Analytical and Technical Aspects of Testing for Drug Abuse: Confirmatory Procedures

Michael A. Peat

Many laboratories are now performing urine drug testing for employers, governmental agencies, and other institutions. It is now recognized that presumptive positive screening results have to be confirmed by an analytical procedure based on a different chemical technique with greater than or equal sensitivity to the screening test. Thin-layer chromatography has been widely used for this; however, it is relatively insensitive for certain drugs, and it cannot satisfy the accuracy and precision requirements needed to determine threshold concentrations reliably. Gas chromatography–mass spectrometry is able to satisfy these threshold requirements and has become the method of choice for confirming initial immunoassay results.

Today's mass drug-screening programs have their historical roots in the methadone maintenance programs and in the analytical procedures used for screening in the emergency room. Each of these areas has had its own requirements. For example, because emergency–room screening requires a rapid turnaround time, simple and reliable immunoassay procedures are used there. Methadone maintenance programs have used thin-layer chromatographic (TLC) procedures because of the wide range of the drugs that can be detected. Historically, confirmatory procedures have not been widely used in such arenas of drug testing.

With the advent of large-scale workplace drug testing, and remembering the early experiences of the Department of Defense drug-testing program, it is now recognized that drug testing is a "forensic toxicology" analysis. This requires a strict chain of custody, security of the laboratory, appropriate and adequate quality-control procedures, and the use of legally defensible analytical procedures. These requirements are necessary because of the potential legal challenges to the laboratory results and the implications of laboratory results for job applicants, employees, and employers. Drug testing of urine for purely clinical purposes is different in that it is used in conjunction with other medical tests to assist in diagnoses.

It is useful at this stage to remember some basic definitions:

Screen: a series of initial tests designed to distinguish negatives from presumptive positive samples.

Confirmation: a second test used to positively identify a drug or drug metabolite.

With today's technology it has also become widely accepted that the confirmation test be based on a different chemical technique with sensitivity at least equal to that of the screening test. Because of this requirement it is inappropriate and unacceptable to perform dual immunoassays; e.g., radioimmunoassay or fluorescence polarization immunoassay cannot be used to confirm an initial enzyme immunoassay result. It is possible that scientific arguments can be made about the specificity of monoclonal antibodies, but because of the public controversy surrounding drug testing, their use would be viewed as unacceptable in legal proceedings.

Two recent articles have focused on different aspects of drug testing. The first, by Hoyt et al. (1), rated the legal defensibility of methods used for drug testing. These methods were rated by scientific experts and arbitrators. The experts generally agreed that two different procedures are required and that immunoassay followed by gas chromatography–mass spectrometry (GC-MS) is rated as fully defensible against legal challenge. In this survey, GC-MS was defined as a procedure involving electron-impact ionization, three-ion monitoring, and ion ratioing. The combination of TLC and GC-MS was also rated very highly. The use of two immunoassays was viewed as a combination that would be difficult to defend in legal challenges. When the experts were asked to rate the single most defensible technique, they chose GC-MS, particularly in the full-scan mode. However, several of them emphasized the need for a second test to ensure that an administrative mistake had not been made in aliquoting the sample for the first analysis.

When the arbitrators were surveyed, comparable uniformity in responses was not seen. Single-procedure and multiple-procedure methods were rated almost equally, although the arbitrators recognized that accurate test results were important in their cases. Obviously other factors such as chain of custody and proficiency testing are important, and these are recognized in the recent Department of Health and Human Services Scientific and Technical Guidelines for Federal Drug Testing Programs (2).

The second article, by Frings et al. (3), was the result of a study by the American Association for Clinical Chemistry to determine the status of drug-abuse testing. Although the data obtained from this study are encouraging, in that the false-positive rate was <0.01% and the false-negative rate was <0.1%, they may be misleading as to the true state of drug testing today. Firstly, there was a "selection" of laboratories, in that laboratories performing both screening and confirmation tests were asked to participate. In addition, the laboratories were current subscribers to the AACC's "Toxicology Survey Plus" quality-control program. Secondly, the concentration of drugs and metabolites used as challenges were significantly higher than the currently suggested thresholds for mass-screening programs. For example, benzylcgonine was included in two of the samples at 1300 and 2000 ng/mL, whereas the widely used immunoassay thresh-

---

Ed. note: Dr. Peat is Vice-President of Clinical and Forensic Services, ChemWest Analytical Laboratories, Inc., Sacramento, CA, and Director of Toxicology, CompuChem Corp., Research Triangle Park, NC. This paper is adapted from a presentation made at the joint AACC/CAP conference on drug testing, Chicago, October 1–2, 1987.
old is 300 ng/mL and that required for confirmation is either 300 or 150 ng/mL. The same criticism could be leveled at the other analytes with the possible exception of the carboxylic acid metabolite of tetrahydrocannabinol.

Some other interesting points also arose from the AACC study. Thirty-four percent of the laboratories said that they used quantitative techniques for their confirmations. Of these, 32% confirmed their cutoff daily, consistent with good quality-control procedures. In contrast, 15% verified the cutoff only when the method was put into place; this is obviously not acceptable. Enzyme immunoassay was the most common screening procedure, and it appears that 60% of the laboratories used either GC-MS or TLC for confirming presumptive positive results. What did the other 40% use?

In a pilot study performed by the National Institute on Drug Abuse (4), in which 50 volunteer laboratories were selected, the rate of false positives observed was similar to that reported by Frings et al. (3). However, a much higher rate of false negatives was found—approximately 20%—which, interestingly, increased to almost 30% when "blind" proficiency test samples were sent.

If you will bear these comments in mind, I would like to discuss analytical techniques that are used for confirmatory testing. These techniques are almost universally based on chromatography, a means of separating components of a mixture based upon its physical-chemical properties. It is useful to develop a simple analogy to explain this, and other analytical techniques, to non-scientifically trained clients, to arbitrators, and to courts. There is no quicker way of losing the attention of a jury or an arbitrator than by using complicated scientific explanations for analytical techniques.

Various chromatographic techniques are available, and used, for confirmations. These include gas-liquid chromatography with a variety of detectors—e.g., nitrogen-phosphorus detection and mass spectrometers (GC-MS)—"high-performance" liquid chromatography, and TLC. For two reasons I will concentrate here on two techniques, TLC and GC-MS: (a) they are now the most widely used, and (b) GC-MS is required in the Federal Drug Testing and Department of Defense programs. I believe that the latter will become the required confirmatory technique in all drug-testing programs related to the workplace. My reasons are as follows.

Firstly, what are the strength and weaknesses of TLC? Two of its major advantages are the ability to detect a wide range of drugs and metabolites and the fact that no expensive instrumentation is required. However, it also suffers from several weaknesses. It is relatively insensitive: for a number of drugs and metabolites, detection limits are in the milligram per liter rather than the microgram per liter range required by today's drug-testing programs. Moreover, interpretation of TLC plates is subjective, in that it is based on experience and pattern recognition. Certainly, in the right hands (or I should say, with the right eyes) it is a valuable procedure in emergency-room screening and methadone maintenance programs. Finally, as used by most laboratories today, it is a non-quantitative technique and therefore it is difficult accurately to establish a pre-designed cutoff and to precisely satisfy this cutoff from batch to batch of samples assayed. Threshold concentration values are widely used in drug-abuse testing (2), and it is imperative that a laboratory can analytically satisfy these thresholds. For a particular analyte, urine samples that contain 75% of the threshold concentration must be differentiated from those containing the threshold concentration or greater. As used in most laboratories, TLC cannot satisfy this important requirement.

Now, what are the advantages and weaknesses of GC-MS? The advantages are that it is specific: it will accurately identify a drug or drug metabolite, either by full-scan or selected-ion monitoring techniques. It is also sensitive, in that <1 μg of certain drugs per liter can be reliably detected by electron-impact and chemical-ionization procedures. It is also an accurate and precise quantitative technique when performed correctly. The use of internal standards allows the analyst to define which samples contain a drug or metabolite in concentration(s) exceeding the pre-designated threshold.

Some of the perceived weaknesses of GC-MS are overrated, considering today's technology. First, I will address the belief that it is labor intensive. Any chromatographic procedure requires extraction of the drug or metabolite and, if GC is to be used, derivatization is often required. This makes the procedure time consuming relative to immunoassays and to batch TLC, but the time per injection depends on the analyte or analytes, the data system, and the calculations required. There is a need for expertise in capillary chromatography and at least some exposure of some of the laboratory staff to mass spectrometry. However, with the software available with most of the GC-MS instruments, programs can be written that perform calibration, quantification, determination of relative retention times, and calculation of ion ratios. These allow the use of GC-MS in high-volume drug-testing laboratories with either autosamplers or untrained personnel. Nevertheless, I would like to stress that the purchase of a mass-spectrometric detector for a GC and the ability to perform, for example, two assays does not make that laboratory a reference laboratory in mass spectrometry!

Another weakness of GC-MS that is often quoted is the initial cost of the instrumentation and its maintenance—certainly not an inconsequential expense. The development of bench-top instruments has lowered the cost considerably during the past decade, and instruments can now be purchased for less than $100,000. Although this is expensive for some laboratories, it is the investment that will be required for a laboratory to continue to perform drug testing. A maintenance contract for such an instrument may cost as much as $10,000 per year.

Several factors are important in choosing a GC-MS procedure. Among these are the extraction procedure, the derivative to be prepared, the ionization method, and the chromatographic and mass-spectrometric conditions. Although chemical ionization has advantages in terms of sensitivity and, under the correct circumstances, is as specific as electron impact, it is not widely used in drug testing. It may however, be applicable to certain analyses, such as that for LSD and its metabolites (5). The procedure most widely used today is that of electron-impact ionization and three-ion monitoring. This technique requires that the ratios of the ion intensities in the unknown must match, within ± 20%, the corresponding ratios in an extracted standard. Note that the ion intensities, and therefore ratios, will vary depending on the amount of drug or metabolite (or derivative) injected onto the column, and these may not be consistent with the extracted standard when the drug or metabolite concentrations are either much higher or lower than those in the standard. The other important identification criterion is either absolute or relative retention time. Full-scan tech-
niques are used by some laboratories, and these are certainly satisfactory. However, in certain cases they suffer from disadvantages of lower sensitivities than selected-ion monitoring, interference from co-eluting compounds, and the lack of automation.

Another important factor in selecting a GC-MS procedure is the choice of internal standard. Ideally, a deuterated analog of the drug or metabolite to be analyzed should be used. Until recently these were difficult to obtain, but several are now available from commercial sources. Before using these, however, the laboratory should determine what the contribution of the nondeuterated form may be to the quantitative accuracy. For example, some of the deuterated analogs have been shown to contain >5% nondeuterated form as an impurity. Ideally, they should contain <1%.

What factors are important in choosing a derivative? Obviously, it should be easily prepared and available in high yield. It should be stable, with good chromatographic and mass-spectrometric characteristics. In particular, it should have three ions, under electron-impact ionization, with m/z values relatively free of interference and with mass separation between the nondeuterated and deuterated compounds. It is also beneficial if the ions are relatively similar in mass, because this will result in greater stability of ion ratios.

Obviously, I cannot describe here in detail all the derivatives that have been formed, and I would refer the interested reader to references 6–12 for more information. Of the five drugs listed in the Department of Health and Human Services guidelines, phencyclidine is the only one that does not require derivatization.

There is a belief that “GC-MS is 100% accurate.” Certainly, when performed correctly, it will result in the positive identification of a drug or metabolite. I emphasize that it will be "legally defensible" when appropriately quality controlled. This quality-control program should include the analysis of standards, blind controls, and drug-free sample with each batch of samples. The standards should include at least one below, one at, and one above the threshold value. In addition to this quality-control program, participation in an appropriate proficiency-testing program is essential.

What about the future techniques for confirmatory analysis? Some drugs may not be suitable for GC-MS identification or, if they are, electron-impact ionization may not be the ideal ionization technique for them. For example, many dimethyl tertiary amines have base peaks at 58 m/z with very little other fragmentation. In these situations, chemical ionization may be more suitable, together with the use of high-resolution capillary columns.

I believe that, as analytical toxicologists, we should approach the future with flexibility and not restrict ourselves to GC-EIMS. The development of lower-cost chemical-ionization instruments and the application of tandem mass spectrometry (MS-MS) and super-critical fluid chromatography to analytical toxicology should be encouraged and applied to confirmatory analysis in drug-abuse testing.

References
2. Department of Health and Human Services, Alcohol, Drug Abuse and Mental Health Administration. Scientific and technical guidelines for federal drug testing programs; standards for certification of laboratories engaged in urine drug testing for federal agencies; notice of proposed guidelines.
4. RL Hawks. Personal communication.

1 Radian, MSD Isotopes, and Research Triangle Institute.