A Statistical Approach to the Prediction of Verifiable Heroin Use from Total Codeine and Total Morphine Concentrations in Urine

REFERENCE: Fuller DC. A statistical approach to the prediction of verifiable heroin use from total codeine and total morphine concentrations in urine. J Forensic Sci 1997;42(4):685–689.

ABSTRACT: There has been much debate in urine drug testing over what criteria should be applied to total codeine and total morphine concentration data to determine the likelihood that a urine donor has used heroin and whether such use can be demonstrated by the presence of 6-acetylmorphine. After determining that the stability of 6-acetylmorphine in frozen urine is adequate for a period of at least two years, a database of over 100 codeine and/or morphine positive urine specimens was subjected to relative operating characteristic analysis to identify a criterion that would indicate a high probability of detecting 6-acetylmorphine in a specimen and thus confirming heroin use. A two-fold criterion was identified. By using a criterion that requires the total morphine concentration to be greater than 5.000 mg/L and the total codeine to total morphine ratio to be less than 0.125, one can predict the presence of 6acetylmorphine with a sensitivity of 92%, a specificity of 79%, and an overall accuracy of 73%. Although this criterion is statistically the most accurate in terms of both sensitivity and specificity for the data analyzed by the author, the results of other criteria are presented to aid toxicologists and medical review officers in determining if analysis for 6-acetylmorphine is likely to produce useful results.

KEYWORDS: forensic science, forensic toxicology, heroin, codeine, morphine, 6-acetylmorphine, statistics

According to Drug Abuse Warning Network (DAWN) the use of heroin has made a resurgence as a popular drug of abuse. DAWN reports that heroin related episodes increased 44% from the first half of 1992 to the first half of 1993 (1). As the incidence of heroin use rises so must the sophistication of urine drug testing to differentiate donors that are using heroin from other types of opiate use. Most urine drug testing programs analyze for total codeine and total morphine in response to a positive opiate immunoassay screen. Interpretation of the results of such assays must take into account the possibility of poppy seed ingestion, codeine use, pharmaceutical morphine use, and heroin use. As a result of the confusing array of possibilities, the Department of Transportation (DOT) estimates that more than 90% of confirmed codeine and morphine positives are reversed after medical review officer review (2). Toxicologists and medical review officers (MRO) need an objective way in which to evaluate quantitative total codeine and total morphine data to determine if the added time and expense of 6-acetylmorphine analysis to demonstrate heroin use is warranted. In the present study the author has applied relative operating characteristic (ROC) analysis to quantitative total codeine and total morphine data from urine donors and has identified criteria that can be utilized to predict the presence of 6-acetylmorphine, which in turn constitutes prima facie evidence of heroin use.

Materials and Methods

Analysis

Immunoassay analyses of urine specimens were performed on an Instrumentation Laboratory Monarch 2000 autoanalyzer using Syva EMIT[®] immunoassay reagents at an opiate cutoff concentration of 300 ng/mL (0.300 mg/L). The manufacturers recommended procedures were followed.

Total codeine and total morphine analysis was performed by GC/MS as follows: 1 mL of 6N hydrochloric acid and 100 mcL of a 2.5 mcg/mL solution of d₃-codeine and d₃-morphine internal standards were added to a 20-mL culture tube containing aliquots of donor urine, typically 100 mcL to 1 mL, depending on the immunoassay results. Specimens less than 1 mL were diluted to approximately 1 mL with deionized water prior to hydrolysis. The loosely capped tubes were then placed in a beaker of boiling water for 1 h. After cooling, the specimens were adjusted to approximately pH 7 by the addition of 900 mcL of 6N NaOH and dry sodium bicarbonate as required. After activating an SPE column (Worldwide Monitoring ZSDAU020) by adding and eluting 3 mL of methanol and 3 mL of deionized water, the specimens were applied to the column at an elution rate of approximately 1-2 mL/ min. The columns were washed by the consecutive addition and elution of 3 mL of deionized water, 3 mL of 0.1M pH 4.5 sodium acetate buffer, and 3 mL of methanol. The columns were then dried at maximum vacuum for 1-2 min. Codeine and morphine were eluted with 3 mL of 80:20:2 methylene chloride:isopropanol:ammonium hydroxide. After evaporating the elution solvent to dryness at 60°C under a stream of nitrogen, codeine and morphine were derivatized by reconstituting with 200 mcL of chloroform and 100 mcL of trifluoroacetic anhydride and heating in a capped tube at 70°C for 15 min. The mixture was then evaporated to dryness under a stream of nitrogen and reconstituted with 100 mcL of chloroform prior to analysis by GC/MS.

Analyses were performed on a Hewlett-Packard Model 5970 MSD with 7673A autosampler and a 12 m by 0.2 mm by 0.33 micron film HP Ultra-1 capillary column. The instrument was programmed from 150° to 300°C at 12°C/min. The injection port temperature was 250°C and operated in the splitless mode. Ions

¹Director of Toxicology, Sierra Nevada Laboratories, Laboratory Corporation of America, Reno, NV.

Received 25 April 1996; and in revised form 5 Sept. 1996; accepted 11 Oct. 1996.

monitored were: codeine m/z 282, 395; morphine m/z 364, 477; d_3 -codeine m/z 285, 398; d_3 -morphine m/z 367, 480. 6-Acetylmorphine analysis was performed by a previously published method (3).

Long Term Stability Study

Thirteen randomly selected unpreserved urine specimens with 6-acetylmorphine concentrations ranging from less than .01 mg/L to approximately 6.000 mg/L were placed in frozen storage at -17° C immediately following their initial 6-acetylmorphine analysis. After storage at these conditions for over two years, the specimens were removed and reassayed for 6-acetylmorphine. The time difference between initial storage and reanalysis ranged from two years and three months to two years and six months. The initial and subsequent 6-acetylmorphine concentrations were plotted on an XY plot and subjected to linear least squares analysis. The correlation coefficient of the resulting line was utilized to assess the reproducibility of the analysis and the slope of the line was utilized to assess the long term stability of 6-acetylmorphine in the stored specimens.

Specimen Selection for ROC Analysis

The criteria for inclusion in the study required that the specimens screen positive for opiates by Syva EMIT immunoassay at a cutoff concentration of 300 ng/mL (0.300 mg/L) and subsequently demonstrate the presence of codeine and/or morphine by gas chromatography/mass spectrometry at a limit of detection of 5 ng/mL (0.005 mg/L). This meant that specimens containing codeine and morphine from various sources would be included in the study. Specimens selected for ROC analysis also included law enforcement urine specimens that are routinely analyzed by the laboratory to insure that the selected specimens would include some donors that were heroin users (Table 1).

ROC Analysis

One hundred and two opiate positive urine specimens selected by the previously described procedure were subjected to ROC analysis as follows: a two-fold decision criteria, comprised of various total morphine concentrations (TM) and total codeine to total morphine ratios (TC/TM) were evaluated for their ability to predict the presence of 6-acetylmorphine in the specimen with a maximum sensitivity and a maximum specificity. The resulting percentage of correctly identified 6-acetylmorphine positive specimens (true positives) was plotted on the y axis of an XY plot and the percentage of incorrectly identified negative specimens (false positives) was plotted on the x axis. As the decision criteria were varied the plotted points delineated a curve. The statistically best decision criterion is the criterion represented by the inflection point of the curve, that is the criterion represented by the point lying nearest the upper left hand corner of the plot.

This point represents the decision criterion that is the best compromise between true positives (specimens containing 6-acetylmorphine) and false positives (specimens that were predicted to contain 6-acetylmorphine, but did not). Sensitivity, the ability of the criteria to identify specimens containing 6-acetylmorphine, is described by the y position of the point, where the greater the value of y corresponds to greater sensitivity. Specificity, the ability of the criteria to rule out specimens not containing 6-acetylmorphine is described by the x position of the point, where the lower the x value the greater the specificity. Accuracy, a term that includes both sensitivity and specificity, is best described by the percentage of area lying beneath the curve at the chosen point, where the greater the area under the curve the greater the accuracy of the described criteria (4). Besides identifying the statistically most accurate criteria, one can determine from the plot how the sensitivity and specificity vary with changing decision criteria, allowing one to choose the criteria that best suits one's particular needs. That is, weighting the criteria toward sensitivity or specificity, if so desired.

Results and Discussion

Long Term Stability

The stability of 6-acetylmorphine in urine specimens frozen at -17° C for a period of up to 12 weeks has already been evaluated in a previous publication and shown to be adequate (3). However, because the results of this study are intended to be applied to urine specimens that have been previously analyzed for total codeine and total morphine and are being evaluated for further analysis after frozen storage for possibly several months, it is necessary to know if subsequent 6-acetylmorphine analysis is capable of yielding qualitative and quantitative results representative of the specimen at the time of the original total codeine and total morphine analysis. Therefore, the author felt it was important to evaluate the long term stability of 6-acetylmorphine in actual donor urine specimens routinely submitted to the laboratory. The results of the long term stability study yielded a regression line with a coefficient of correlation of 0.999 and a slope of 0.952. The excellent correlation coefficient of 0.999 indicates that the 6-acetylmorphine concentration is quite reproducible for a period of greater than two years. The calculated slope of 0.952, very near 1.00 which would correspond to perfect stability, indicates that 6-acetylmorphine is exceptionally stable for over two years when frozen at -17°C (Fig. 1).

ROC Analysis

Relative operating characteristic analysis, sometimes called receiver operating characteristic analysis, or ROC for short, is a statistical method derived from signal detection theory. The general theory of signal detection is that any diagnostic system is looking for a "signal," however defined, and attempting to ignore or reject "noise." However, no diagnostic system is perfect, therefore at times noise will imitate signal and at times signal will be disguised by noise. In a diagnostic system in which a signal event is considered to be "positive" and a noise-alone event is considered to be "negative," there exists two possible ways for the actual event and the diagnosis to agree, that is "true positive" and "true negative." There are also two possible ways for the actual event and the diagnosis to disagree, that is "false positive" and "false negative." All the diagnostic data from the system being studied can be compiled in a two by two contingency table consisting of the observed frequencies for these four possible outcomes (Table 2). The contingency table tells much about a diagnostic system. However, if one considers the proportions rather than the raw frequencies, one soon discovers that just two proportions contain all the information regarding the diagnostic system. If one takes the first column, for example, whenever a positive event occurs, it is diagnosed by the system as either a true positive or a false negative. If one divides the frequency of true positives by the sum of the true positives and the false negatives, TP/(TP+FN), one obtains

Case #	Total Codeine	Total Morphine	MAM Present	Case #	Total Codeine	Total Morphine	MAM Present
1	0	5	0	52	2783	22553	28
2	490	70	0	53	1853	22654	298
3	750	94	0	54	2700	23580	396
4	4140	129	0	55	1587	24853	47
5	1363	254	0	56	1475	26073	643
6	0	296	0	57	1348	27682	934
7	0	563	0	58	1611	29217	521
8	16261	624	0	59	2014	29450	332
9	91651	625	0	60	1143	32687	54
10	9	634	0	61	1738	32878	832
11	43323	753	0	62	1177	33795	247
12	15	1062	0	63	1418	34413	2613
13	0	1100	0	64	1573	34816	654
14	18186	1182	0	65	1237	34955	2019
15	192	1588	143	66	1860	36064	1101
16	0	2116	0	67	2003	37418	245
17	0	2126	0	68	3648	37692	64
18	2205	2205	89	69	1806	37863	227
19	179	2783	95	70	3500	38716	801
20	48	2932	0	71	1740	40950	639
21	38166	3925	0	72	1397	42936	824
22	83	4125	0	73	11767	43330	269
23	175	4150	417	74 75	4302	44948	626
24	0	4654	0	75	2411	47825	650
25	232	4823	Ō	76	2840	48150	750
26	0	5421	57	77	3788	49935	2827
27	778	6321	130	78	1393	50047	2915
28	501	6345	28	79	3300	52190	37
29	316	6417	489	80	2693	52292	1018
30	40205	6956	0	81	5132	54478	297
31	144	7237	22	82	3302	54569	1021
32	43409	7304	0	83	39170	58620	805
33	373	7957	0	84	3352	59270	997
34	329	8330	187	85	2993	61112	2222
35	518	9657	242	86	4376	64650	1950
36	483	9878	306	87	4038	66666	2211
37	590	10429	34	88	2060	66880	951
38	282	10787	0	89	4750	74743	503
39	542	11457	62	90	2470	76370	1259
40	630	12268	9	91	5902	77332	0
41	527	13527	0	92	5543	85569	2136
42	541	13660	49	93	4225	85842	1615
43	1114	15253	0	94	3326	86647	980
44	837	15809	391	95	7245	89480	135
45	1109	16411	580	96	4591	123897	0
46	694	17439	14	97	5029	126435	1674
47	536	18228	29	98	4870	130540	1279
48	1156	18979	317	99	6050	210490	507
49	975	20007	16	100	9520	223560	5613
50	466	20786	141	101	11560	296750	2685
51	839	22437	392	102	71800	1445800	7700

TABLE 1—Database of analytical results of total codeine, total morphine, and 6-acetylmorphine analyses (ng/mL).

the true positive proportion. Likewise, in the second column, whenever no event occurs, noise-alone, it is diagnosed by the system as either a false positive or a true negative. By dividing the frequency of false positives by the sum of the false positives and the true negatives, FP/(FP+TN), one obtains the false positive proportion. Although the complementary proportions could be calculated, it is the true positive and false positive proportion that is usually of interest in a diagnostic system. The true positive proportion is sometimes referred to as sensitivity and the false positive proportion is 1-specificity. By converting raw frequencies to these two proportions, our measure of accuracy is normalized to a percentage and thus is independent of the relative frequency of an event. That is, the calculated accuracy of a diagnostic system to predict verifiable heroin use from total codeine and total morphine concentrations is independent of the frequency of heroin use in the population (4).

ROC analysis of a system is achieved by changing the characteristics of the system and observing how the true positive and false positive proportions vary. The true positive and false positive proportions tend to vary in the same direction. That is, as one increases the true positive proportion, one also increases the false positive proportion and vice versa (4). However, the true positive and false positive proportions do not necessarily follow each other in a linear fashion. The purpose of ROC analysis is to identify this non-linearity and utilize this information to choose a decision criterion that represents the best compromise between sensitivity and specificity. In the case of the system under discussion, the various decision criteria consisted of a minimum threshold of total

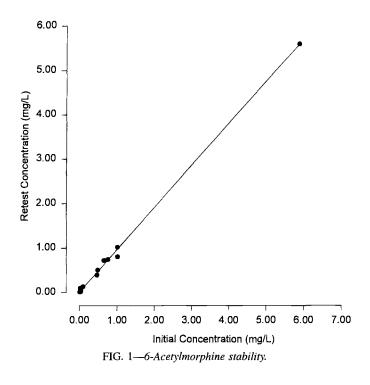


TABLE 2—Two-by-two contingency table for a diagnostic system.

		Event		
		Positive	Negative	
Diagnosis	Positive	True Positive	False Positive	
Diagnosis	Negative	False Negative	True Negative	

morphine, and a maximum ratio of total codeine to total morphine. As the author varied total morphine minimum threshold values and the total codeine to total morphine ratio maximum threshold values the true positive and false positive proportions varied as well (Table 3). By plotting these proportions on an XY plot where

TABLE 3—Calculated ROC results	for the analyzed criteria.
--------------------------------	----------------------------

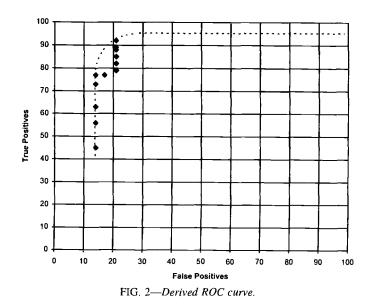
Criteria	ТМ	TC/TM	False Positive Rate	True Positive Rate
1	>5.0 mg/L	< 0.050	14%	45%
2	>5.0 mg/L	< 0.055	14%	56%
3	>5.0 mg/L	< 0.060	14%	63%
4	>5.0 mg/L	< 0.065	14%	73%
5	>5.0 mg/L	< 0.070	14%	77%
6	>5.0 mg/L	< 0.075	17%	77%
7	>5.0 mg/L	< 0.080	21%	79%
8	>5.0 mg/L	< 0.085	21%	82%
9	>5.0 mg/L	< 0.090	21%	82%
10	>5.0 mg/L	< 0.095	21%	85%
11	>5.0 mg/L	< 0.100	21%	88%
12	>5.0 mg/L	< 0.105	21%	88%
13	>5.0 mg/L	<0.110	21%	89%
14	>5.0 mg/L	< 0.115	21%	89%
15	>5.0 mg/L	< 0.120	21%	89%
16	>5.0 mg/L	< 0.125	21%	92%

the X axis is the false positive proportion, or 1-specificity, and the Y axis is the true positive proportion, or sensitivity, one derives a curve (Fig. 2).

Diagnostic Criteria Selection

The ROC analysis was undertaken by first requiring that the total morphine concentration of specimens be greater than or equal to 5.000 mg/L. This value was selected as a starting point for two reasons; to exclude most poppy seed ingestion (5), and the fact that it has become an industry standard in determining which specimens may be specimens from donors using heroin. For example, the College of American Pathologists uses the "greater than or equal to 5000 ng/mL" criteria in the instructions of the UDC proficiency tests to instruct the laboratory whether specimens should be assayed for 6-acetylmorphine (6). Furthermore, preliminary investigation demonstrated that as one increases the 5.000 mg/ L total morphine minimum criteria, the false negative proportion increases rapidly, and as one decreases the 5.000 mg/mL total morphine minimum criteria, the false positive proportion increases rapidly. Therefore, it was determined that the 5.000 mg/L total morphine minimum criterion was the best place to begin ROC analysis. Because codeine use results in the presence of morphine as does heroin use and poppy seed ingestion, a criterion for total codeine concentration must be included as an adjunct to a minimum total morphine threshold to rule out specimens that are positive for morphine due to codeine use. The author hypothesized that utilizing the total codeine to total morphine ratio (TC/TM) would ameliorate any transient concentration anomalies resulting from differing levels of donor hydration and therefore would be more diagnostic than total codeine concentrations alone. The lower the TC/TM the less likely it was that the morphine positive was the result of codeine ingestion. The problem for ROC analysis thus became "In specimens containing total morphine at concentrations greater than 5.000 mg/L, what total codeine to total morphine (TC/ TM) maximum ratio does the statistical analysis indicate to be the best criteria for predicting the presence of 6-acetylmorphine?"

ROC Analysis



Criteria Analysis

A computer program (spreadsheet) was utilized to calculate the true positive and false positive proportions from the database in Table 1 and plot the resulting points as the TC/TM was varied. The most useful curve was obtained when the maximum TC/TM ranged from specimens with TC/MC less than 0.125 to specimens with TC/TM less than 0.05 (Table 3). Figure 2 summarizes the findings of ROC analysis. Although not shown in Fig. 2, at TC/ TM less than 0.05 the true positive proportion or sensitivity suffers because most heroin cases in the area serviced by the author's laboratory, resulted from the use of black tar or Mexican heroin which contains substantial quantities of codeine. Therefore, at the stringent criteria of TC/TM less than 0.05, few specimens would be predicted by the criteria as being from a heroin user, when in fact they contained 6-acetylmorphine. Taken to the extreme the curve approaches the origin as the TC/TM minimum ratio approaches zero. At a maximum TC/TM threshold greater than 0.125 the false positive proportion increases indicating a loss of specificity due to the inclusion of more cases that resulted from codeine use. Taken to the extreme, as the TC/TM increases beyond 0.125 the curve approaches a 95% true positive proportion and a 28% false positive proportion, the proportions that one would achieve by using a greater than 5.000 mg/L threshold alone. In a geographical area, such as the east coast of the United States, where black tar heroin is not prevalent and a purer heroin is commonly used, a more stringent (lower) TC/TM criteria could be used and achieve equal sensitivity and greater specificity.

The criteria that produced the statistically best compromise between sensitivity and specificity for the type of heroin use seen in the author's geographic area was a total morphine concentration greater than 5.000 mg/L and a total codeine to total morphine maximum ratio (TC/TM) less than 0.125. This is represented by the result obtained from the decision criteria nearest the inflection point and therefore is the point that yields the greatest sensitivity while retaining the greatest specificity. At the point described, the true positive proportion or sensitivity is 92%, indicating that the criteria accurately predicted 92% of those specimens containing 6-acetylmorphine to contain 6-acetylmorphine. The false positive proportion is 21% which corresponds to a specificity of 79%, that is, the criterion incorrectly predicted 21% of the specimens containing no demonstrable 6-acetylmorphine to contain 6-acetylmorphine. One can easily calculate the area under the curve at any chosen point and therefore the overall accuracy of the decision criteria represented by that point. The inflection point mentioned above has an x axis location of 21% therefore 79% of the x axis lies under the curve and likewise the inflection point has a y axis location of 92%, therefore 92% of the y axis lies under the curve. The total area under the curve is $79 \times 92 = 7268$. The total area of the graph is $100 \times 100 = 10,000$. Therefore the percentage of the graph under the curve is 72.7% hence the accuracy of the criteria can best be described as 72.7%. Although the previously mentioned criteria gave the best statistical accuracy for the author's data consisting of 102 cases, given enough cases to analyze, the best minimum TC/TM criteria would likely lie somewhere between 0.07 and 0.125, the area of greatest non-linearity. This is indicated by the theoretical ROC curve superimposed on the graph.

On November 16, 1995 the Substance Abuse and Mental Health Services Administration (SAMHSA) proposed that the initial screening cutoff for opiates be raised from 300 ng/mL to 2000 ng/mL and that confirmatory testing include 6-acetylmorphine analysis in order to more adequately target true heroin use and rule out the majority of poppy seed use (6). The data presented herein tends to support that proposal. Because a significant portion of morphine in urine is in the form of morphine glucuronide, which is not very cross reactive in many popular immunoassay screening systems, the proposed 2000 ng/mL (2.000 mg/L) opiate cutoff would correspond quite closely to a total morphine concentration of 5000 ng/mL (5.000 mg/L) in many cases.

Conclusions

6-Acetylmorphine has historically been viewed as an somewhat labile compound, however, as has been demonstrated by this and other studies, at normal urinary pH's between 6 and 8 and storage at -17° C, 6-acetylmorphine is exceptionally stable (3). Furthermore, the author has described a criterion by which to predict whether further toxicological analysis of a morphine positive urine specimen is likely to demonstrate the presence of 6-acetylmorphine and thus verify a suspicion of heroin use. The criteria correctly diagnosed 92% of the specimens known to contain 6-acetylmorphine and incorrectly diagnosed only 21% of the specimens as containing 6-acetylmorphine that in fact did not. This criteria may be of use to toxicologists and medical review officers in determining if analysis for 6-acetylmorphine is warranted or as a model for determining their own criteria for the type of heroin use common to their service areas.

Acknowledgments

The author wishes to thank Dr. William H. Anderson for his collaboration in the development of the 6-acetylmorphine assay used in these analyses, the staff of the Forensic Toxicology Department of Sierra Nevada Laboratories for their aid in the analysis of the specimens included in the study, and Dr. Vina Spiehler for her presentation "Statistical Approaches to Accuracy in Toxicology" presented at the American Academy of Forensic Sciences Annual Meeting in 1993.

References

- 1. Drug Detection Report, Pace Publications, April 20, 1994;20.
- 2. Drug Detection Report, Pace Publications, June 20, 1993;1.
- 3. Fuller DC, Anderson WH. A simplified procedure for the determination of free codeine, free morphine, and 6-acetylmorphine in urine. J Anal Toxicol 1992;16:315-8.
- 4. Swets JA. Measuring the accuracy of diagnostic systems, Science, 1988, June 3;1285–93.
- ElSohly MA, Jones AB. Morphine and codeine in biological fluids: Approaches to source differentiation. Forensic Sci Rev 1989;1:13–22.
- 6. Federal Register, Volume 60, No. 221, 57587-57589.

Additional information and reprint requests: Dwain C. Fuller Sierra Nevada Laboratories Laboratory Corporation of America 888 Willow Street Reno, NV 89502