Survey on Drugs-of-Abuse Testing in the European Union*

Roser Badia, Jordi Segura, Anna Artola, and Rafael de la Torre†
Drug Abuse Research Unit, Institut Municipal d'Investigació Mèdica (IMIM), Autonomous University of Barcelona, Dr. Aiguader 80, E-08003 Barcelona, Spain

Abstract

A survey on the quality of drugs-of-abuse testing in European laboratories was performed in 1993 (Part I) and 1994 (Part II). A total of 195 laboratories participated in Part I and 228 in Part II. There were 154 repeater laboratories. In each part of the survey, six urine samples were sent for analysis under routine conditions. A set of reference materials, including deuterated drug-standard solutions, was provided in Part II of the survey for optimization of analytical procedures. Screening for groups of substances was mainly performed by immunological techniques. Rates of false-positive (FP) results were 0 and 0.5%, respectively, in Parts I and II, and rates of false-negative (FN) results were 7.7% in Part I and 3.4% in Part II. Identification and quantitation of specific substances was mainly carried out by chromatographic methods, particularly gas chromatography coupled to mass spectrometry. Application of chromatographic methods showed a lack of sensitivity (14.9% and 16.7% FN in Parts I and II, respectively) and a lack of specificity (0.8% FP in Part I versus 1.1% in Part II). Repeater laboratories improved their results (17.4% FN in Part I versus 14.2% Part II). The improvement of repeater laboratories emphasizes the need of external quality control programs. The experience is being used as a basis for a recommendation on drug-testing guidelines by a group of European toxicologists.

Introduction

The consumption of drugs of abuse during recent decades is a matter of rising concern to society. Drug consumption can be detected by analyzing urine samples, also known as drug testing. Analytical results help determine if drug use has taken place within a certain length of time before collection of the urine specimen. Drug testing has been progressively extended from the workplace, forensic medicine, prison, insurance, driving, sports, and scientific research. In some cases, positive results may strongly affect the individual's future, especially when associated with disciplinary measures. Laboratories performing tests for nonclinical purposes should have an appropriate background in forensic medicine and internal quality-control programs in place. Because some laboratories may not meet these standards, various measures were proposed to enforce reliability of results.

In Europe, there are several approaches to drug testing, nearly as many as there are countries in the European Union. Consequently, there are different proposals for the enforcement of measures for assuring the quality of results. However, this diversity in policies and practices became a problem when an internal common market and free circulation of workers were suggested in the European Union for the last decade of this century. The most prominent initiatives directed to promote the quality of drug testing have been proficiency-testing programs on drugs-of-abuse urinalyses developed in Italy (1), Spain (2), and the United Kingdom (3).

In 1989, the Council of the European Union showed interest in the evaluation of the extent to and circumstances under which drug testing was developing in Europe, the consequences for individuals with positive tests, and the compatibility of the new situation created by the single European market (4). In this context, several studies in this field have been supported by the Commission of the European Union (5,6). The current report corresponds to a survey, which was developed in two parts during 1993 and 1994, on the quality of analytical results provided by a representative number of European laboratories and is a comprehensive description of the project called "Survey undertaken in the European Community to examine the reliability of urinalyses carried out to detect the use of illicit drugs".

Materials and Methods

Survey design

Laboratories in the European Union member states that perform drug testing participated in a two-step survey (Part I in 1993, Part II in 1994). In Part I of the survey, six samples of sterile...
spiked urine that contained several drugs and/or their metabolites were sent to laboratories for analysis under routine conditions. Participating laboratories were made aware of a list of substances potentially included in the samples. These substances included amphetamines, opiates (morphine related), methadone, cocaine, and cannabinoids. A comprehensive data collection form was requested for return within five weeks of sample receipt. In addition to representative data on preliminary screening, confirmation, and quantitation of drugs, the structure of the data collection questionnaire afforded an opportunity to gain insight into analytical cutoff values currently used, toxicological criteria applied to establish final results, and a number of other analytical aspects.

In Part II of the survey, the set of samples to be analyzed included three spiked urines with different mixtures but concentrations similar to the first step and three real clinical urines. In addition, a set of reference materials, which consisted of several aliquots of freeze-dried urine of known composition, drug standards solutions, and a brochure with a description of the materials provided, including physicochemical characteristics and precautions for safe handling and use, were provided. All laboratories that had participated in Part I of the survey using gas chromatography coupled to mass spectrometry (GC-MS) for quantitation also received a set of deuterated drug standards for optimization of the quantitation procedures.

A group of reference laboratories was in charge of validating sample content. The contents of each sample was disclosed only after the results from all laboratories have been received. A European Toxicological Working Group participated in the survey design (see Acknowledgments).

Participating laboratories

More than 300 laboratories from a list supplied by the European Commission were invited to take part in the study. The list resulted from the previous questionnaire on performances and capabilities of European analytical toxicology laboratories (6). In the first step of the survey, 207 laboratories received samples. In the second step, as many as 242 laboratories received samples. All laboratories participated on a voluntary basis, and no costs were derived for their participation in the study. Laboratories received a code from the European Commission, and their anonymity was preserved during the whole exercise (individual results of participating laboratories were not available to either the European Commission or the other laboratories).

Preparation of samples

For spiked samples, a single pool of human urine from healthy volunteers (selected at the Clinical Trials Unit of the Department of Pharmacology and Toxicology in the IMIM) was used as a matrix for both liquid urine samples and freeze-dried controls. The pool was tested by fluorescence polarization immunoassays (FPIA, TDx instrument, Abbott Laboratories), GC with a nitrogen phosphorus specific detector (GC-NPD), and GC–MS for amphetamines, opiate, methadone, cocaine, cannabinoids, benzodiazepines, barbiturates, propoxyphene and related substances, and other medicines (7-10) and was found to be free of drugs and their metabolites.

High-purity reference standards used for spiking samples were as follows: d-methamphetamine·HCl and morphine-3-O-glucuronide (Sigma Chemical, St. Louis, MO); dl-amphetamine·SO4·H2O, morphine·HCl, codeine·PO4·H2O, and dl-methadone·HCl (Dirección General de Farmacia y Productos Sanitarios, Madrid, Spain); 6-acetylmorphine, dl-2-ethyl-1,5-dimethyl-3,3-diphenylpyrrolinium perchlorate (EDDP·ClO4), benzoylcoecgonine, and eocgonine methyl ester (Radian, Austin, TX); and (−)-11-nor-9-carboxy-Δ9-tetrahydrocannabinol (Research Triangle Institute, NIDA Drug Supply System, Research Triangle Park, NC).

Liquid urine samples were added to combinations of the aforementioned substances at concentrations that attempted to mimic the real intake of drugs. The freeze-drying process of the control samples was performed by the Department of Pharmacology and Therapeutics of the University of Cardiff (Wales, U.K.).

In the case of clinical samples, three samples were obtained.

Table I. Substances Present in the Test Samples

<table>
<thead>
<tr>
<th>Group of drugs</th>
<th>Categorization</th>
<th>Substance</th>
<th>Spiked samples (ng/mL)</th>
<th>Clinical samples (ng/mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Part I</td>
<td>Part II</td>
</tr>
<tr>
<td>Amphetamines</td>
<td>Low</td>
<td>Amphetamine</td>
<td>1076</td>
<td>971</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>Amphetamine</td>
<td>328</td>
<td>286</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Methamphetamine</td>
<td>978</td>
<td>1077</td>
</tr>
<tr>
<td>Opiates</td>
<td>Heroin sample</td>
<td>Total morphine</td>
<td>1664</td>
<td>1496</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total codeine</td>
<td>186</td>
<td>178</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6-Acetylmorphine</td>
<td>426</td>
<td>413</td>
</tr>
<tr>
<td></td>
<td>Codeine sample</td>
<td>Total morphine</td>
<td>307</td>
<td>293</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total codeine</td>
<td>553</td>
<td>502</td>
</tr>
<tr>
<td>Methadone</td>
<td>High</td>
<td>Methadone</td>
<td>977</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EDDP</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>Methadone</td>
<td>490</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EDDP</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Cocaine</td>
<td>High</td>
<td>Benzoylcoecgonine</td>
<td>862</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ecgonine methyl ester</td>
<td>512</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>Benzoylcoecgonine</td>
<td>345</td>
<td>317</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ecgonine methyl ester</td>
<td>180</td>
<td>135</td>
</tr>
<tr>
<td>Cannabinoids</td>
<td>Low</td>
<td>11-nor-9 COOH-Δ9-THC</td>
<td>31</td>
<td>–</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>11-nor-9 COOH-Δ9-THC</td>
<td>122</td>
<td>–</td>
</tr>
</tbody>
</table>

* Mean of results obtained by reference laboratories was calculated excluding the highest and lowest values for each parameter.

* Clinical sample.
One sample belonged to an excretion study of codeine in healthy male volunteers and corresponded to 72-h pooled urine from three volunteers given 30 mg of codeine·PO₄·₃· orally (authorized by the Ethical Committee of Hospital del Mar, Barcelona, and the Spanish Ministry of Health, Madrid, reference 88/135). The remaining two samples were pooled urines from drug-dependent subjects admitted to the Department of Psychiatry of Hospital del Mar for detoxification treatment.

The same methodology currently in use in the Spanish Quality Control Program on Drugs-of-Abuse Testing for sample preparation was applied to the present study (2). In summary, samples were analyzed by immunological, chromatographic, and spectrometric methods as mentioned previously. All liquid samples (including those freeze dried later) were sterilized by filtration (0.2-µm pore size) and re-analyzed before being delivered to the participating laboratories by a courier service less than 48 h after their dispatch.

Samples content
Sterile liquid samples were sent to participating laboratories in glass capped bottles containing approximately 50 mL of urine. Table I shows the concentrations of drugs in the test samples corresponding to Parts I and II of the survey. Samples were classified into "HIGH" or "LOW" according to the overall substance concentration for each group of drugs. In Part II of the survey, an additional set of two human freeze-dried urines was sent to the laboratories in four glass-capped bottles labeled "Reference Sample LOW" and four glass-capped bottles labeled "Reference Sample HIGH" to be reconstituted into 25 mL with distilled water.

Concentrations (equivalent to the free base) in the freeze-dried reference samples based on the gravimetric addition of constituents were 100 ng/mL of dl-amphetamine·SO₄·, dl-methamphetamine·HCl, codeine·PO₄·, dl-methadone·HCl, and benzoylcegonine; 50 ng/mL of 6-acetylmorphine; and 62 ng/mL of morphine-3-O-glucuronide for "Reference Sample LOW" and four glass-capped bottles labeled "Reference Sample HIGH" to be reconstituted into 25 mL with distilled water.

Samples content validation
Reference laboratories were responsible for validating the samples' contents from both qualitative and quantitative points of view. The means of quantitative results reported by the reference laboratories, calculated excluding the highest and the lowest values, were taken as the true concentrations present in the liquid urine samples. Drug concentrations in the freeze-dried reference samples were validated by the coordinating center and by reference laboratories and were found to be present within an acceptable range of at least ±20% of the target value.

Drug standard solutions
In Part II of the survey, laboratories received a set of 11 vials of separated drug standard solutions. There were 10 vials, each of which contained 1 mL of one of the following substances at a 100-µg/mL concentration: dl-amphetamine, dl-methamphetamine, morphine, codeine, dl-methadone, EDDP·ClO₄·, benzoylcegonine, 6-acetylmorphine, eegomine methyl ester, or morphine-3-O-glucuronide 2.5 H₂O. The first six compounds were diluted in methanol, the next two in acetonitrile, and the last in dimethylformamide. Solutions were prepared specifically for the study by Radian. One vial containing 1 mL of (±)-11-nor-9-COOH·Δ⁹·THC at 10 µg/mL in ethanol prepared at IMIM from reference materials provided by the Research Triangle Institute (NIDA Drug Supply System) was also included.

A total of 31 laboratories that participated in Part I of the survey and used GC–MS techniques for quantitation also received a set of 10 vials, each of which contained 1 mL at 10 µg/mL concentration of the following separated deuterated drug standard solutions: dl-amphetamine-d₅, dl-methamphetamine-d₅, morphine-d₁₀, codeine-d₅, dl-methadone-d₃, EDDP-d₃ ClO₄-, benzoylcegonine-d₅, and (±)11-nor-9-COOH·Δ⁹·THC-d₃ diluted in methanol, and 6-acetylmorphine-d₃ and eegomine methyl ester-d₃ diluted in acetonitrile. These solutions were also prepared specially for this study by Radian.

Past results reporting form
The form for reporting results required laboratories to provide information on the following aspects: type of laboratory (forensic, clinical, research, other), type of institution (commercial or non-commercial), analytical results obtained with the test samples when screening for groups of substances, and results obtained in the identification and quantitation of specific substances. Laboratories were asked to report the analytical methodology applied in the screening, identification, and quantitation of each substance or group of substances. A coded list of the techniques most commonly used in drug testing was provided. Laboratories were also asked to provide information on a number of relevant methodologic aspects, routine use and availability of reference substances, and sample preparation procedures. Moreover, information regarding the strategy used to evaluate analytical results, including cutoff values applied when screening for groups of drugs and when identifying specific substances, and usual toxicological criteria for considering positive results were also recorded.

After analytical and toxicological data had been interpreted, laboratories were asked to select one of the following statements to categorize the final result of each sample and for each class of drugs: "positive", "negative", "inconclusive", "analysis not performed", or "would be referred to another laboratory for confirmation". In case of positivity for opiates, the definition of the result being compatible with heroin intake was added in Part I of the survey, and compatibility with codeine intake in addition to heroin intake was added in Part II.

Evaluation of results
The criterion established for a positive or negative finding was the presence or absence of a substance in the urine sample tested, regardless the concentration. In fact, urinary concentrations in terms of response when analyzed by common immunological methods (EMIT, FPIA) were above cutoffs currently in use in most laboratories. A false-negative (FN) test was defined as a result indicating that neither the drug nor its metabolite was present when, in fact, the drug or metabolite was present in the sample. Rates of FN results were calculated as (FN/TP)*100, where FN is the...
number of analyses that originally produced FN results for a given analyte, and true positive (TP) is the total number of analyses for such analyte performed on a sample containing the drug or metabolite. A false-positive (FP) test was defined as a result that indicated that a drug or its metabolite was present when, in fact, the drug or metabolite was absent in the sample. Rates of FP results were calculated as \((FP/TN)\times 100\), where FP is the number of analyses originating FP results and true negatives (TN) is the total number of analyses for such analyte performed in the samples not containing the drug or metabolite.

**Repeater laboratories**

Because some of the concentrations for amphetamines, opiates, and cocaine in spiked urines were the same throughout the study, it was possible to compare the performance of the subgroup of laboratories that participated in both parts of the survey and to assess whether some actions introduced (i.e., supply of reference substances and freeze-dried urines of known content) in Part II contributed to the improvement of performance.

In order to evaluate “final results”, repeater laboratories were classified in two groups: laboratories performing only screening and laboratories performing additional identification of specific substances. For those laboratories performing only screening, it was considered that the “inconclusive” and “referred for confirmation” options were acceptable. For laboratories performing identification of specific substances, the previously mentioned options were considered incorrect results.

**Statistics**

The chi-squared test was used for the comparison of proportions and the Z test for the comparison of means. Statistical significance was set at \(p < 0.05\). The SPSS-PC+ (version 4.0) computer program was used for statistical analysis.

**Results and Discussion**

**Laboratories**

In Part I of the survey, results from 195 laboratories were provided before the deadline, whereas results from 228 laboratories were available in Part II. This represented an active participation of 94% of contacted laboratories at each step of the study. As many as 154 laboratories, homogeneously distributed through the member states, were repeater laboratories.

| Table II. Rates (%) of False-Positive and False-Negative Results Observed for each Analytical Technique when Screening for Groups of Drugs |
|---|---|---|---|---|---|---|---|---|---|---|
| Analytical method type | Rate of use Part I | Rate of use Part II | Specific analytical method | Rate of use Part I | Rate of use Part II | % FN Part I | % FN Part II | % FP Part I | % FP Part II |
| Immunological Agglutination | 1.4 | 4.8 | CEDIA dau | 0 | 0.3 | 0 | 0 | 0 | 0 |
| | | | ONTRAK | 0.7 | 1.4 | 33.3 | 10.8 | 0 | 1.9 |
| | | | ONLINE | 0.7 | 3.1 | 16.7 | 9.1 | 0 | 0 |
| EIA | 59.4 | 56.9 | EMIT dau | 51.2 | 45.0 | 10.3 | 3.2 | 0 | 0.7 |
| | | | EMIT st | 2.1 | 1.7 | 21.0 | 6.5 | 0 | 4.6 |
| | | | EMIT II | 3.7 | 8.3 | 14.3 | 6.5 | 0 | 0.3 |
| | | | Du Pont ACA | 0.6 | 0.8 | 0 | 0 | 0 | 0 |
| | | | Others | 1.8 | 0.9 | 3.1 | 4.3 | 0 | 0 |
| FPIA | 26.6 | 28.4 | Abbott | 26.0 | 28.4 | 3.0 | 1.5 | 0 | 0.1 |
| | | | Merck | 0.6 | 0 | 0 | – | 0 | 0 |
| RIA | 1.8 | 1.0 | DPC | 1.5 | 0.8 | 3.8 | 0 | 0 | 0 |
| | | | Abuscreen | 0.3 | 0.2 | 0 | 0 | 0 | 0 |
| Other | 0 | 1.8 | TRIAGE | 0 | 1.7 | – | 11.9 | – | 1.5 |
| | | | Colorimetric | 0 | 0.1 | – | 0 | – | 0 |
| Chromatographic TLC | 6.8 | 3.6 | Conventional | 1.3 | 1 | 20.8 | 7.7 | 0 | 0 |
| | | | Toxilab | 5.1 | 1.9 | 23.1 | 28.3 | 0.6 | 6.3 |
| | | | HPTLC | 0.4 | 0.6 | 20.0 | 5.6 | 0 | 4.3 |
| HPLC | 0.5 | 0.6 | UV, FLU, EC | 0.5 | 0.6 | 22.2 | 13.3 | 0 | 4.0 |
| GC | 2.6 | 1.4 | NPD, FID, ECD | 2.6 | 1.4 | 2.2 | 0 | 0 | 0 |
| GC-MS | 0.7 | 1.4 | GC-MS | 0.7 | 1.4 | 9.1 | 3.0 | 25.0 | 0 |
Most participating laboratories were clinical (51.3% and 64% of laboratories in Parts I and II, respectively), and a lower number were forensic (32.3% versus 28.1%). The remaining laboratories were included in the categories: "research" (7.2% versus 2.2%) and "other" or "no answer" (9.2% versus 5.7%). In Part II of the survey, there were more clinical laboratories ($X^2, p < 0.007$) and fewer research-oriented laboratories ($X^2, p < 0.03$) than in Part I. Laboratories belonged mainly to noncommercial institutions (85.4% versus 88.2%).

**Qualitative and quantitative analyses**

Qualitative screening of the full menu of drugs of abuse was covered by nearly all participating laboratories. The compliance rate was particularly high for opiates (98% and 99.6% of laboratories in Parts I and II, respectively). Fewer laboratories tested for methadone (81% versus 87%). Overall, more screening analyses were performed in Part II of the survey (95.2%) than in Part I (92.6%).

Figure 1 compares the percentage of laboratories performing identification and quantitation in relation to each specific substance and in both parts of the survey. Nearly 80% of laboratories performed qualitative identification of almost one substance. In Part I of the survey, 62.3% of the total possible identification analyses were done by participating laboratories as compared with 64.5% in Part II. In Part I of the survey, 28.9% of laboratories reported at least one quantitative result as compared with 33.8% in Part II. Overall, more quantitative analyses were performed by laboratories in Part II (25.6%) than in Part I (18%). In general terms, the proportion of laboratories performing identification and quantitation was almost the same for both parts of the survey, but a larger portion of the menu of substances was identified and quantitated in Part II as compared with Part I.

### Table III. Rates (%) of False-Positive and False-Negative Results Observed for the Identification of Specific Substances According to the Analytical Method Used*

<table>
<thead>
<tr>
<th>Analytical method type</th>
<th>Rate of use</th>
<th>Specific method</th>
<th>Rate of use</th>
<th>% FN</th>
<th>% FP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Part I</td>
<td>Part II</td>
<td>Part I</td>
<td>Part II</td>
<td>Part I</td>
</tr>
<tr>
<td><strong>Immunological</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EIA</td>
<td>1.5</td>
<td>0.7</td>
<td>EMIT dau</td>
<td>1.2</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>EMIT II</td>
<td>0.3</td>
<td>0.1</td>
</tr>
<tr>
<td>FPIA</td>
<td>2.1</td>
<td>0.7</td>
<td>Abbott</td>
<td>2.1</td>
<td>0.7</td>
</tr>
<tr>
<td>RIA</td>
<td>0.2</td>
<td>0.6</td>
<td>DPC</td>
<td>0.2</td>
<td>0.6</td>
</tr>
<tr>
<td>Other</td>
<td>0.1</td>
<td>0.8</td>
<td>Other</td>
<td>0.1</td>
<td>0.8</td>
</tr>
<tr>
<td><strong>Chromatographic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLC</td>
<td>29.2</td>
<td>22.3</td>
<td>Conventional</td>
<td>6.9</td>
<td>6.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Toxilab</td>
<td>19.5</td>
<td>11.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>HPTLC</td>
<td>3.0</td>
<td>3.7</td>
</tr>
<tr>
<td>HPLC</td>
<td>7.9</td>
<td>9.2</td>
<td>UV, FLU, EC</td>
<td>7.9</td>
<td>9.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>HPLC-MS</td>
<td>0</td>
<td>0.1</td>
</tr>
<tr>
<td>GC</td>
<td>18.3</td>
<td>16.0</td>
<td>NPD, FID, ECD</td>
<td>18.3</td>
<td>16.0</td>
</tr>
<tr>
<td>GC-MS</td>
<td>40.7</td>
<td>49.7</td>
<td>GC-MS</td>
<td>40.7</td>
<td>49.7</td>
</tr>
</tbody>
</table>

* Results shown in the table refer to those obtained by laboratories as if they were reporting to apply a single analytical method independently of the use of a previous method, if any.
Analytical techniques

The use of either immunoassay or chromatographic techniques was highly dependent on the purpose of the analysis. Screening for classes of drugs was mainly performed by immunological techniques (88.6% of analyses in Part I versus 92.8% in Part II). By contrast, identification and quantitation of specific substances was mainly carried out with chromatographic methods (96.5% versus 98.6% in Parts I and II for drug identification; 83.8% versus 92.1% in Parts I and II, respectively, for quantitation). In Part II of the survey, there was a tendency towards using more immunological techniques for screening purposes and chromatographic techniques for the identification and quantitation of specific substances.

Tables II and III include an extensive list of the analytical procedures used for drug screening and identification. The use of more than one analytical technique was common in the identification of specific substances (i.e., 17.2% of cases in Part II). There was an increase in the number of analyses performed by GC-MS in Part II of the survey. Although thin-layer chromatography (TLC) was the second most frequently applied technique, its use declined. A low but non-negligible number of laboratories reported the use of immunological methods for the identification of specific substances.

With regard to the quantitation of specific substances, GC-MS (57% versus 62.6% in Parts I and II), followed by GC (16%) and high-performance liquid chromatography (HPLC) (11%) were the most common. Immunological methods, especially FPIA, were used frequently. Some laboratories used this approach for substances for which antibody of the immunoassay is supposed to be quite specific (e.g., benzoylecgonine). In other cases, however, quantitative results were reported with immunoassays whose antibodies had a wide range of cross-reactivities. Before the use of immunological methods for quantitation, chromatographic methods were selected for drug identification. In Part II of the survey, a marked decrease in the use of immunological methods for quantitation purposes was found (16% versus 5.5%).

Qualitative analytical results by technique

When screening for groups of drugs (Figure 2), the frequency of FP and FN results was lower when using immunoassays than chromatographic methods, the latter being used only by a few number of laboratories. The rates of FP and FN results for each analytical technique are shown in Table II. When evaluating results for the most commonly used immunological methods, a better performance was observed for laboratories using FPIA analyses. Commercial TLC, the most common chromatographic method used for screening procedures, appeared to have poor performance. The overall rates of FP and FN results when identifying specific substances by chromatographic methods are depicted in Figure 2 (no data for immunological methods are shown because they were used by few laboratories). Table III shows FP and FN results of each technique for the different specific drugs. Interestingly, the FP results of GC-MS in Part I of the survey were well above the mean value for chromatographic techniques.
However, a lower rate of FP results for GC-MS was observed in Part II. Rates of FN results by GC-MS were lower than those found for other chromatographic techniques in spite of being relatively high. TLC had the highest rate of FN results in both parts of the survey.

Qualitative analytical results by drug

When qualitative data for each class of drugs were assessed, no FP results were found for immunoassays in Part I of the survey or for immunological methods when screening for opiates and methadone in Part II. However, FP results occurred in 0.2% and 0.5% in parts I and II, respectively, for amphetamines, and in 1.5% in Part II for cannabinoids. The frequency of FN results was higher in samples with low drug concentration (Table IV). The percentage of FN results for amphetamines showed an increase in Part II of the survey. By contrast, the percentage FN results for cannabinoids showed a statistically significant decrease ($X^2, p < 0.05$). This finding may be explained in part by a larger proportion of laboratories in Part II of the survey using a cutoff of less than 100 ng/mL (68% versus 80%). The rates of FP results for specific substance identification using chromatographic techniques are shown in Table V. A combined evaluation of FP results in both parts of the survey does not reveal a trend of poor performance for any of the drugs tested. Performance in clinical and spiked samples was similar (1.2% versus 1.1% of FP in clinical and spiked samples in Part II, respectively).

Results reported by laboratories performing identification of specific substances showed that the percentage of FN results was higher as drug concentrations to be tested were lower. Table VI shows the percentages of FN for the different concentrations of each specific substance. FN results were especially high for opiates, particularly in clinical samples in Part II of the survey.

Evaluation of final reports

Final results corresponding to both surveys are shown in Table VII. Positive results were provided by laboratories involved in either drug screening or in drug screening plus identification. In some cases, when a positive result was obtained in screening, but no substances were found when performing identification, an unexpected positive final result was given; some laboratories appeared to have more confidence in screening than in identification results.

Taking into consideration the percentages of FN results obtained in the identification of specific substances, a larger percentage of negative results could be expected. Interestingly, some FN results in the identification step had little impact on the final results. Final result rates of error were closer to those observed in screening results than to those observed in the identification of specific substances. Laboratories would generally refer the samples to another laboratory for confirmation when a positive result was found in screening (90% of laboratories chose this option), but no further identification was performed. Some laboratories performing both analytical steps correctly (screening and identification) and other laboratories that obtained contradictory results in screening and identification would also refer the samples to another laboratory for confirmation. Regarding inconclusive results, 22% are from cases were only screening was performed. Another 25% of inconclusive results corresponds to cases in which screening and identification results were in contradiction. The 28% of the cases that resulted in inconclusive results corresponded to cases in which at least one of the substances for which analysis was performed in a given group was not identified (FN). The rest of the inconclusive results (23%) came from laboratories performing both screening and identification of all metabolites.

### Table V. Rates (%) of False-Positive Results Reported for the Identification of each Specific Substance by Chromatographic Methods

<table>
<thead>
<tr>
<th>Substance</th>
<th>Spiked samples</th>
<th>Clinical samples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Part I</td>
<td>Part II</td>
</tr>
<tr>
<td>Amphetamine</td>
<td>1.0</td>
<td>1.7</td>
</tr>
<tr>
<td>Methamphetamine</td>
<td>1.4</td>
<td>4.0</td>
</tr>
<tr>
<td>Morphine</td>
<td>0.5</td>
<td>0.0</td>
</tr>
<tr>
<td>Codeine</td>
<td>0.0</td>
<td>0.8</td>
</tr>
<tr>
<td>6-Acetylmorphine</td>
<td>1.1</td>
<td>1.5</td>
</tr>
<tr>
<td>Methadone</td>
<td>1.2</td>
<td>0.0</td>
</tr>
<tr>
<td>EDDP</td>
<td>-</td>
<td>0.0</td>
</tr>
<tr>
<td>Benzoylcegonine</td>
<td>1.3</td>
<td>0.5</td>
</tr>
<tr>
<td>Egonine methyl ester</td>
<td>0.6</td>
<td>0.8</td>
</tr>
<tr>
<td>11-nor-9-COOH-DTHC</td>
<td>0.8</td>
<td>0.0</td>
</tr>
</tbody>
</table>

### Table VI. Rates (%) of FN Results Obtained when Performing Identification of Specific Substances for each Concentration

<table>
<thead>
<tr>
<th>Approximate concentration (ng/mL)*</th>
<th>Substance</th>
<th>Spiked samples</th>
<th>Clinical samples</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Part I</td>
<td>Part II</td>
</tr>
<tr>
<td>30</td>
<td>11-nor-9-COOH-DTHC</td>
<td>7.4</td>
<td>-</td>
</tr>
<tr>
<td>50</td>
<td>6-Acetylmorphine</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>100</td>
<td>11-nor-9-COOH-DTHC</td>
<td>-</td>
<td>0.8</td>
</tr>
<tr>
<td>180</td>
<td>11-nor-9-COOH-DTHC</td>
<td>9.8</td>
<td>-</td>
</tr>
<tr>
<td>200</td>
<td>Codeine</td>
<td>31.2</td>
<td>30.1</td>
</tr>
<tr>
<td></td>
<td>Egonine methyl ester</td>
<td>34.6</td>
<td>22.4</td>
</tr>
<tr>
<td>300</td>
<td>Amphetamine</td>
<td>17.8</td>
<td>23.7</td>
</tr>
<tr>
<td></td>
<td>Morphine</td>
<td>38.8</td>
<td>34.3</td>
</tr>
<tr>
<td>400</td>
<td>Benzoylcegonine</td>
<td>13.6</td>
<td>16.5</td>
</tr>
<tr>
<td>500</td>
<td>6-Acetylmorphine</td>
<td>18.2</td>
<td>10.8</td>
</tr>
<tr>
<td></td>
<td>Methadone</td>
<td>6.7</td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td>Egonine methyl ester</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>600</td>
<td>Codeine</td>
<td>17.7</td>
<td>9.2</td>
</tr>
<tr>
<td></td>
<td>Egonine methyl ester</td>
<td>20.4</td>
<td>-</td>
</tr>
<tr>
<td>1000</td>
<td>EDDP</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Amphetamine</td>
<td>8.1</td>
<td>4.9</td>
</tr>
<tr>
<td></td>
<td>Methamphetamine</td>
<td>8.3</td>
<td>4.2</td>
</tr>
<tr>
<td></td>
<td>Morphine</td>
<td>7.2</td>
<td>7.4</td>
</tr>
<tr>
<td></td>
<td>Methadone</td>
<td>1.4</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Benzoylcegonine</td>
<td>5.1</td>
<td>-</td>
</tr>
<tr>
<td>1300</td>
<td>Methadone</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1500</td>
<td>Codeine</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2600</td>
<td>EDDP</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3000</td>
<td>Benzoylcegonine</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5200</td>
<td>Morphine</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

* See Table I for specific concentrations.

† Clinical sample.
(5%) or the main metabolites (18%) correctly.

It should be noted that drugs in the opiate group were the only substances included in the survey for which, in addition to analytical data, toxicologic criteria had to be applied for the interpretation of results.

Repeater laboratories

Almost 15% of screening analyses were performed by a different analytical method in Part II of the survey as compared with Part I. This was particularly noticeable with the use of FPIA and agglutination techniques instead of EIA and a decrease in the use of chromatographic techniques, especially TLC. In respect to identification procedures, a low rate of change (nearly 5%) was observed. This corresponded to a decrease in the use of immunological methods and TLC on the benefit of GC and GC-MS. Twenty-eight percent of laboratories changed the analytical method used for drug quantitation, that is, there was a decrease in the number of immunoassay tests and an increase in the number of GC-MS tests.

The number of centers performing drug identification tests decreased slightly (85.7% in Part I versus 83.1% in Part II), although there was an overall increase in the number of tests carried out for identification purposes in relation to a larger number of substances per laboratory assayed. Results reported for methamphetamine and ecgonine methyl ester showed a statistically significant increase ($X^2, p < 0.05$). The number of laboratories reporting quantitative results raised from 36% in Part I to 42% in Part II. Moreover, a larger number of drugs per laboratory were quantitated.

When results of analyses performed in both parts of the survey for samples with common content, there was no improvement data reported for screening. It should be noted, however, that results obtained in Part I of the survey were particularly successful. In Part I of the survey, no FP results were encountered, and the percentage of FN results was 3% (95% confidence interval, 1.9 to 4.6%). In Part II, the corresponding figures were 0.5 and 4.9% (95% confidence interval, 3.5 to 6.8%). No statistically significant differences were found between the percentages of both FN and FP results of drug identification by chromatography between Parts I and II of the survey. The percentage of FP in Part I was 0.7% (95% confidence interval, 0.38 to 1.07%) and 1% in Part II (95% confidence interval, 0.5 to 1.8%). The rates of FN results were 17.4% in Part I of the survey (95% confidence interval, 15 to 19.7%) and 14.2% in Part II (95% confidence interval, 12.9 to 16.3%). With regard to the identification of each particular analyte (Figure 3), differences in FN results between Parts I and II showed substantial improvements for ecgonine methyl ester and statistically significant improvements for codeine high and morphine low ($X^2, p < 0.05$).

With regard to final reports (Table VIII), laboratories performing screening and identification of substances, for opiates and cocaine in particular, showed better performance. The apparently poorer performance for final results of amphetamines in Part II is partially due to the use of 1000 ng/mL instead of 300 ng/mL as the cutoff value for screening purposes. The rate of correct “final results” for laboratories performing screening for sample compatibility with codeine ingestion was very low because opiate immunological screening procedures have broad cross-reactivities that prevent the distinction between codeine and heroin intake; however, a large proportion of laboratories reported these samples as compatible with a heroin intake.

With respect to laboratories performing identification of specific substances, a significant improvement was observed for the urine that mimicked codeine ingestion. There was a general trend towards improvement because there was a reduction in the number of results reported as inconclusive (6.6% versus 4.5% in parts I and II), whereas the number of results reported as confirmation was almost the same (5.1% versus 4.8%).

<table>
<thead>
<tr>
<th>Table VII. Final Results Released by Laboratories for Each Sample</th>
<th></th>
<th>Survey Part I</th>
<th></th>
<th>Survey Part II</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Categorization*</td>
<td>positive</td>
<td>negative</td>
<td>inconclusive</td>
</tr>
<tr>
<td>Amphetamines</td>
<td>Low</td>
<td>86.5</td>
<td>6.4</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>90.9</td>
<td>1.6</td>
<td>2.7</td>
</tr>
<tr>
<td>Opiates</td>
<td>Codeine sample</td>
<td>35.8</td>
<td>37.9</td>
<td>19.5</td>
</tr>
<tr>
<td></td>
<td>Heroin sample</td>
<td>85.3</td>
<td>3.4</td>
<td>5.3</td>
</tr>
<tr>
<td>Methadone</td>
<td>Low</td>
<td>90.5</td>
<td>6.5</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>95.3</td>
<td>1.8</td>
<td>0.6</td>
</tr>
<tr>
<td>Cocaine</td>
<td>Low</td>
<td>75.8</td>
<td>12.4</td>
<td>7.0</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>88.2</td>
<td>3.7</td>
<td>2.1</td>
</tr>
<tr>
<td>Cannabinoids</td>
<td>Low</td>
<td>69.8</td>
<td>22.9</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>67.6</td>
<td>25.7</td>
<td>2.2</td>
</tr>
</tbody>
</table>

* See Table I for specific concentrations.
† Ref. conf. = sample would be referred for confirmation to another laboratory.
Conclusion

The number of actively participating laboratories in both parts of the survey (94%) was high and reflected the interest for this type of study.

When screening for groups of substances, immunological techniques appear to have a much better reliability than the chromatographic techniques. Immunological techniques results obtained in Part II were probably closer to their real performance (11) as compared with Part I, probably because of the inclusion of clinical samples in Part II of the survey.

Because specific regulations and a consensus between European laboratories on cutoff or threshold concentrations for reporting positive results have not been developed, it should be taken into account when evaluating qualitative analytical results obtained in this survey that the presence of a substance in the test urines, regardless of its concentration, was considered as a criterion for positivity. Accordingly, some FN results may be due to the use of a cutoff value higher than the actual analyte concentration present for reporting positive results or an inadequacy of the cutoff value selected (relatively high) and not to inappropriateness of the technique itself. It should be noted that there is a great variety of cutoff levels for the same compound in Europe, which may account for differences in the interpretation of results with important consequences for individuals.

Application of chromatographic analytical techniques for the identification of specific substances was shown to be problematic for some laboratories. These results should be toned down by the fact that, except for opiates, many identification FN results are not translated in the final result released by laboratories. For example, the rate of FN results for ecgonine methyl ester does not match the high rate of correct "final results" released by most laboratories for cocaine because many of them identified benzoylecgonine correctly. This high rate of FN results cannot be generalized to the whole population of laboratories. Although there were participating laboratories with a very high rate of FN results, a non-negligible proportion (about 20%) was free of errors in the identification of substances. Repeater laboratories slightly improved their results.

The changes observed between Parts I and II of the survey in the type of analytical methodologies used seem to indicate that the technology used by laboratories for drug testing is not consolidated. Commercial pressures (e.g., particularly for immunological screening methods) and an attempt to correct past errors in some aspects of drug testing (e.g., especially those regarding quantitation) may explain these changes.

The assessment of final results shows that laboratories performing only screening tests have a lower performance than laboratories carrying out the identification of specific substances. This observation is particularly true in the case of the codeine samples. Despite an improvement in Part II, there are still important problems related to the interpretation results for opiates.

The provision of reference materials was followed by an increase in the menu of substances identified and in the number of laboratories reporting quantitative results and improved the performance of laboratories for their identification. Availability of reference substances should be mandatory in any laboratory performing drug testing routinely. The improvement of the repeater laboratories emphasizes the need for the implementation of periodic external quality control programs for drugs-of-abuse testing.

Except for the issue of the interpretation of results, some of the questions addressed in this study on drugs-of-abuse testing in the EU were already covered by pilot studies conducted in the United States in the 1980s (12,13) before the development and implementation of NIDA guidelines. Results obtained by European laboratories in the present survey study are, in general terms, better than those observed in the older studies. This is probably due to the experience accumulated in drug

![Table VIII. Rates (%) of Final Correct Results Released in Positive Samples by Repeater Laboratories Screening Analytical Step and by Laboratories Performing Identification of Specific Substances](image)

<table>
<thead>
<tr>
<th>Groups of substances</th>
<th>Categorization*</th>
<th>Only Screening</th>
<th>Screening and Identification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Part I</td>
<td>Part II</td>
<td>Part I</td>
</tr>
<tr>
<td>Amphetamines</td>
<td>Low</td>
<td>95.5</td>
<td>83.3</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>100</td>
<td>95.8</td>
</tr>
<tr>
<td>Opiates (compatible with heroin intake)</td>
<td>Heroin sample</td>
<td>90.9</td>
<td>96.3</td>
</tr>
<tr>
<td></td>
<td>Codeine sample</td>
<td>45.5</td>
<td>44.4</td>
</tr>
<tr>
<td>Cocaine</td>
<td>Low</td>
<td>72.0</td>
<td>81.3</td>
</tr>
</tbody>
</table>

*See Table I for specific concentrations.

![Figure 3. Rates of errors performed by repeater laboratories when identifying each specific substance at the same concentrations in Parts I and II of the survey.](image)
testing in the past 10 years, the indirect influence of standards set by the NIDA guidelines, and to other recommendations issued by the scientific community (14,15). However, the present results show that areas for improvement still exist for European laboratories. In addition, some specific problems have been identified that should guide future programs in the EU. As a first step and based mainly in the present results, a specific recommendation document has been prepared by a group of European toxicological experts (16) for urine testing for drugs of abuse.

Acknowledgments

These studies were supported both by the DG VII/1 (Directorate General V, Employment, Industrial Relations and Social Affairs. Public health and safety at work. Implementation of disease-specific programmes and actions) of the European Commission and by the Institut Municipal d’Investigació Médica (IMIM), Barcelona, Spain (Coordinating Center of the survey).

We are indebted to M.T. Van der Venne and J.C. Berger (Commission of the European Union); Ch. Gilliard and F. Parmentier (Belgium); K. Worm (Denmark); P. Lafargue and P. Mangin (France); M.R. Möller and R.K. Müller (Germany); H. Tsoukalis (Greece); A. Pierce (Ireland); U. Avico, E. Sternieri, and S.D. Ferrara (Italy); R. Wenning (Luxembourg); R.A. de Zeeuw and I.C. Dijkhuis (The Netherlands); D.A. Carrondo Tome dos Reis (Portugal); and M.D. Osselton and J. Williams (U.K.), who are members of the European Toxicological Working Group, for their valuable assistance in the design of the study. The following laboratories are also acknowledged for their support as reference laboratories: Institute of Toxicology (Oslo, Norway), Institute of Legal Medicine (Padova, Italy), National Poisons Unit (London, UK), Psychiatric Diagnostic Laboratories America (New Jersey), Smithkline Beecham Laboratory (Atlanta, GA), Foothill Hospital (Calgary, AB, Canada), Institute für Rechtsmedizin (Homburg/Saar, Germany), and National Laboratory of Forensic Chemistry (Linköping, Sweden). We are grateful to the Dirección General de Farmacia y Productos Sanitarios (Spain) and to the NIDA Drug Supply System for their supply of clinical urines, to Alicia Redón, Marta Carnicero, and Marisa González for excellent technical assistance, to Javier Morano for computer work and Marco Pavesi for statistical analysis of data, to Rosa Herrera for secretarial assistance, and to Marta Pulido for editing the manuscript and editorial assistance.

References

5. Commission of the European Community questionnaire on the legal context and governmental regulations of the testing body fluids to detect the use of illicit drugs in the EU Member States, Luxembourg, 1993.

Manuscript received March 20, 1997; revision received September 9, 1997.