4. Themes included (1) layout and language of interpretative reports; (2) utility in aiding clinical decision-making and overcoming knowledge deficits; (3) impact on clinician-patient interactions; (4) human factors and workflow considerations; and (5) effects of external factors on interpretive report utility.

Regarding layout and language of the reports, clinicians stressed the importance of balancing descriptive accuracy while using nonjudgmental language, clearly highlighting normal from abnormal results, and

Table 4. *Description of key themes and subthemes of laboratory-generated interpretations.*

Themes	Supporting Quotation(s)
Layout and language of interpretive reports	
Balancing descriptive accuracy with nonstigmatizing language	"Aberrancy…it's a stronger word…and I think you gotta remember that patients might be…looking at their own records and eventually somehow this would be populating the records, right? So, I think consistent/inconsistent might be a softer way of kind of approaching that..." [Pain Medicine] "I think data and language that sounds like data, sounds like, sort of, a nonjudgmental presentation of fact can be helpful sometimes because…at the end of the day, [an aberrant result] is not what any of us wish. It's what's happening." [Primary Care] "We don't really wanna leave language in the chart to say that they're guilty." [Primary Care]
Highlighting normal vs. abnormal results	"If there's something wrong with the [urine] creatinine, highlight that…cause honestly I'm not gonna notice that." [Primary Care] "The most abnormal should go on top" [Primary Care] "Highlight or bold something that's aberrant or if it's not aberrant." [Pain Medicine]
Clinician desire for greater certainty in the interpretations	"I never know whether there's any value in the numeric quantity when it's found…basically you'd say 'positive' or 'negative'…and I don't know whether a number ought to be conveying more information to me." [Primary Care] "I'm not sure how they would word this, but the percentage of confidence that it's legitimate…some statement that says, 'with high degree of certainty, this is a positive finding,' or somehow to let us know that…the risk of false positive is extremely low, if not nil." [Palliative Care] "In the medication notes…when you say 'active,' I'm not sure what the time frame is." [Primary Care]
Utility in aiding clinical decision-making and overcoming knowledge deficits	
Concise summary of findings	"I find that they're a helpful, concise, one-line summary of the findings." [Palliative Care] "…it was a…informative and quick scan." [Pain Medicine] "They were pretty clear…pretty easy to read." [Primary Care]
Accessible reference that compensates for knowledge gaps	"…taking the sum of the PCPs, where you know, the education is all over and just having an interpretation [is] a huge plus." [Primary Care] "…the labs are somewhat complicated…most of the time you end up pulling up the PowerPoint or trying to figure it out myself with lots of…mixed messages and unclear data, which made [the interpretation] uncertain. So it just made that part a lot easier." [Primary Care]
Prevents misinterpretation by flagging and reinforcing aberrant results	"I think the greatest benefit…it's just flagging it in my mind. I'm like, 'What the hell? That was a positive? I can't believe it!'…it will stay with me because those I will remember." [Pain Medicine] "We miss stuff sometimes. I found it very good that…I was getting some…reinforcing things to look at this in person." [Pain Medicine] "It just stops me a little bit in my tracks and makes me look at things." [Primary Care]
Helpful for managing high-risk populations	"It's more the newer Suboxone patients that are…toying with their sobriety that…it does make a difference to have an interpretation." [Primary Care] "Dr. 1: …four years ago, this would have been the greatest gift to humankind. Dr. 2: Right…when we were actually weaning the people off [of opioids] who were abusing." [Primary Care]
Impact on clinician-patient interactions	
Helps clinicians prepare for future patient encounters	"It was a good heads up for when you see the patient again…I get the e-mail, then I either remember it or make a notation." [Pain Medicine] "It wakes me up to like, I gotta like…we either got to get him in sooner…" [Pain Medicine] "They potentially affect the decision-making at the next cycle of prescribing." [Primary Care]
Impact on clinician communication when discussing results	"It gives me more confidence having those conversations, um, and more like, I have a…scientific leg to stand on than I think I might have felt before." [Primary Care] "It celebrated the communication with them somehow, but it didn't change anything." [Pain Medicine] "There's just sort of [an] intangible…comfort and confidence of…known that…in the back of your mind you didn't just mess something up or…that you didn't just miss something…" [Primary Care]
Interplay of human factors, systems, and workflow considerations	
Desire for greater visibility in the EMR	"If this becomes standard, I'd want to roll out. It's beyond primary care. It's anybody who would look at a tox screen." [Primary Care] "…it might be nice to have the interpretation…tied to the result to the benefit of someone else to also have the interpretation." [Primary Care]
Delayed timing of receiving interpretations	"The lag sometimes was significant that I had to make a clinical decision before the interpretation came through. I was comfortable doing that but someone else might not have the expertise I did." [Pain Medicine] "The only problem would be that [the interpretations] took a lot of time to turn up." [Primary Care] "I just wanted to…either get rid of the [toxicology result] or to respond, or do the action, so the wait was a slight bother…" [Primary Care]

Table 4 (con't). *Description of key themes and subthemes of laboratory-generated interpretations.*

Themes	Supporting Quotation(s)
Layout and language of interpretive reports	
Minimal time saving	"That's the only time saver is if I were to have to spend a lot of time trying to interpret something or track somebody down [to help interpret]." [Primary Care] "On rare occasions, if I would have had to make a phone call to say, 'What does this mean?'…I haven't had to do that since it started, and I maybe had done that—I don't know—4 or 5 times a year." [Primary Care] "I don't think it's so much time saving because when we see the patient, we still have…the fellow, who we usually see people with, or the resident usually does that, so personally it might not be relevant because we're still going to have to look it up." [Pain Medicine]
Effects of external factors on interpretive report utility	
Incomplete or inaccurate clinical information	"Sometimes the medication lists aren't accurate…a few of the times…you guys commented on benzos—there was no evidence—but…you don't reference the PMP or anything right?...it's a big deal." [Pain Medicine] "It's like with radiology, 'concerning for cancer' but I don't have a biopsy, so I can't say it's cancer. You guys can't be 100% sure…you'd need all the other background info." [Primary Care] "Patients in our practice certainly are seen elsewhere…the [EMR] is not the comprehensive source of that [medication] information." [Palliative Care]
Fentanyl contamination as a confounder	"…it's still not giving us more confidence because of this issue with the dusting of marijuana with everything… and it's kind of interfering with the confidence that we can address things with them." [Pain Medicine] "…fentanyl began to appear everywhere…so your patients might not actually be using the fentanyl but things [are] getting cut with it." [Primary Care]
Imperfect urine drug test and urine collection process	"It's difficult now because we don't have the marijuana [in the urine drug test] that we can just say, 'Did you buy it on the street?' because it's all contaminated now…from fentanyl to cocaine to God knows what." [Pain Medicine] "Your thresholds are so low…we're definitely getting a lot of what I would consider, probably not false positives but false clinical relevancies because of probably some other explanation for what's going on." [Pain Medicine] "There are standards for how the process is collected that we don't have the means to put in place, so it's very easy to provide a tampered specimen. And that's…one big reason why I lost faith…in sending them and trying to…make sense of the interpretations." [Palliative Care]

Abbreviations: PCP, primary care physician, PDMP, prescription drug monitoring program.

having the interpretations express greater certainty. When making clinical decisions, the interpretations served as an accessible reference and concise summary of the findings that is helpful for managing high-risk populations and prevents errors by flagging and reinforcing results. When interacting with patients, the interpretive reports helped clinicians prepare for future encounters, but had variable impact on clinician confidence in discussing results. Clinicians acknowledged the intervention's modest time-saving features but also expressed a desire for wider access of the interpretations within the EMR and improving the timeliness of receiving the interpretations. Ultimately, other factors impacted the utility of the intervention, including the quality and completeness of medication information in the EMR on which the interpretation is based.

Discussion

Clinicians receiving expert laboratory interpretations of UDT results found this service valuable. Intervention clinicians felt these interpretations provided a reliable, concise summary of findings that helped minimize errors, prepared them for patient encounters, and assisted with opioid management of high-risk patients. They felt that the interpretation service was easy to use, enhanced comprehension of urine toxicology results, and helped them interpret results more accurately, quickly, and confidently. However, although there was a trend toward improvement in our primary outcome, the intervention failed to show a statistically significant difference between the intervention and control groups regarding clinician-laboratory interpretation concordance, interpretation documentation, or opioid prescribing behavior during follow-up. Our study identified areas for improvement in the intervention and how external factors affect the intervention's utility.

Our qualitative data elucidated how and why clinicians strongly agreed that the laboratory-generated interpretations were useful, easy to use, and improved their interpretation abilities. First, an expert toxicologist summarized the complex results using clear and concise language and conveniently delivered the interpretation to their in-basket. Previously, clinicians who were unsure about the meaning of test results would have contacted the laboratory by e-mail, telephone,

or pager, or painstakingly looked up the drug metabolic pathway themselves. Second, the interpretation brought aberrant results to the clinician's attention, reducing the likelihood that they would overlook them, which could easily occur in cases with subtle abnormalities (e.g., urine dilution). Third, the clinicians felt the interpretations provided reassurance from a trusted source, thereby helping them engage more confidently in difficult conversations with patients to discuss aberrant results.

We found no significant difference in clinician-laboratory interpretation concordance or documentation between intervention and control groups. Although we calculated an a priori sample size to detect a clinically significant difference, low rates of documentation could have underpowered our study. Also, importing text from prior notes is common (33), and sending clinicians personalized interpretations may not have been a strong enough behavioral nudge to change clinician documentation habits (34,35). Therefore finding ways to incorporate written toxicology interpretations into clinician documentation workflow (e.g., auto-population of interpretations into the note) may help improve interpretation documentation in the future.

The laboratory toxicologists did not have direct access to the Massachusetts Prescription Drug Monitoring Program (PDMP) because review of this database is not permitted for research. As a result, they relied on prescriptions in the EMR without verification of which medications were actually filled and/or being taken at the time of UDT. This may have led to an overestimation of the aberrant subcategory "taking medication not prescribed" because a controlled substance could have been prescribed by an outside clinician. Therefore if the laboratory is to be included as part of the diagnostic team (36-38), ability to review the PDMP should be permitted to increase their interpretation accuracy (39,40).

Clinicians also cited concerns about fentanyl lacing, highly sensitive UDT cutoffs, and absence of THC confounding their ability to interpret the test accurately. Fentanyl-laced cocaine, heroin, methamphetamine, and counterfeit pills have been increasingly reported (41), but there is little evidence to support fentanyl-laced marijuana (42). The lack of universal standards for determining cutoffs in urine toxicology results leads to significant results variability and can complicate interpretations (43-45). Furthermore, consensus among toxicologists and prescriber input are needed to standardize panel components (e.g., whether to include THC) to provide optimal and accurate interpretations for frontline clinicians (46).

Future laboratory-generated interpretations ideally should be included as an impression section on urine toxicology results similar to radiology and pathology reports. This logical and convenient location of the interpretation would support busy clinicians quickly viewing and understanding the results. However, issues related to EMR privacy for patients with SUD receiving alcohol or drug abuse treatment in federally assisted treatment programs has led to the 42 CFR (Code of Federal Regulations) Part 2, a law that keeps such records confidential (47,48). To the extent that the 42 CFR Part 2 applies to these laboratory-generated interpretations, it may preclude the implementation of such interpretations because complete drug use data would not be accessible in the chart. Thus confirmation with a hospital's office of general council may be required before implementing this service widely.

If laboratory-generated toxicology interpretations become more widespread, guidelines to standardize reporting language will be necessary to ensure that information is conveyed accurately and succinctly while minimizing patient stigma. Substantial variation in interclinician reporting currently exists within radiology and pathology (49,50). These specialties are working to reduce unwarranted variation by synoptic reporting, which is the use of a structured report with coded concepts that support the discrete input and storage of clinical data and enables direct extraction in a machine-readable format (51,52). Synoptic laboratory and imaging reporting has been associated with increased physician satisfaction (53) and more complete reporting (54).

This study has several limitations. First, the small sample size may have led to an inability to detect statistically significant change between the intervention and control groups. Second, clinician interpretation documentation does not directly measure clinician knowledge or comprehension of toxicology results. Third, the intervention was implemented in ambulatory clinics of 2 large academic institutions, which may limit generalizability in the community setting. Fourth, a convenience sampling of the highest UDT ordering clinicians meant to optimize ordering volume in the intervention group may not be representative of clinicians who order UDT less frequently. However, the baseline scores of these experienced clinicians in the intervention group showed suboptimal scores in regard to speed, results comprehension, and self-confidence when interpreting UDT results. Fifth, the clinician char-

acteristics between both groups were well matched except for age. Because the intervention cohort was older, the accumulated experience of interpreting toxicology results because of more years of experience may have conservatively biased the results toward the null. Finally, toxicologists based their interpretations on the results and medications listed in the EMR and did not have access to the PDMP. Such access would have provided a more comprehensive account of prescribed and dispensed controlled substances.

Conclusions

Clinicians who received these laboratory-generated UDT interpretations found that the service was useful, easy to use, enhanced comprehension of toxicology results, and helped them interpret results more accurately, quickly, and confidently. However, the intervention did not improve laboratory-clinician interpretation concordance, results interpretation documentation, or change documented opioid-prescribing behavior in a statistically significant manner. Larger prospective studies are needed to assess the efficacy of laboratory-generated interpretations on improving clinician knowledge, skills, and attitudes when interpreting UDT results.

Author Contributions

Drs. Chua, Orav, Katznelson, Virk, Ms. Ransohoff, and Dr. Demetriou had access to the data in the study and take responsibility for the integrity of the data and the accuracy of the data analyses. Drs. Melanson, Petrides, Chua, Schiff, Ms. Ransohoff, and Dr. Ehlrich designed the study protocol. Dr. Chua and Ms. Ransohoff managed the literature searches and summaries of previous related work and wrote the first draft of the manuscript. Drs. Chua, Melanson, and Petrides provided revision for intellectual content and final approval of the manuscript.

References

1. National Institute on Drug Abuse. Overdose death rates. Available at: www.drugabuse.gov/related-topics/trends-statistics/overdose-death-rates. Updated Sept 2017. Accessed 03/03/2021.
2. Manchikanti L, Atluri S, Trescot AM, Giordano J. Monitoring opioid adherence in chronic pain patients: Tools, techniques, and utility. *Pain Physician* 2008; 11(2 Suppl):S155-S180.
3. 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; 9:123.
4. Trescot 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.
5. Melanson SE, 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; 47:971-976.
6. 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; 23:173-179.
7. 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 Inter Med* 2010; 152:712-720.
8. Katz NP, Sherburne S, Beach M, et al. Behavioral monitoring and urine toxicology testing in patients receiving long-term opioid therapy. *Anesth Analg* 2003; 97:1097-1102.
9. Heit HA, Gourlay DL. Urine drug testing in pain medicine. J Pain Symptom Manage 2004; 27:260-267.
10. 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; 29:1663-1671.
11. Argoff CE, Alford DP, Fudin J, et al. Rational urine drug monitoring in patients receiving opioids for chronic pain: Consensus recommendations. *Pain Med* 2017; 19:97-117.
12. Dowell D, Haegerich TM, Chou R. CDC Guideline for Prescribing Opioids for Chronic Pain--United States, 2016. *JAMA* 2016; 315:1624-1645.
13. Starrels JL, Fox AD, Kunins HV, 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; 27:1521-1527.
14. Reisfield G, Bertholf R, Barkin R, Webb F, Wilson G. Urine drug test interpretation: What do physicians know? *J Opioid Manag* 2007; 3:80-86.
15. Reisfield GM, Webb FJ, Bertholf RL, Sloan P, Wilson G. Family physicians' proficiency in urine drug test interpretation. *J Opioid Manag* 2007; 3:333-337.
16. Durback L, Scharman E, Brown B. Emergency physicians perceptions of drug screens at their own hospitals. *Vet Hum Toxicol* 1998; 40:234-237.
17. Levy S, Harris SK, Sherritt L, Angulo M, Knight JR. Drug testing of adolescents in ambulatory medicine: Physician practices and knowledge. *Arch Pediatr Adolesc Med* 2006; 160:146-150.
18. 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 Abuse* 2018; 39:518-521.
19. Chua I, Petrides AK, Schiff GD, et al. Provider misinterpretation, documen-

tation, and follow-up of definitive urine drug testing results. *J Gen Intern Med* 2020; 35:283-290.

20. 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; 166:506-513.
21. 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. *J Appl Lab Med* 2018; 2:489-526.
22. Cross TG, Hornshaw MP. Can LC and LC-MS ever replace immunoassays? *J Appl Bioanal* 2016; 2:108.
23. 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* 2010; 13:273-281.
24. Petrides A, Melanson S. LC-MS/MS testing for monitoring compliance in pain management: The cost and technical benefits of a phased approach. *Clin Lab News* 2016; 42:22-23.
25. 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.
26. Darragh A, Synyder 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* 2014; 17:359-366.
27. 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; 14:10684.
28. 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; 12:389-395.
29. Arthur J, Edwards T, Reddy S, et al. Outcomes of a specialized interdisciplinary approach for patients with cancer with aberrant opioid-related behavior. *Oncologist* 2018; 23:263-270.
30. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Q* 1989; 13:319.
31. Chau PY. An empirical assessment of a modified technology acceptance model. *J Manag Inf Syst* 1996; 13:185-204.
32. Rahimi B, Nadri H, Afshar HL, Timpka T. A systematic review of the technology acceptance model in health informatics. *Appl Clin Inform* 2018; 9:604-634.
33. Wang MD, Khanna R, Najafi N. Characterizing the source of text in electronic health record progress notes. *JAMA Intern Med* 2017; 177:1212-1213.
34. Patel MS, Volpp KG, Asch DA. Nudge units to improve the delivery of health care. *N Engl J Med* 2018; 378:214-216.
35. Lorenzetti DL, Quan H, Lucyk K, et al. Strategies for improving physician documentation in the emergency department: A systematic review. *BMC Emerg Med* 2018; 18:36.
36. 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; 163:893-900.
37. Graber ML, Rusz D, Jones ML, et al. The new diagnostic team. *Diagnosis* 2017; 4:225-238.
38. Verna R, Velazquez AB, Laposata M. Reducing diagnostic errors worldwide through diagnostic management teams. *Ann Lab Med* 2019; 39:121-124.
39. Haffajee RL, Jena AB, Weiner SG. Mandatory use of prescription drug monitoring programs. *JAMA* 2015; 313:891-892.
40. Haffajee RL. Prescription drug monitoring programs--friend or folly in addressing the opioid-overdose crisis? *N Engl J Med* 2019; 381:699-701.
41. Drug Enforcement Administration. National drug threat assessment, 2019. Available at: https://dea.gov/sites/default/files/2020-01/2019-NDTA-final-01-14-2020_Low_Web-DIR-007-20_2019.pdf. Accessed March 7, 2020.
42. Kasprak A. *Is fentanyl-laced marijuana use on the rise?* Available at: www.snopes.com/fact-check/fentanyl-laced-marijuana-rise/. Published 2017. Accessed March 7, 2020.
43. Petrides A, Melanson S. Urine drug testing: Debates over best practices to assess compliance and manage the opioid crisis. *Clin Lab News* 2019:8-12.
44. Clarke WA. Are standard cutoff concentrations sufficient for pain management screening? Yes! *J Appl Lab Med* 2018; 2:653-656.
45. McMillin GA. Drug detection in urine for evaluating exposure—No limits! *J Appl Lab Med* 2018; 2:648-652.
46. Melanson SE, Ptolemy AS, Wasan AD. Optimizing urine drug testing for monitoring medication compliance in pain management. *Pain Med* 2013; 14:1813-1820.
47. Electronic Code of Federal Regulations. *Part 2: Confidentiality of substance use disorder patient records.* Available at: www.ecfr.gov/cgi-bin/text-idx?c=ecfr&sid=b7e8d29be4a2b815c404988e29c06a3e&rgn=div5&view=text&node=42:1.0.1.1.2&idno=42. Updated March 5, 2020. Accessed March 8, 2020.
48. Substance Abuse Mental Health Services Administration. Substance abuse confidentiality regulations. Available at: www.samhsa.gov/about-us/who-we-are/laws-regulations/confidentiality-regulations-faqs. Updated April 18, 2019. Accessed March 8, 2020.
49. Cochon LR, Kapoor N, Carrodeguas E, et al. Variation in follow-up imaging recommendations in radiology reports: Patient, modality, and radiologist predictors. *Radiology* 2019; 291:700-707.
50. Persing S, James TA, Mace J, Goodwin A, Geller B. Variability in the quality of pathology reporting of margin status following breast cancer surgery. *Ann Surg Oncol* 2011; 18:3061.
51. Srigley JR, McGowan T, MacLean A, et al. Standardized synoptic cancer pathology reporting: A population-based approach. *J Surg Oncol* 2009; 99:517-524.
52. Goel AK, DiLella D, Dotsikas G, Hilts M, Kwan D, Paxton L. Unlocking radiology reporting data: An implementation of synoptic radiology reporting in low-dose CT cancer screening. *J Digit Imaging* 2019; 32:1044-1051.
53. Lankshear S, Srigley J, McGowan T, Yurcan M, Sawka C. Standardized synoptic cancer pathology reports—So what and who cares?: A population-based satisfaction survey of 970 pathologists, surgeons, and oncologists. *Arch Pathol Lab Med* 2013; 137:1599-1602.
54. Sluijter CE, van Lonkhuijzen LR, van Slooten H-J, Nagtegaal ID, Overbeek LI. The effects of implementing synoptic pathology reporting in cancer diagnosis: A systematic review. *Virchows Arch* 2016; 468:639-649.

Appendix 1. *Definitions of aberrant and nonaberrant results and their subcategories.*
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 "nonaberrant" 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 nonaberrant.
See the following table for definitions of aberrant and nonaberrant subcategories.

Term	Definition
Illicit drug use	A UDT panel was considered indicative of illicit drug use if any of the following drugs were detected without evidence of a prescription: 6-acetylmorphine (6-AM, the unique heroin metabolite), amphetamine, benzoylecgonine (cocaine metabolite), fentanyl, methamphetamine or morphine. Our pain management panel does not include THC; therefore the presence of THC, indicative of marijuana use, was not considered illicit drug use. An example of an illicit drug interpretation is "Suggestive of cocaine use."
Simulated compliance	A UDT panel was considered indicative of simulated compliance if (a) the urine creatinine concentration was <20 mg/dL, or (b) a high concentration of a drug was detected with absence or very low levels of metabolites. Examples of these types of interpretations are "The presence of abnormally high concentrations of buprenorphine and naloxone suggest simulated compliance (i.e., dropping a Suboxone tablet or film directly into the urine sample)" or "The creatinine concentration is between 5 and 20 mg/dL, indicating a very dilute specimen, which may artificially lower drug concentrations below the detection limit of the assay. This typically indicates the ingestion of large amounts of fluids or the addition of fluids directly to the urine specimen at the time of collection."
Not taking a prescribed drug(s)	The absence of a prescribed drug or its metabolite(s) in the patient's urine was interpreted as "Not taking a prescribed drug." An example of this type of interpretation is "No evidence of recent hydromorphone use despite an existing prescription."
Taking a drug not prescribed	The presence of a drug or its metabolite(s) in the patient's urine, which was not prescribed in the EMR, was interpreted as "Taking a drug not prescribed." An example of this type of interpretation is "Suggestive of oxycodone use, a prescription for which was not found in EMR."
Results consistent with prescriptions	Results were considered consistent with prescriptions if the UDT panel was positive for the prescribed drug(s) and/or its metabolite. An example of this type of interpretation is "Consistent with buprenorphine use, as prescribed."
PRN medications	If the patient was prescribed a medication PRN, both the presence or absence of the drug and/or metabolites was considered nonaberrant. An example of this type of interpretation is "No evidence of oxycodone use, despite an existing PRN prescription."

Abbreviations: PRN, as needed; UDT, urine drug testing; THC, tetrahydrocannabinol; EMR, electronic medical record.

Appendix 2. *Control matching algorithm*.

Intervention Clinician	Number of Pretest Controls	Overall % Match Pretest Controls*	Number of Posttest Controls	Overall % Match Posttest Controls*
A	1	85%	1	70%
B	1	80%	2	83%
C	7	80%	7	80%
D	4	100%	5	99%
E	1	70%	1	90%
F	1	90%	1	90%
G	1	80%	1	80%
H	1	100%	1	100%
Average		86%		87%

Clinicians were matched based on location, degree, age, and gender. Control clinicians were assigned points based on the characteristics that matched with the intervention clinician's. The control clinician was assigned 4 points for location, 3 for degree, 2 or 1 for age, and 1 for gender. If age matched within 5 years, providers were assigned 2 points. If age matched within 10 years, they were assigned 1 point.

For each assigned control, "percentage match" with an intervention clinician was calculated by dividing the number of points assigned by the total number of points possible. For a given intervention clinician, the algorithm identified the highest percentage matched controls in a stepwise manner, until the required number of cases was obtained.

Providers who matched equally well with an intervention provider were assigned a "priority group" and cases were randomly pulled from the highest "priority group." In a couple of cases, a second "priority group" had to be created to obtain the remaining number of cases.

*The "overall percent match" was calculated by adding the "percentage match" of all controls assigned to an intervention clinician then dividing this sum by the total number of control cases.
For example, let's say that intervention clinicians needed to be matched with 20 cases. Priority group 1 had 3 clinicians and 18 cases, and priority group 2 had 5 clinicians and 10 cases. We would then select all 18 cases from priority group 1 then randomly select the remaining cases from priority group 2.

If priority group 1 had a "percentage match" of 100%, and priority group 2 had a "percentage match" of 80%, then the "overall percent match" = [[(1*18) + (0.8*2)]/20]*100 = 98%.

The total number of clinicians used was all 3 from priority group 1 and one clinician happened to be associated with the 2 cases from priority group 2, therefore 4 control clinicians total were matched with the intervention clinician.

Appendix 3. *Drugs and metabolites included in the urine toxicology panel.*

Drug Class	Drug	Cutoff (ng/mL)	Qual/Quant Results
Amphetamines	Amphetamine	25	Qual
	MDA	25	Qual
	MDMA	25	Qual
	Methamphetamine	25	Qual
Benzodiazepines	7-Aminoclonazepam	25	Quant
	Alpha-hydroxy-alprazolam	25	Quant
	Clonazepam	5	Qual
	Diazepam	5	Qual
	Lorazepam	25	Quant
	Nordiazepam	25	Quant
	Oxazepam	25	Quant
	Temazepam	25	Quant
Buprenorphine	Buprenorphine	5	Quant
	Norbuprenorphine	5	Quant
	Buprenorphine-glucuronide	5	Quant
	Norbuprenorphine-glucuronide	5	Quant
	Naloxone	100	Quant
Cocaine metabolite	Benzoylecgonine	5	Qual
Fentanyl	Fentanyl	2	Qual
	Norfentanyl	2	Qual
Methadone	Methadone	5	Qual
	Methadone metabolite (EDDP)	5	Qual
Opiates/opioids	6-Acetylmorphine (heroin metabolite)	5	Qual
	Codeine	25	Quant
	Hydrocodone	25	Quant
	Hydromorphone	25	Quant
	Morphine	25	Quant
	Morphine-3-beta-glucuronide	25	Quant
	Morphine-6-beta-glucuronide	25	Quant
	Noroxycodone	25	Quant
	Oxycodone	25	Quant
	Oxymorphone	25	Quant
Tramadol	O-desmethyltramadol	25	Qual
	Tramadol	5	Qual

Abbreviations: Qual, qualitative; Quant, quantitative; MDA, 3,4-methylenedioxyamphetamine; MDMA, 3,4-methylenedioxymethamphetamine; EDDP, 2-ethylidene-1,5-dimethyl-3,3-diphenylpyrrolidine.

Appendix 4.

1. In which ways were the lab interpretations helpful?
2. What was problematic or not helpful about the lab interpretations?
3. Tell us about any other barriers you encountered while using the lab interpretations?
4. How much time do the interpretations save you?
5. How did the interpretations change your communication with the patient, compared with the previous lab reports?
   a. How did the interpretations change your communication with the patient when the test detected illicit drug use?
6. Tell us how the new interpretations changed your management of patients who require urine toxicology in general?
   a. How did the lab interpretations change your development of a plan of care?
   b. Can you think of any specific examples in which the interpretation changed your management of a patient?
7. We would like to know about communications you had with other health care professionals about the interpretations.
   a.First, we would like to know about your communications with the laboratory. For instance, how did they affect how you communicated with the lab?
   b. Second, how did the lab interpretations affect your communications with other health care providers caring for the patient.
8. Given that it would take an extra day to finalize the lab interpretation, how do you feel about waiting a few days to receive the lab results and interpretation together?
   a. If not, what is the best time frame within which to receive the interpretation?
9. In the interpretations, we try to frame the comments using a scale from "nonaberrant" to "aberrant." For example, "consistent as prescribed" comes before any "illicit use" comments. How does this ordering sound to you?
   a. For example:
      i. "Consistent with prescription as prescribed"
      ii. "No evidence of prescribed prescription"
      iii. "Suggestive of using a prescription medication that is not prescribed"
      iv. "Suggestive of illicit drug use"
      v. "Suggestive of simulated compliance"
10. Which word would be most helpful in describing illicit drug use?
    a. For example, what are your thoughts on using the word "suggestive" of illicit drug use as opposed to "consistent with" illicit drug use?
11. How else would you change the interpretative report? What comments do you have on layout of the interpretation report and how it could be improved?

If time permits display some generic result scenarios and ask the provider to tell us how the interpretation could have changed communication with the patient or changed management.