Performance of Five Non-Instrumented Urine Drug-Testing Devices with Challenging Near-Cutoff Specimens*

Leo J. Kadehjian†
Biomedical Consulting, Palo Alto, California

Abstract

A comparison of five non-instrumented urine drug-testing devices was performed using a challenging clinical specimen set with drug concentrations close to the immunoassay screening cutoffs. The five devices were Syva RapidTest d.a.u. 8, Syva RapidCup d.a.u. 5, Roche TesTcup 5, Biosite Triage, and Casco-Nerl microLINE Drug Screen Card. Sixty clinical specimens for each of the five SAMHSA-specified drug categories were tested by both a scientist and a non-scientist with each result independently read by both. All specimens were also tested on a benchtop automated immunoassay analyzer (Syva ETS using Emit d.a.u. reagents) for comparison and by gas chromatography–mass spectrometry (GC–MS). The non-instrumented devices demonstrated an overall accuracy of 70% (66–74%), based on standard GC–MS confirmation cutoffs, comparable to the Syva ETS analyzer (80%). There was also little difference in overall accuracy between the scientist (71%) and non-scientist (69%), although the non-scientist reported 10 false-positive results (0.7% of 1490 total results or 3.8% of 260 results for drug-free specimens), and the scientist reported only 1 false-positive result (0.07% of 1490 total results or 0.38% of 260 results for drug-free specimens). When device performance was assessed according to drug presence/absence criteria, accuracy generally improved with all devices demonstrating extremely high positive predictive values (0.98–1).

Introduction

There has been much recent interest in the utility of rapid, easy-to-use, non-instrumented urine drug-testing devices for a variety of clinical and non-clinical applications: emergency room (1–3), perinatal (4), drug treatment, regulated and non-regulated workplaces (5,6), children and students (7), a variety of criminal justice settings (8), drugged driving (9,10), and even in the home (11–13). Since their introduction in the 1980s, there have been numerous technological advancements and performance improvements. The current devices utilize well-established immunoassay technologies with antigen-antibody reactions in part occurring on chromatographic test strips and the test results being read visually within a few minutes as the presence or absence of a colored line. The test strips also have control lines so each test is controlled. A wide variety of these devices is now available in both single- and multi-assay configurations and in different formats, such as cup-type devices where the specimen may be provided directly into the device, card- or cassette-type devices where the specimen is applied to the enclosed test strip using a pipette, and dipstick-type devices where the test strip is immersed in an aliquot of the specimen. These devices are generally easy to use and rapid with results generally available within 10 min, and often within 5 min. The great interest in these devices is evidenced by the rapid proliferation of the variety of these devices and the numerous positive performance evaluations of these devices presented at scientific meetings and in scientific publications (14–34).

Furthermore, the utility and application of these and other unit test devices are undergoing active regulatory review by numerous agencies: the U.S. Food and Drug Administration (FDA) establishing regulatory clearance criteria for marketing, with at least one non-instrumented drug-testing device already cleared by the FDA for at-home use; the Substance Abuse and Mental Health Services Administration (SAMHSA) for potential application in federally regulated workplace drug-testing programs; a wide variety of criminal justice agencies; and laboratory accreditation and standards organizations such as the College of American Pathologists (CAP), the Healthcare Financing Administration (HCFA), and the National Committee for Clinical Laboratory Standards (NCCLS).

However, there has been ongoing concern that these simple, visually read devices may not provide sufficient scientific or forensic accuracy for use in some of the mentioned applications. Currently, there appears to be a consensus that these devices' results should not be used without further confirmation testing, regardless of the setting. Although many of the published performance studies have found these devices to perform surprisingly well, some published studies have been critical of their performance (1,35–47). It is interesting to note that although these devices are now being widely used in

* This work was presented at the annual meeting of the Society of Forensic Toxicologists, San Juan, Puerto Rico, October 1999.

† Address correspondence to Dr. Leo Kadehjian, Biomedical Consulting, 765 Chimalus Drive, Palo Alto, CA 94306.

numerous settings, there is little significant case law where there has been detailed judicial scrutiny of the accuracy of these devices and whether use of these devices fulfills the due process requirements in each of the variety of settings.

The concerns of the accuracy of these devices are based mainly on the subjective nature of visually reading the test results as the presence or absence of a colored line. Visual determination of the test results is considered to be the most challenging when specimen drug concentrations are near the immunoassay cutoffs, and accordingly the clear presence or absence of a line may be difficult for the operator to determine. The device package inserts indicate that any such equivocal or borderline results should be reported as negative, taking a conservative approach. There are also concerns about accuracy since the tests may likely be performed by non-technical staff without formal laboratory experience.

In order to address these concerns of accuracy of these non-instrumented drug-testing devices, a comparison study of five current devices was performed to assess their relative performance using a challenging clinical specimen set specifically selected to be weighted around the standard SAMHSA immunoassay screening cutoffs. A few other studies have also specifically addressed device performance with specimens with drug concentrations near the assay cutoffs (5,8,36,37). In addition, the study assessed device performance in the hands of both scientifically trained and non-technical operators.

Materials and Methods

The five devices included in the study were Syva RapidTest d.a.u. 8, Syva RapidCup d.a.u. 5, Roche TesTcup 5, Biosite Triage, and Casco-Nerl microLINE Drug Screen Card. All devices were obtained directly from the respective manufacturers; the Syva RapidTest and Syva RapidCup were provided at no cost, and the other devices were purchased.

Product descriptions

Syva RapidTest d.a.u. 8. This cassette-type device requires the dropwise addition of three drops of urine to each of the cassette's two specimen wells using the provided disposable squeeze-type plastic pipette. The specimen immediately wicks along the two parallel test strips, with each strip testing for four drugs. Each strip also has a single control line. Results are read between 5 and 10 min. Positive results are indicated by the absence of a colored line across the test strip for each assay. This device required about 6 min to perform each test.

Syva RapidCup d.a.u. 5. This combination collection/test cup device has five single-assay test strips built into the wall of the cup. Each strip also has a single control line. A specimen is provided directly into the collection/testing device and the cap is partly secured. The test is not activated until the cap is further screwed down to its full stop position. The cap has a tamper-evident ratcheting mechanism that precludes the cap from later being removed without evidence. The cap presses down a dual O-ring sealed plunger that delivers a small portion of the specimen to the test well in the base of the device, isolating the test aliquot from the remainder of the specimen. The test sample wicks up the test strips, and the results are read through the window in the side of the collection/test cup. Positive results are indicated by the absence of a colored line across the test strip. This device required about 4 min to perform each test.

Roche TesTcup 5. This combination collection/test cup has five single-assay test strips built directly into the wall of the device. Each strip also has a single control line. The specimen is provided directly into the cup, and the cap is secured. To activate the test, the cap is rotated further to the test position, and the cup is partially inverted. Inverting for 10 s allows the specimen access to the test strips. The cup is returned to the upright position, and results are read by removing a protective adhesive strip over the results window. Positive test results are indicated by the absence of a colored line. This device required about 4 min to perform each test.

Biosite Triage. This cassette-type device requires the addition of a fixed volume of specimen (via a specially provided pre-calibrated syringe with disposable tips) to a well in the cassette that contains three lyophilized reagent beads. The reaction mixture is allowed to incubate in the cassette well for 10 min. The tip on the syringe is then replaced, and the incubated reaction mixture is drawn into the syringe and spread across the cassette's multi-assay test strip. The test strip also has both a positive and negative control line. After rapid absorption of the mixture into the strip, three drops of a wash solution are added and the results read within 5 min. For this device, positive results are indicated by the presence of a colored line across the test strip (unlike the other devices, which are read in the opposite manner, with positive results indicated by the absence of a colored line). This device took the longest to perform, about 13 min, because of the 10-min incubation time.

Casco-Nerl microLINE Drug Screen Card. This cassette-type device has five single-assay test strips extending below the cassette. Each assay strip also has a single control line. The test kit includes a plain specimen cup into which the strips protruding from the cassette device are inserted into the specimen. After the wicking process is observed in the results windows (about 30 s), the device is removed from the specimen cup and the protruding test strips are covered with a plastic protective cover. Results are read within 3–8 min. Positive test results are indicated by the absence of a colored line. This device required about 4 min to perform each test.

Study design

The study design, using a selected clinical specimen set weighted around the device immunoassay cutoffs, is effectively the same as that used by this author in two previous comprehensive studies of non-instrumented devices performed by Duo Research, Inc., one for the Administrative Office of the U.S. Federal Courts in 1996 (8) and one for SAMHSA in 1998 (5).

It is clear that drug-free specimens, as well as strongly positive specimens, should be easily and correctly identified by the devices. Accordingly, testing such specimens provides little guidance on any performance differences between devices. In contrast, testing specimens with immunoreactive concentrations of

drugs and metabolites near the devices' immunoassay cutoffs provide a more rigorous challenge of these devices' capabilities. In particular, such a challenging specimen set allows an assessment of how well the manufacturers have established device cutoffs claiming to match the screening cutoffs specified by SAMHSA. It is important to know whether a device's cutoff is functionally shifted to higher levels such that the device proves conservative (i.e., not identifying as positive specimens with drug concentrations just above but near the cutoff) or aggressive (i.e., identifying as positive specimens containing drug below but near the cutoff). Furthermore, such a challenging specimen set allows a determination of how sharp or well-defined the cutoff is and its ease of readability by the operators.

Clinical specimens submitted by U.S. Federal Probation sites for routine drug testing at a SAMHSA-certified laboratory were chosen for inclusion in the study based on their immunoassay screening results from a high-volume automated immunoassay analyzer (Hitachi 747 using Diagnostic Reagents, Inc. immunoassay reagents). Once the specimen set was chosen, results from this analyzer were not used further in the study. For each drug category, the specimens were either from frozen storage or freshly submitted. The specimens were selected such that 10 specimens were drug-free (with immunoassay screening rates at the level of a negative control), 20 were close to but below the immunoassay screening cutoff, 20 were close to but above the immunoassay screening cutoff, and 10 were more clearly above the immunoassay screening cutoff but yet not strongly positive. Thus, two-thirds of all specimens were chosen to have immunoreactive levels around the screening cutoff. All selected specimens had gas chromatographic–mass spectrometric (GC–MS) analyses performed for the specific analyte in question. The devices were tested with 60 clinical specimens each for amphetamines, cannabinoids, and cocaine and 59 clinical specimens each for opiates and PCP. The final specimen distribution by GC–MS levels and confirmation criteria is shown in Table I.

Having all of the selected specimens analyzed by GC–MS before inclusion in the study was important. This ensured that all specimens showing immunoreactivity for amphetamines actually demonstrated the presence of amphetamine and/or methamphetamine and that any of the observed screening test immunoreactivity was not due solely to the presence of high levels of potentially cross-reacting amphetamine-like materials. In addition, specimens showing immunoreactivity for opiates were included only if they contained morphine and/or codeine and not other opiates, so as to avoid cross-reactivity issues from other immunoreactive opiates such as hydrocodone or hydromorphone. Finally, as clinical specimens actually containing PCP were rare, 35 of the PCP-containing specimens used the study were prepared by diluting PCP-containing specimens with drug-free urine.

All of the selected specimens were stored refrigerated at 4°C pending testing in the study.

Experimental protocol

The tests were performed independently by two operators, one with a strong laboratory background (B.S. in Biochemistry) and the other with no laboratory or scientific experience. Before the start of the study, the operators were given each the device's package insert to read, instructed in the use and operations of all devices, and tested five levels of control specimens to become familiar with the testing and reading operations. All test results were independently read and recorded by both operators. The total number of results was 2980 (298 specimens tested on each of five devices, with each result independently read by two operators).

In addition, a benchtop automated immunoassay analyzer, the Syva ETS using Emit d.a.u. reagents, was also included to compare the performance of the objectively read analyzer against the subjectively read non-instrumented test devices. The Syva ETS analyzer was operated by the principal investigator, independently from the non-instrumented device

Table I. Specimen Distribution

	# Negatives	# GC–MS Confirmed Positives	Range (ng/mL)	Average (ng/mL)
Cocaine (BE)	20 (11 at 0)	40 (≥ 150 ng/mL)	0–765	320
THC (THC-COOH)	30 (10 at 0)	30 (≥ 15 ng/mL)	0–25	15
Amphetamines	42 (11 at 0)	18		
Amphetamine		11 (≥ 500 ng/mL Amphetamine)	0–1959	488
Methamphetamine		16 (≥ 500 ng/mL Methamphetamine + ≥ 200 ng/mL Amphetamine)	0–1861	697
Opiates	31 (10 at 0)	28		
Morphine		18 (≥ 300 ng/mL)	0–871	312
Codeine		12 (≥ 300 ng/mL)	0–986	370
PCP	36 (10 at 0)	23 (≥ 25 ng/mL)	0–55	27
Total	159 (52 at 0)	139		

testing by the two other principal operators.

Each day, specimens from one of the five drug classes were tested on all non-instrumented devices as well as on the Syva ETS analyzer. The test operators knew which drug they were testing for each day but not whether the specimens were drug-free, borderline, or clearly positive. This was done solely for expediency in requiring assessment and recording of only one drug assay's results from the multi-assay devices. Specimens were tested in 2 groups of 10. Refrigerated specimens, identified only by barcode labels, were brought to room temperature, and 1-mL aliquots were removed for testing on the Syva ETS analyzer. The remainder of the specimen was tested by the two operators on the five different non-instrumented devices, one type of device at a time. Thus, each operator had 10 specimens, and these specimens were tested first with one type of device. After reading and recording the results for their 10 specimens, the operators quickly moved to each other's stations and read the other operator's developed devices, ensuring that all testing and reading was performed within the manufacturer's specified time limits. Then the next type of device was tested with the same 2 sets of 10 specimens, and so forth. Although each specimen's test on a given device was performed by only one of the operators, every test result was independently read and recorded on their own daily test sheets by both operators. Results from the Syva ETS analyzer were recorded separately from the device results, and the device operators were not informed of the Syva ETS analyzer's results nor of each other's readings.

Test results as read by the operators were recorded as one of three options: clearly negative, borderline, or clearly positive. Although all the test devices' package inserts indicate that any equivocal or borderline results should be reported as negative to be conservative, for the purposes of this study it was important to capture information on when the test results were viewed as equivocal by the operators. Thus, the operators were given the additional reporting option of the borderline result designation. However, for performance assessment, any recorded borderline results were scored as negative per package insert instructions.

For the primary performance assessment, the standard was confirmability using SAMHSA GC–MS confirmation cutoff criteria in place at the time of the study (Table II).

Table II. GC–MS Confirmation Cutoff Criteria

	GC–MS cutoff (ng/mL)
Cocaine (BE)	150
THC (THC-COOH)	15
Amphetamines	
Amphetamine	500
Methamphetamine	500 with 200 Amphetamine
Opiates	
Morphine	300
Codeine	300
PCP	25

In addition to assessing device performance using SAMHSA GC–MS confirmation criteria, device performance was also assessed by drug presence or absence in the specimen by GC–MS limit of detection criteria. The reason for making this alternative performance assessment is that in certain testing scenarios it may be useful to know that a device's positive screening result correctly identified the presence of drug in a specimen, even though the level of the drug may not be sufficient to be confirmed under SAMHSA GC–MS confirmation criteria. When appropriate in a given testing program, confronting a drug user with such an initial positive result may lead to an admission of drug use, even though subsequent confirmation testing, if performed, may return an unconfirmed result. Alternatively, it may be useful to know if a non-instrumented device's negative result truly indicates no drug present, rather than levels simply below confirmation cutoffs.

Results

The performance of the test devices was assessed in terms of five standard criteria: sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and overall accuracy.

Performance of non-instrumented devices versus Syva ETS analyzer

Overall the non-instrumented urine drug-testing devices performed well relative to the Syva ETS analyzer, even considering the especially challenging near cutoff specimen set. Using GC–MS confirmation criteria, the five non-instrumented devices had an overall accuracy of 70% compared to the Syva ETS analyzer's overall accuracy of 80%. The non-instrumented devices as a group demonstrated comparable accuracy to the Syva ETS analyzer for cocaine (82% vs. 80%) and for amphetamines (71% vs. 72%), but lower accuracy than the Syva ETS analyzer for cannabinoids (63% vs. 72%), opiates (72% vs. 90%), and PCP (62% vs. 85%). However, in six individual device/assay instances, the non-instrumented devices outperformed the Syva ETS analyzer. For cocaine, the accuracy of the Syva RapidTest (93%), Syva RapidCup (86%), and Roche TesTcup (88%) outperformed the Syva ETS analyzer (80%). For amphetamines, the Roche TesTcup (80%), Biosite Triage (73%), and Casco-Nerl microLINE (73%) outperformed the Syva ETS analyzer (72%).

Individual device performance for five drugs

The overall accuracy of each non-instrumented device for the five drugs was similar, varying between 74% for the Syva RapidTest and 66% for the Syva RapidCup. The Syva RapidCup had the best overall sensitivity at 95%, reflective of this device's relatively aggressive nature, whereas Biosite Triage had very low sensitivity at only 32%, reflecting this device's relatively conservative nature, at least with these near cutoff specimens (Table III).

Device performance by drug class

Amphetamines. Overall the non-instrumented devices performed comparably to the Syva ETS analyzer for amphetamines (71% accuracy for the devices vs. 72% for the analyzer). The Roche TesTcup had the highest accuracy (80%) and Syva RapidCup the lowest (58%) when using SAMHSA GC–MS confirmation criteria. A key performance determinant among the devices was their detection of specimens containing methamphetamine above 500 ng/mL but less than 200 ng/mL amphetamine as required for methamphetamine-confirmed results under current SAMHSA criteria. Sensitivity for all devices except Biosite Triage was excellent (92–94%) with accordingly very high negative predictive values (0.95–1); Biosite Triage had a very low 19% sensitivity (Table IV).

Because the devices have varying antibody sensitivity to amphetamine and methamphetamine, it is important to fully describe the specimen set utilized to fairly interpret these performance results. For the 60 amphetamines specimens, there were 11 drug-free specimens with no amphetamine or methamphetamine detected by GC–MS and 18 confirmable specimens fulfilling the SAMHSA criteria of either 500 ng/mL amphetamine or a combination of 500 ng/mL methamphetamine along with 200 ng/mL amphetamine. Of these 18 confirmable specimens, 9 were confirmable under both the amphetamine and methamphetamine confirmation criteria. Of the 31 remaining specimens, all contained amphetamine and/or methamphetamine, but at levels insufficient to meet the SAMHSA confirmation criteria. Two of these unconfirmable specimens had amphetamine only at levels below 500 ng/mL. Eighteen of the unconfirmable specimens had methamphetamine only (with six of those with methamphetamine levels > 500 ng/mL). The other 11 unconfirmable specimens had both amphetamine and methamphetamine, 7 with both methamphetamine and amphetamine below 500 ng/mL and 4 with methamphetamine levels above 500 ng/mL and amphetamine levels below 200 ng/mL. Thus, there was a total of 10 specimens with > 500 ng/mL of methamphetamine, but with insufficient amphetamine to be confirmable under current SAMHSA GC–MS confirmation criteria (six of these had no amphetamine and four had amphetamine present but < 200 ng/mL). Detection of these 10 specimens that contained over 500 ng/mL of methamphetamine may not be considered a deficiency for these screening devices, even though the screening results would not be confirmable under SAMHSA criteria. For these 10 specimens, Syva RapidCup detected 10 of 10, microLINE and the Syva ETS analyzer 9 of 10, Syva RapidTest 8 of 10, Roche TesTcup 4 of 10, and Triage 1 of 10.

Cannabinoids. Overall, the non-instrumented devices performed slightly more poorly than the Syva ETS analyzer for cannabinoids (63% accuracy for the non-instrumented devices vs. 72% for the Syva ETS analyzer). Casco-Nerl microLINE had the highest accuracy (71%), and Biosite Triage had the lowest (50%). Surprisingly, Biosite Triage did not detect any of the 30 GC–MS-confirmable cannabinoid specimens in the study. In fact, there were no positive cannabinoids results recorded by either operator for any cannabinoids specimen. However, it should be noted that the highest GC–MS level of THC-COOH in any specimen was only 25 ng/mL, relatively close to the GC–MS confirmation cutoff of 15 ng/mL. Syva RapidCup had the highest sensitivity at 85%, whereas Syva RapidTest's sensitivity was quite low at 35% (Table V).

Cocaine. Overall, the non-instrumented devices performed slightly better than the Syva ETS analyzer for cocaine (82% accuracy for the non-instrumented devices vs. 80% for the Syva ETS analyzer). The Syva RapidTest had the highest accuracy

Table III. Device Performance for All Five Drugs versus GC–MS

	Sensitivity	Specificity	PPV	NPV	Accuracy
Syva RapidTest	83%	66%	0.68	0.82	74%
Syva RapidCup	95%	41%	0.58	0.90	66%
Roche TesTcup	89%	58%	0.65	0.86	73%
Biosite Triage	32%	98%	0.93	0.62	67%
Casco-Nerl microLINE	72%	67%	0.66	0.74	70%
Syva ETS analyzer	81%	79%	0.77	0.82	80%

Table IV. Device Performance for Amphetamines versus GC–MS

	Sensitivity	Specificity	PPV	NPV	Accuracy
Syva RapidTest	94%	60%	0.5	0.96	70%
Syva RapidCup	94%	43%	0.41	0.95	58%
Roche TesTcup	94%	74%	0.61	0.97	80%
Biosite Triage	19%	95%	0.64	0.73	73%
Casco-Nerl microLINE	92%	65%	0.53	0.95	73%
Syva ETS analyzer	100%	60%	0.51	1	72%

Table V. Device Performance for Cannabinoids versus GC–MS

	Sensitivity	Specificity	PPV	NPV	Accuracy
Syva RapidTest	35%	83%	0.68	0.56	59%
Syva RapidCup	85%	48%	0.62	0.76	67%
Roche TesTcup	62%	73%	0.70	0.66	68%
Biosite Triage	0%	100%	No positive results	0.50	50%
Casco-Nerl microLINE	82%	60%	0.67	0.77	71%
Syva ETS analyzer	53%	90%	0.84	0.66	72%

(93%), and Biosite Triage had the lowest (65%). Three of the devices showed very high sensitivity with the Syva RapidTest at 100%, Syva RapidCup at 99%, and the Roche TesTcup at 93%. However, Biosite Triage had a very low sensitivity of 48%. All the devices had high positive predictive values (0.83–1) (Table VI).

Opiates. For opiates the non-instrumented devices performed slightly more poorly than the Syva ETS analyzer (72% accuracy for the devices vs. 90% accuracy for the Syva ETS analyzer). Biosite Triage had the highest accuracy (84%), and Syva RapidCup had the lowest (64%). Casco-Nerl microLINE had very low sensitivity (32%) (Table VII).

PCP. For PCP, the non-instrumented devices demonstrated lower accuracy than the Syva ETS analyzer (62% for the non-instrumented devices vs. 85% for the Syva ETS analyzer). Syva RapidTest had the highest accuracy (75%), and Syva RapidCup had the lowest (56%). All devices except Biosite Triage had very high sensitivity (93–100%), with Biosite Triage demonstrating only 11% sensitivity (Table VIII).

Table VI. Device Performance for Cocaine versus GC–MS

	Sensitivity	Specificity	PPV	NPV	Accuracy
Syva RapidTest	100%	78%	0.9	1	93%
Syva RapidCup	99%	60%	0.83	0.96	86%
Roche TesTcup	93%	78%	0.89	0.84	88%
Biosite Triage	48%	100%	1	0.49	65%
Casco-Nerl microLINE	71%	95%	0.97	0.62	79%
Syva ETS analyzer	70%	100%	1	0.63	80%

Table VII. Device Performance for Opiates versus GC–MS

	Sensitivity	Specificity	PPV	NPV	Accuracy
Syva RapidTest	96%	55%	0.66	0.94	75%
Syva RapidCup	96%	34%	0.57	0.91	64%
Roche TesTcup	100%	42%	0.61	1	69%
Biosite Triage	71%	95%	0.93	0.79	84%
Casco-Nerl microLINE	32%	98%	0.95	0.62	67%
Syva ETS analyzer	96%	84%	0.84	0.96	90%

Table VIII. Device Performance for PCP versus GC–MS

	Sensitivity	Specificity	PPV	NPV	Accuracy
Syva RapidTest	93%	64%	0.62	0.94	75%
Syva RapidCup	100%	28%	0.47	1	56%
Roche TesTcup	100%	32%	0.48	1	58%
Biosite Triage	11%	100%	1	0.64	65%
Casco-Nerl microLINE	96%	33%	0.48	0.92	58%
Syva ETS analyzer	100%	75%	0.72	1	85%

Performance by scientist and non-scientist operators

The tests were performed and independently read by two operators: one with strong scientific laboratory training and the other without any laboratory experience. However, there was little overall performance difference between them: overall accuracy for the scientist was 71% and for the non-scientist was 69%. The two operators completely agreed on 81% of their 1490 determinations (i.e., both reporting negative, borderline, or positive for a particular specimen with a particular device). Only 1.1% of their results showed complete disagreement (i.e., one reporting a clear positive with the other reporting a clear negative). The degree of operator agreement did vary somewhat between devices: Syva RapidCup 88%, Biosite Triage 80%, Syva RapidTest and Casco-Nerl microLINE 79%, and Roche TesTcup 78%. The most significant performance difference between the operators was in the rate of false-positive results, although the rate was quite low for both operators. The non-scientist reported 10 false-positive results (0.7% of 1490 total results or 3.8% of 260 results for drug-free specimens), and the scientist had only 1 false-positive result (0.07% of 1490 total results or 0.38% of 260 results for drug-free specimens).

Although there was impressive agreement between the operators' results and performance, the non-scientist was occasionally confused with the varied testing and reading procedures for the different devices. The reading of the Biosite Triage device differed distinctly from the others in that for Biosite Triage the presence of a line is read as a positive result, whereas for all the other devices, the presence of a line is read as a negative result. This potential for confusion when using several devices with inconsistent reading criteria could

Table IX. False-Positive Test Results (for 52 specimens at 0 ng/mL by GC–MS)

	Amphetamines	Cannabinoids	Cocaine	Opiates	PCP	Total
Syva RapidTest	0	0	2	0	0	2
Syva RapidCup	1	0	1	0	0	2
Roche TesTcup	0	0	1	1	4	6
Biosite Triage	0	0	0	0	0	0
Casco-Nerl microLINE	0	0	1	0	0	1
Syva ETS analyzer	0	0	0	0	0	0

Table X. Scientist/Non-Scientist Performance

	Overall accuracy	# False positive results	% Borderline results	Accuracy of borderline results
Scientist	71%	1	31%	62%
Non-scientist	69%	10	20%	67%

Table XI. Percentage and Accuracy of Borderline Results versus GC–MS

	% Borderline results	Accuracy of borderline results
Syva RapidTest	27%	75%
Syva RapidCup	10%	79%
Roche TesTcup	26%	82%
Biosite Triage	31%	34%
Casco-Nerl microLINE	32%	64%

account for the fact that of the 11 false-positive results (positive results for specimens with no detectable analyte by GC–MS), 10 were recorded by the non-scientist (although there were no false-positive reports for Biosite Triage).

It is important to carefully define the term false positive as used in this study. False-positive results are positive test results for specimens that had none of the analyte in question by GC–MS. It is important to distinguish such false-positive results from unconfirmed positive results, where device results were reported as positive and the analyte in question was in fact present by GC–MS, but at levels below the confirmation criteria.

There were 52 specimens that were drug-free by GC–MS. Of the 520 total results for these drug-free specimens, only 11 (2.1%) were incorrectly reported as positive. It is important to note that for all of the 11 false-positive results, the two readers' results were discordant. That is, for none of 11 specimens with false-positive results did both readers report a false positive result. For four of the false-positive results, the operators completely disagreed, with the non-scientist reporting a clear positive and the scientist reporting a clear negative. Because of the varied reading formats of the devices, the non-scientist was occasionally confused about whether a line on the device meant a positive or a negative result. Such confusion would not be expected in actual field situations where it was likely that only one type of device would be used. The scientist's one false-positive result was with the Roche TesTcup for cocaine, for which the non-scientist reported the result as borderline (Table IX).

The proportion of all results reported as borderline was 25% across all devices and drugs, but the rate of reporting borderline results did vary between the operators, with the scientist reporting 31% of all results as borderline and the non-scientist reporting only 20% as borderline. It is surprising that the number of borderline results was so low given that two-thirds of the specimen set was selected to be borderline. The overall accuracy of the borderline results (when scored as negative per package insert instructions) was 64%, which was comparable to the accuracy of clearly negative results (75%) and to the accuracy of clearly positive results (66%). The accuracy of the borderline results also varied slightly between the operators, with 62% accuracy for the scientist, and 67% accuracy for the non-scientist (Table X).

The proportion of borderline results and their accuracy also varied among the different devices. The Syva RapidCup had the fewest borderline results (10%), and Casco-Nerl microLINE had the most (32%). The accuracy of borderline results was highest for the Roche TesTcup (82%) and lowest for Biosite Triage (34%) (Table XI). There were also differences in the proportion and accuracy of borderline results between the five various drug classes with the proportion of borderline results varying between 18 and 34% among the drugs and the accuracy of borderline results scored as negative varying between 53 and 85%, but no significant patterns were noted.

Performance according to drug presence/absence

The accepted standard for assessing screening test performance has been confirmability using SAMHSA GC–MS confirmation criteria. However, depending on a particular testing program's goals, a screening test may not be considered deficient in identifying the presence of drug in a specimen simply because that result may not fulfill secondary confirmation testing criteria. The question asked by a drug-testing program may simply be whether or not the donor has recently used drugs, not specifically whether or not a test result is confirmable. In some testing programs, effectively identifying specimens with drug present, even if below confirmation cutoffs, may still prove useful, such as leading to an admission of use when the donor is confronted with positive screening test results. It was deemed valuable to assess the performance of these devices when the criteria are whether the specimen actually contained the analytes in question or not. Devices that are aggressive in nature, that is, having high sensitivity and identifying as positive specimens samples that have the drug in question but perhaps below cutoff, should show an improvement in performance when assessed by drug presence/absence criteria. On the other hand, devices that are conservative in nature, that is, with relatively low sensitivity and only identifying as positive specimens that are at or enough above the cutoff to ensure that device positive results are confirmable, would be expected to show a decrease in performance. Of course with specimens well below or above the cutoffs, these aggressive/conservative performance differences would not be noticed or significant.

When performance was assessed according to drug presence/absence, for the Syva RapidTest and Casco-Nerl microLINE there was little difference in overall accuracy than when using standard SAMHSA GC–MS confirmation criteria. For Roche TesTcup, the accuracy improved slightly, whereas for Syva RapidCup there was a significant increase in overall accuracy from 66% to 93%, reflecting an aggressive nature with these near cutoff specimens. In contrast, for Biosite Triage, accuracy declined significantly, reflecting a conservative nature with these near cutoff specimens. The overall specificity and

positive predictive values for all devices were very high (specificity 94–100%, positive predictive value 0.98–1). This is effectively a reflection of the fact that these devices had almost no false-positive results (positive results with drug-free specimens). Rather, the devices (other than the conservative Biosite Triage) often demonstrated positive results when there was in fact drug present even at levels below confirmation cutoff criteria. In this study, positive test results for these devices had virtually 100% accuracy in correctly indicating the presence of drug (Table XII).

Discussion

Overall, this study has demonstrated a high level of performance for these easy to use, rapid, non-instrumented drug-testing devices, especially when considering the challenging near cutoff specimen set used for the study.

It is well appreciated that such a specifically selected specimen set weighted around the immunoassay cutoff does not represent what would be expected in actual clinical or field situations. In fact, even better device performance would be expected in actual clinical and field situations than demonstrated with this challenging specimen set. The specimen distributions in typical drug testing scenarios would not be expected to be weighted around the cutoff, but rather would have much greater proportions of the more easily correctly identified drug-free and strongly positive specimens, with only a small proportion of the specimens around the immunoassay cutoff. One study of urine drug testing in an emergency room setting reported that less than 4% of specimens submitted for suspected drug use had immunoassay rates that were within three standard deviations (assay precision) of the immunoassay cutoff rate (48). The devices performed extremely well with drug-free and clearly positive specimens. The overall accuracy of the devices with the drug-free specimens was 97.9% (results reported as clearly negative). With clearly positive specimens with drug levels at least 50% above the confirmation cutoffs, device accuracy was 82.9% (results reported as clearly positive), and this performance improves to 93.8% for the four devices other than the relatively low sensitivity Biosite Triage.

Of course, one important issue of assay performance not addressed by this study is the role of potentially cross-reacting substances in actual clinical specimens. This study intentionally selected clinical specimens to contain only the analyte(s) in question. Assessing the devices' cross-reactivities to other substances was beyond the intended scope of this study. It should be noted that the devices' package inserts do specify assay cross-reactivities to a wide variety of other analytes.

Overall, four of the five devices performed comparably with each other, with the Biosite Triage device notably less sensitive (more conservative) than the others, with a very low overall sensitivity of 32%. Triage missed all 30 of the confirmable cannabinoid specimens in the study, with no positive results reported by either operator. The Syva RapidTest had the highest overall accuracy at 74%, followed by Roche TesTcup (73%), Casco-Nerl microLINE (70%), Biosite Triage (67%), and Syva RapidCup (66%).

Only 2.1% (11) of the 52 drug-free specimens (with no analyte detected by GC–MS) yielded false-positive test results, with all but one of these reported by the non-scientist. Furthermore, for none of these false-positive results did the operators agree.

Borderline or equivocal results were a surprisingly low 25% of all results considering the specimen set being weighted around the immunoassay cutoffs. When such borderline results were conservatively scored as negative, per package insert instructions, they had comparable overall accuracy (64%) to clearly negative results (75% accuracy) and clearly positive results (66%). Thus the concerns that operator uncertainty in reading borderline results would lead to poor performance appear unfounded.

Overall, both the scientist and non-scientist operators performed comparably, although the non-scientist appeared to have more difficulty switching between devices with opposite reading requirements. The main performance difference between the scientist and non-scientist was in the number of false-positive results, with the non-scientist reporting 10 false-positive results (0.7% for 1490 total results or 3.8% of 260 results for drug-free specimens) and the scientist reporting only 1 false-positive (0.07% of 1490 total results or 0.38% of 260 results for drug-free specimens). Nonetheless, both operators completely agreed on their results 81% of the time, in spite of the near cutoff specimen set that would have been expected to lead to a large number of equivocal results. This indicates that the device manufacturers have done a good job in ensuring that the devices provide reasonably clear and distinct endpoints.

When using drug presence/absence instead of standard confirmation cutoffs as the performance criteria, the Syva RapidCup had the highest accuracy overall (increasing from 66% to 93%), and Biosite Triage had the lowest (decreasing from 67% to 34%), with all the devices demonstrating ex-

Table XII. Device Performance for Five Drugs versus Drug Presence/Absence

	Sensitivity	Specificity	PPV	NPV	Accuracy	Accuracy GC–MS criteria
Syva RapidTest	68%	98%	0.99	0.40	73%	74%
Syva RapidCup	92%	98%	1	0.71	93%	66%
Roche TesTcup	76%	94%	0.98	0.45	79%	73%
Biosite Triage	20%	100%	1	0.21	34%	67%
Casco-Nerl microLINE	62%	99%	1	0.35	68%	70%
Syva ETS analyzer	59%	100%	1	0.34	66%	80%

tremely high positive predictive values (0.98–1). In drug-testing programs where it is important not to miss drug users, the more aggressive devices would be preferred, especially considering that as currently practiced there is the opportunity for some form of confirmation, for example, by an admission of use, clinical observations, or further confirmation testing. Alternatively, if the device screening test results might be used without any form of confirmation, which is unlikely, or a testing program wanted to ensure that positive test results are confirmable, then a conservative device would be preferred, but with the trade-off of missing drug users. The results of this comparative study allow each drug-testing program to choose the type of device, aggressive or conservative, that best fits its goals. It should be emphasized that although these devices demonstrated their respective aggressive or conservative natures with these near cutoff specimens, in actual field use these differences may be minimized as the specimen distributions should be quite different.

The results from the present study are comparable to those observed in two similar comprehensive studies, also performed by the author with Duo Research, Inc. The first of these was performed in 1996 commissioned by the Administrative Office of the U.S. Courts in response to the issuance of President Clinton's Executive Order in 1995 that all federal arrestees be subjected to drug testing.(8) That study of 15 non-instrumented devices found that many of these devices performed amazingly well, especially considering a challenging specimen set weighted around the immunoassay cutoffs, with accuracies judged by standard GC–MS confirmation comparable to a benchtop automated immunoassay analyzer (Syva ETS using Emit d.a.u. reagents). The non-instrumented devices demonstrated an overall accuracy of 71% (52–79%) compared with the Syva ETS analyzer's average of 80% (78–82%). A second similar study was commissioned in 1998 by the Substance Abuse and Mental Health Services Administration (SAMHSA) with similar impressive results, with the 15 devices demonstrating an overall accuracy of 70% (61–78%) versus the Syva ETS analyzer's 76% (5).

These three studies together clearly demonstrate the impressive capabilities of these non-instrumented urine drug-testing devices and serve as a testament to the advances in immunoassay technology. These devices are expected to demonstrate even higher accuracies with specimen populations actually encountered in routine clinical, workplace, and corrections settings, where specimens should not be weighted around the immunoassay cutoffs as in these studies.

Conclusions

This study has demonstrated the impressive performance capabilities of five non-instrumented drug-testing devices with a challenging near cutoff specimen set. Both scientifically trained and non-scientist operators were able to demonstrate comparable overall performance, although the non-scientist reported 10 false-positive results (0.7% of 1490 total results or 3.8% of 260 results for drug-free specimens), and the scientist reported only 1 false-positive result (0.07% of 1490 total results or 0.38% of 260 results for drug-free specimens). Overall performance using GC–MS confirmability criteria was comparable to an automated immunoassay analyzer (Syva ETS using Emit d.a.u. reagents), with some devices occasionally matching or even exceeding the performance of the Syva ETS analyzer. Furthermore, when assessed against the criteria of drug presence and absence, device performance generally improved, with all devices showing extremely high positive predictive values. It is clear that such non-instrumented devices should be able to fulfill the due process requirements in a wide variety of settings. What remains is for there to be significant case law establishing judicial recognition of the accuracy of these devices and acceptability of their use in different testing contexts.

Acknowledgments

The author thanks the following: the test operators Jeffrey Bedeaux and Phyllis Melmon-Martinez; Mr. Neil Fortner and the staff at PharmChem Laboratories, Menlo Park, CA, for their generous support including provision of laboratory space, initial immunoassay screening results for specimen selection, and GC–MS analyses; the Administrative Office of the U.S. Courts for authorizing use of specimens which were to be discarded; and Syva Company, a Division of Dade Behring Diagnostics, Inc., for their financial support of this study and provision of the Syva ETS analyzer and Emit d.a.u. reagents.

References

1. R.A. Belfer, B.L. Klein, D.A. Boenning, and S.J. Soldin. Emergency department evaluation of a rapid assay for detection of cocaine metabolites in urine specimens. *Pediatr. Emerg. Care* **12(2):** 113–115 (1996).
2. J.G. Schwartz, I.L. Hurd, and J.J. Carnahan. Determination of tricyclic antidepressants for ED analysis. *Am. J. Emerg. Med.* **12:** 513–516 (1994).
3. A.H.B. Wu, S.S. Wong, K.G. Johnson, J. Callies, D.X. Shu, W.E. Dunn, and S.H.Y. Wong. Evaluation of the Triage system for emergency drugs-of-abuse testing in urine. *J. Anal. Toxicol.* **17:** 241–245 (1993).
4. E. Welch, L.E. Fleming, I. Peyser, W. Greenfield, B.W. Steele, and E.S. Bandstra. Rapid cocaine screening of urine in a newborn nursery. *J. Pediatr.* **123:** 468–470 (1993).
5. R. Willette and L. Kadehjian. An evaluation of non-instrumented drug test devices for the Division of Workplace Programs, Center for Substance Abuse Prevention, Department of Health and Human Services. Summary Report for Workplace Programs, Duo Research, Inc., 1998. Available in part at www.health.org/workplace/summary.htm. Presented at the annual meeting of the Society of Forensic Toxicologists, San Juan, Puerto Rico, October 1999.
6. R.H. Schwartz, H.W. Clark, and P.S. Meek. Laboratory tests for rapid screening of drugs of abuse in the workplace: a review. *J. Addict. Dis.* **12(2):** 43–56 (1993).
7. J.L. Valentine and E.M. Komorowski. Use of a visual panel detection method for drugs of abuse: clinical and laboratory experience with children and adolescents. *J. Pediatr.* **126(1):** 135–140 (1995).
8. R. Willette and L. Kadehjian. An evaluation of non-instrumented drug tests for the Administrative Office of the U.S. Courts, Duo Research, Inc. Presented at the annual meeting of the Society of Forensic Toxicologists, Salt Lake City, UT. October 1997.

9. B.J. Buchan, J.M. Walsh, and P.E. Leaverton. Evaluation of the accuracy of on-site multi-analyte drug testing devices in the determination of the prevalence of illicit drugs in drivers. *J. Forensic Sci.* **43(2):** 395–399 (1998).
10. D. Brookoff, C.S. Cook, C. Williams, and C.S. Mann. Testing reckless drivers for cocaine and marijuana. *N. Engl. J. Med.* **331:** 518–522 (1994).
11. S. Nightingale. New rule proposed for home drug abuse test kits. *J. Am. Med. Assoc.* **279(17):** 1339 (1998).
12. S. Nightingale. Over-the-counter test system for drugs of abuse approved. *J. Am. Med. Assoc.* **277(9):** 703 (1997).
13. J. Wise. Drug testing kit approved for sales in U.S. *Br. Med. J.* **314:** 324 (1997).
14. R. Schilling, B. Bidassie, and N. El-Bassel. Detecting cocaine and opiates in urine: comparing three commercial assays. *J. Psychoactive Drugs* **31(3):** 305–313 (1999).
15. D. Crouch, J.L. Frank, L.J. Farrell, H.M. Karsch, and J.E. Klaunig. A multiple-site laboratory evaluation of three on-site urinalysis drug-testing devices. *J. Anal. Toxicol.* **22:** 493–502 (1998).
16. S.D. Ferrara, L. Tedeschi, G. Frison, G. Brusini, F. Castagna, B. Bernardelli, and D. Soregaroli. Drugs-of-abuse testing in urine: statistical approach and experimental comparison of immunochemical and chromatographic techniques. *J. Anal. Toxicol.* **18:** 278–291 (1994).
17. R. de la Torre, A. Domingo-Salvany, R. Badia, G. Gonzalez, D. McFarlane, L. San, and M. Torrens. Clinical evaluation of the Triage analytical device for drugs-of-abuse testing. *Clin. Chem.* **42(9):** 1433–1438 (1996).
18. A.D. Fraser and R. Meatherall. Comparative evaluation of five immunoassays for the analysis of alprazolam and triazolam metabolites in urine: effect of lowering the screening and GC–MS cut-off values. *J. Anal. Toxicol.* **20:** 217–223 (1996).
19. J. Röhrich, K. Schmidt, and H. Bratzke. Application of the novel immunoassay TRIAGE to a rapid detection of antemortem drug abuse. *J. Anal. Toxicol.* **18:** 407–414 (1994).
20. T.R. Koch, R.L. Raglin, S. Kirk, and F.J. Bruni. Improved screening for benzodiazepine metabolites in urine using the Triage Panel for drugs of abuse. *J. Anal. Toxicol.* **18:** 168–172 (1994).
21. L.E. Edinboro and A. Poklis. Detection of benzodiazepines and triazolams by Triage: confirmation by solid-phase extraction utilizing SPEC•3ML•MP3 microcolumns and GC–MS. *J. Anal. Toxicol.* **18:** 312–316 (1994).
22. K.F. Buechler, S. Moi, B. Noar, D. McGrath, J. Villela, M. Clancy, A. Shenhav, A. Colleymore, G. Valkirs, T. Lee, J.F. Bruni, M. Walsh, R. Hoffman, F. Ahmuty, M. Nowakowski, J. Beuchler, M. Mitchell, D. Boyd, N. Stiso, and R. Anderson. Simultaneous detection of seven drugs of abuse by Triage Panel for drugs of abuse. *Clin. Chem.* **38(9):** 1678–1684 (1992).
23. A.J. Jenkins, W.D. Darwin, M.A. Huestis, E.J. Cone, and J.M. Mitchell. Validity testing of the AccuPinch THC test. *J. Anal. Toxicol.* **19:** 5–12 (1995).
24. O. Beck, M. Kraft, M.R. Moeller, B.L. Smith, S. Schneider, and R. Wennig. Frontline immunochromatographic device for on-site urine testing of amphetamines: laboratory validation using authentic specimens. *Ann. Clin. Biochem.* **37:** 199–204 (2000).
25. R. Wennig, M.R. Moeller, J.M. Haguenoer, A. Marocchi, F. Zoppi, B.L. Smith, R. de la Torre, C.A. Carstensen, A. Goerlach-Graw, J. Schaeffler, and R. Leinberger. Development and evaluation of immunochromatographic rapid tests for screening of cannabinoids, cocaine, and opiates in urine. *J. Anal. Toxicol.* **22:** 148–155 (1998).
26. A.J. Jenkins, L.C. Mills, W.D. Darwin, M.A. Huestis, E.J. Cone, and J.M. Mitchell. Validity testing of the EZ-SCREEN cannabinoid test. *J. Anal. Toxicol.* **17:** 292–298 (1993).
27. R.H. Schwartz, S. Bogema, and M.M. Thorne. Evaluation of the EZ-SCREEN enzyme immunoassay test for detection of cocaine and marijuana metabolites in urine specimens. *Pediatr. Emerg. Care* **6(2):** 147–149 (1990).
28. D.J. Crouch, M.L. Cheever, D.M. Andrenyak, D.J. Kuntz, and D.L. Loughmiller. A comparison of ONTRAK TESTCUP, Abuscreen ONTRAK, Abuscreen ONLINE, and GC/MS urinalysis test results. *J. Forensic Sci.* **43(1):** 35–40 (1998).
29. J. Towt, S.-C.J. Tsai, M.R. Hernandez, A.D. Klimov, C.V. Kravec, S.L. Rouse, H.S. Subuhi, B. Twarowska, and S.J. Salamone. ONTRAK TESTCUP: a novel, onsite, multi-analyte screen for detection of abused drugs. *J. Anal. Toxicol.* **19:** 504–510 (1995).
30. D.A. Armbruster and J.M. Krolak. Screening for drugs of abuse with the Roche ONTRAK assays. *J. Anal. Toxicol.* **16:** 172–175 (1992).
31. R.H. Schwartz, S. Bogema, and M.M. Thorne. Evaluation of a rapid latex-particle agglutination-inhibition screening assay for cocaine in urine. *J. Pediatr.* **117(4):** 670–672 (1990).
32. E.R.S. Brown, D.R. Jarvie, and D. Simpson. Evaluation of Bionike One-Step tests for the detection of drugs of abuse in urine. *Ann. Clin. Biochem.* **34:** 74–80 (1997).
33. T. Korte, J. Pykäläinen, P. Lillsunde, and T. Seppälä. Comparison of RapiTest with Emit d.a.u. and GC–MS for the analysis of drugs in urine. *J. Anal. Toxicol.* **21:** 49–53 (1997).
34. P. Kintz, B. Godelar, and V. Cirimele. Evaluation of the Rapidrog cannabis noninstrumental immunoassay. *J. Anal. Toxicol.* **21:** 243 (1997).
35. J.H. Lewis. Interference of gemfibrozil with Roche TesTcup. *J. Anal. Toxicol.* **23:** 384 (1999).
36. E.H. Taylor, E.H. Oertli, J.W. Wolfgang, and E. Mueller. Accuracy of five on-site immunoassay drugs-of-abuse testing devices. *J. Anal. Toxicol.* **23:** 119–124 (1999).
37. J.J.W. Ros, M.G. Pelders, and A.C.G. Egbert. Performance of Abusign drugs-of-abuse slide tests with particular emphasis on concentrations near the cut-off: comparison with FPIA–ADx and confirmation of results with GC–MS. *J. Anal. Toxicol.* **22:** 40–44 (1998).
38. A. Poklis and C.L. O'Neal. Potential for false-positive results by the Triage™ panel of drugs-of-abuse immunoassay. *J. Anal. Toxicol.* **20:** 209–210 (1996).
39. H.R. Kranzler, J. Stone, and L. McLaughlin. Evaluation of a point-of-care testing product for drugs of abuse; testing site is a key variable. *Drug Alcohol Depend.* **40:** 55–62 (1995).
40. S. George and R.A. Braithwaite. A preliminary evaluation of five rapid detection kits for on-site drugs of abuse screening. *Addiction* **90:** 227–232 (1995).
41. E.J. Cone and J. Mitchell. Validity testing of commercial urine cocaine metabolite assays: II. Sensitivity, specificity, accuracy, and confirmation by gas chromatography/mass spectrometry. *J. Forensic Sci.* **34(1):** 32–45 (1989).
42. R.H. Schwartz, S. Bogema, and M.M. Thorne. Evaluation of InstaScreen cannabinoid drug screen test. *Arch. Pathol. Lab. Med.* **113:** 1204 (1989).
43. R.M. Kaplan, F. Fochtman, P. Brunett, and C. White. An analysis of clinical toxicology urine specimens using the KDI Quik Test. *Clin. Toxicol.* **27(6):** 369–373 (1989).
44. R.H. Schwartz, S. Bogema, and M.M. Thorne. Evaluation of the Keystone Diagnostic Quik Test. *Arch. Pathol. Lab. Med.* **113:** 363–364 (1989).
45. H. Wells. A discussion of the KDI Quik Test drug screen. *J. Anal. Toxicol.* **12:** 111 (1988).
46. E.J. Cone and S.L. Menchen. Lack of validity of the KDI Quik Test drug screen for detection of benzoylecgonine in urine. *J. Anal. Toxicol.* **11:** 276–277 (1987).
47. S. Bogema, R. Schwartz, and I. Godwin. Evaluation of the Keystone Diagnostics Quik test using previously screened urine specimens. *J. Anal. Toxicol.* **12:** 272–273 (1988).
48. M.W. Linder and R. Valdes, Jr. Distribution of the magnitude of Emit drug assay signals from emergency room patients. *Clin. Chem.* **41(6):** S117 (1995).

Manuscript received October 5, 2000;
revision received January 8, 2001.