Pharmacology Review: Testing for Drugs in Horses

Thomas Tobin, D.V.M., Ph.D.

From the Kentucky Equine Drug Research Program, Department of Veterinary Science, University of Kentucky, Lexington, KY 40506.

The testing of body fluids from horses for the presence of drugs was first introduced into racing in Europe about 1910. In the year 1912, a horse called Bourbon Rose won the gold cup at Maisons Lafitte in France, but was disqualified because it yielded the first "positive" dope test. Following a pattern, which has since become all too familiar, the owners sued, but lost their case. Routine testing for drugs in horses had arrived and had been supported in the courts.

Although equine drug testing has been with us now for more than 70 years, most veterinarians know relatively little about the testing process and how an analyst arrives at the conclusion that he has a "positive." Consequently, most veterinarians are not likely to be in a position to even begin to evaluate an analyst's evidence or advise a client in the event of a "positive" being called. This review is to outline the process of drug testing as it is currently conducted and to help the veterinarian in assessing the significance of the kind of evidence on which an analyst may "call" a "positive."

Before we outline the various methods used by analysts, it should be clearly understood that analytical methods which are quite acceptable in experimental work may be of very limited value in forensic work. In an experimental situation, where one administers a known drug to a horse, relatively simple analytical methods will often suffice.[9] This is because these horses can be tested before and after the experiment to insure that they are not yielding "false positives." In the forensic situation, however, no control (i.e. predrug or postdrug) samples are available from the horse, and samples are being drawn from a large number of horses under widely differing conditions. These horses will have been treated with many different drugs, feed additives, and domestic remedies, and exposed to many different varieties of plant life, and may have been exposed to any one of the approximately 4.2 million known chemical entities.[3a] On scientific grounds, therefore, scrupulous care must be taken in any forensic test merely to insure accuracy, quite apart from the consideration that personal and professional reputations and livelihoods are usually involved. Because of these considerations, drug "positives" should be "called" only on the strongest scientific grounds. The purpose of this article is to help the veterinarian recognize these grounds and to differentiate between excellent and perhaps marginal analytical work. In writing this article I have drawn heavily on a review of the subject by Dr. Bruce Stein[6] and his co-workers in *The Wisconsin Law Review*.

On the assumption that the biological samples in question were properly taken, numbered, sealed with evidence tape, and delivered to the laboratory, their receipt in the laboratory should be recorded by identifying numbers in a bound laboratory log dated for that day. The log should record the approximate volume received, any comments on their appearance, and the pH of a urine sample. The laboratory should also have an up-to-date loose-leaf manual on the testing procedures in use on that particular day, as methods in all laboratories change with time, and sometimes rapidly. In addition, the laboratory should have clear-cut written criteria set up in advance to be adhered to for differentiating between data points. For example, in thin layer chromatography work by Sunshine, *et al.*,[7] all Rf's differing by more than +0.05 were considered distinguishable. Similar criteria should also be established for GC retention times and other quantative data.

If the procedure of the International Olympic Committee is followed, as it is by many European laboratories, the samples arrive in two bottles labelled "A" and "B." The analysis is started on the A sample and the B sample is stored frozen. In the event of a "positive," the B sample is then available for confirmatory analysis either in the presence of a referee or by an independent laboratory. This procedure safeguards the interests of the owners, and in many European laboratories analysis is not started in the absence of a referee's sample.[1] This system has proven its worth in Europe, because at least one "positive" called this spring (1978) was not confirmed by an unquestionably competent refereeing laboratory in another country. Unfortunately, in the United States many jurisdictions do not require a referee sample, and it is all too easy to dispose of the remains of a test sample.

[a] There are about 4,000 active drug ingredients, 63,000 chemicals in everyday use, and, as of November 1977, about 4,000,000 chemical entities, a number which was increasing at the rate of about 6,000/week.[3]

The Testing Procedure

The first step in drug analysis is usually the taking from the sample two aliquots of about 10 to 20 ml. One of these aliquots is made acidic with about 5 ml of saturated acid phosphate buffer, and one is made basic with a saturated solution of sodium tetraborate. Then an equal volume of an organic solvent, such as benzene or dichloromethane, is added to each system and the whole shaken up together. When the organic solvent settles out, it is found that the acidic drugs will have migrated into the organic solvent under acidic conditions. In this way, acidic drugs such as phenylbutazone, furosemide, naproxen, and so forth will predominate in the acidic extract, while basic drugs such as amphetamine, ritalin, cocaine, and the local anaesthetics will predominate in the basic extract.[4] This process is technically called solvent extraction. Other drugs, such as caffeine and reserpine, which do not carry any dominant electrical charge, are not easy to separate from other drugs at this stage (Figure 1).

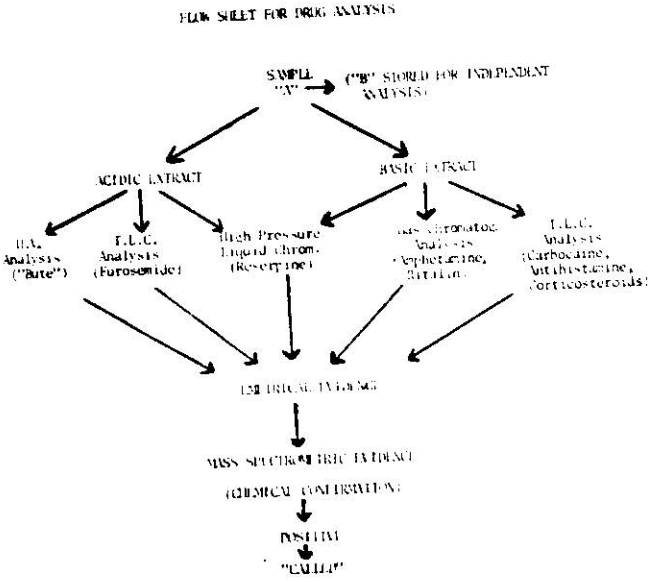

Figure 1. General flow sheet for drug analysis.

Once the solvent extraction has been made, the solvent is evaporated down to a very small volume to concentrate the extracts. A small portion of the acidic extract is then spotted on a "thin layer plate" and chromatographed by a process called Thin Layer Chromatography (TLC). Other common manuevers with acidic extracts include Ultraviolet Spectrometry (UV analysis) and, increasingly, High Performance Liquid Chromatography (HPLC). If positive indications for the presence of a drug appear in any or all of these test systems, the urine extract may be examined by Gas Chromatography-Mass Spectrometry (GCMS).[4]

The same general procedure is followed with the basic extract. While a portion of the extract may be subjected to thin layer chromatography, an increasingly common procedure is to react the extract with a highly reactive "marker" chemical such as pentafluropropionic anhydride (PFA) and examine the result by gas chromatography. This method is useful for picking up very small quantities of central nervous system stimulants.[2] Again, if suspicious peaks are found, the material may be subjected to gas chromatography-mass spectrometry. A general flow sheet for drug analysis is presented schematically in Figure 1.

Horses excrete some drugs such as apomorphine and pentazocine in their urine by linking these drugs to highly water-soluble sugar molecules. Before analysis for these drugs can commence, the drug portion of this complex must be split out by enzymatic hydrolysis.[10] Once this process, which takes two to three hours, is complete, testing for these drugs proceeds by acidic or basic extraction as previously.

The Analytical Methods

The essence of all these procedures is that they give varying kinds of evidence about the presence of a drug. At the first sign of something suggesting the presence of a drug, the analyst has what he calls a "suspicious sample." As he continues to gather evidence, he must ask himself how good or useful the evidence is and at what point he should call it a "positive."[b] Before we can second-guess him on this judgement, we will have to look more closely at the analytical techniques he uses and the quality of information that has been provided him.

Ultra Violet Spectrometry (UV Spectrum)

In ultraviolet spectrometry, light of shorter and shorter wavelengths is directed through the drug solution. If a drug is present in sufficient concentration and absorbs UV light, the instrument will plot out a graph of the drug's light absorbance at each wavelength. A typical UV absorbance spectrum for flunixin, extracted from horse urine, is shown in Figure 2. From the shape of the UV curve and the wavelengths at which the peaks occur, the analyst may suspect the presence of a certain drug. He will then change the pH of the system, and if the shape of the absorbance curve changes

[b] For the purpose of this report, a "positive" is a drug finding which violates a medication rule. An analyst reports a positive when there is sufficient data to substantiate the presence of a specific drug or foreign chemical substance. Partial data which suggests but does not satisfy the analyst that a specific drug is present constitutes a "suspicious." A "suspicious" may be upgraded to a positive by the acquisition of more data depending on the medication rules of the particular jurisdiction.[11]

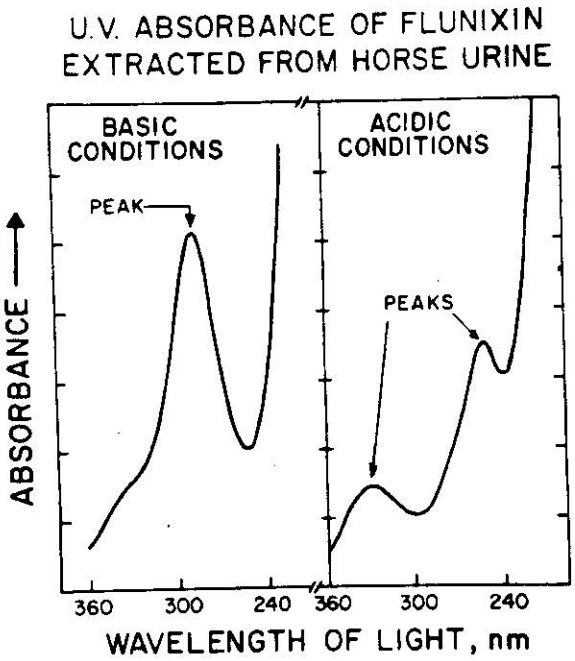

Figure 2. UV absorbance pattern of flunixin. The right-hand panel shows two absorbance peaks for flunixin at 250 nm and 330 nm under acidic conditions, while the left-hand panel shows a single peak at 290 nm under basic conditions.

in a way which supports his opinion as to which drug is present, he will have made a tentative drug identification.

There are two principal problems with UV data as a basis for drug identification. The first is that many compounds share broadly similar UV absorption patterns. For example, all barbiturates generate similar UV absorption spectra, so the test cannot distinguish between barbiturates. Since not all barbiturates are depressants, the method cannot be relied on to identify barbiturate drugs.

The second problem is that the portion of the UV region used by analysts is only about 200 nm wide, which means that there are at the most about 200 different values for UV absorption peaks. If only about 20% of the compounds absorb in the UV, then there are about 4,000 different candidate compounds for each UV peak. The problem is further complicated by the fact that extracts of horse urine contain unknown compounds which also absorb in the UV and will tend to confuse the picture. Because no control (drug free) sample is available, it is not possible to run matched controls. For these reasons UV data may suggest the presence of a compound but cannot positively identify it. UV is therefore considered a screening technique by most analysts.[5]

Thin Layer Chromatography (TLC)

Thin layer chromatography is another useful screening technique. It derives its name from the fact that the experiment is performed in a thin layer of silica gel or other adsorbent coated on a glass or metal plate. The drug extract is spotted about 1 cm from the edge of the plate, and the plate is stood on edge in a solvent system (Figure 3). As the solvent system migrates up the plate, the compounds in the spot move at varying speeds depending on their affinity for either the solvent (mobile phase) or the gel (stationary phase). Compounds which spend about half their time in each phase will migrate about one-half the way up the plate and are spoken of as having an Rf of about 0.5. Other compounds may spend most of their time absorbed to the silica gel and barely move at all from the point of origin. Others may spend most of their time in the solvent and migrate close to the "solvent front"[6] and thus have an Rf in the region of 0.9 to 1.0.

When an analyst suspects the presence of a drug, he makes an educated guess as to which drug it might be. He then reruns the experiment with his best guess running right beside his unknown. If the spots do not migrate the same distance relative to the solvent front, he can be quite sure that the drug and the unknown are not the same substance. If the spots do migrate together, they may be the same substance, but they are not necessarily the same substance.

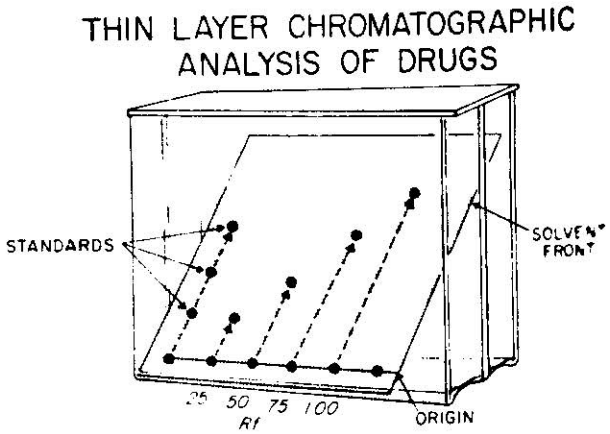

Figure 3. Thin layer chromatography. Drug standards and unknowns are spotted at the origin and the plate stood on end in the solvent. The fraction of the distance migrated by the drug in comparison with the distance migrated by the solvent front is the Rf of the drug.

The reason one can never be sure about identity with TLC is statistical. The maximum number of spots that one can hope to separate physically on a TLC

[6] The solvent front is the leading edge (front) of the solvent as it moves up the TLC plate.

plate is about 20. With 4.2 million chemical candidates, one has about 210,000 possible chemicals for each spot. A 99% specific color or other marker reaction would reduce this number 100-fold, and leave only 2,000 possible candidates for each spot. Specific extraction conditions would reduce this number even further, but a considerable probability of overlap in TLC systems still exists. This is the basic problem with trying to use TLC for definitive identification.

Since most forensic chemists are aware of the lack of specificity of TLC, they usually elect to compare the drug and the unknown in a number of different TLC systems. If both spots again migrate the same distance, the chances that the analyst is right are improved. In an experimental test of this procedure, Sunshine and his co-workers[7] chromatographed 138 drugs in four different solvent systems. They were unable to separate 25 of these drugs in four solvent systems, and in experiments with even seven TLC systems, overlaps were still found. A reasonable conclusion from these experiments is that the number of solvent systems required to separate just the four thousand drugs in common use without risk of overlap is both astronomic and unlikely to be run by the average drug-testing laboratory. TLC is therefore a test which can readily prove the absence of a drug but can never prove its presence. It is consequently not considered a specific test by most forensic authorities.[6]

Gas Chromatography (GC)

Gas chromatography functions on the same principle as TLC with the difference that the mobile phase is a gas, and the stationary phase may be any one of a huge variety of materials. Because the mobile phase is gaseous, the substance to be chromatographed must be volatile, which restricts GC analysis to the approximately 400,000 volatile chemicals and perhaps another 400,000 which can be volatilized after appropriate treatment.

Gas chromatography is like TLC in that some drugs will flow right along with the gas and some will stick to the column (stationary phase) near the origin and never come out. However, if one has chosen the right column pack (stationary phase), mobile phase (gas) temperature, gas flow, and so forth, one can arrange for a nice separation of drugs by GC. To start the experiment, the unknown is injected onto the column and allowed to percolate through the column. Given a suitable detector at the end of the column, one can record the exit from the column of everything the detector will detect. Some typical gas chromatographic records for procaine in horse urine are presented in Figure 4.

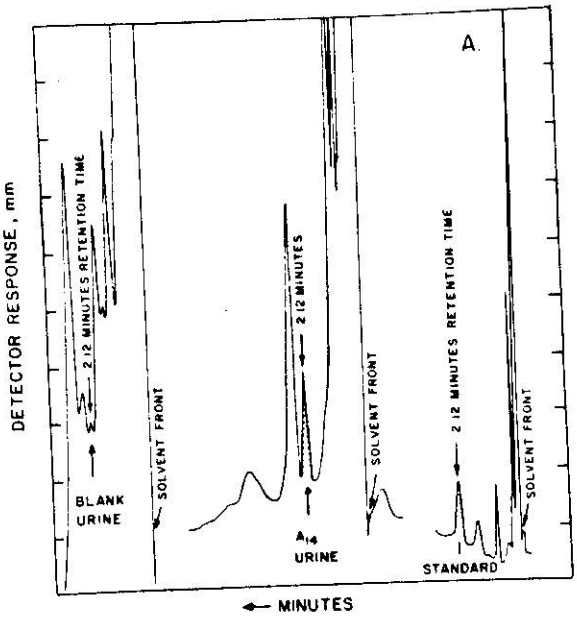

Figure 4. Gas chromatographic analysis of procaine derivatives in horse urine. The right-hand chromatogram is that of a standard solution of 10 ng ml of procaine derivatized with heptafluorobutyric anhydride (HFBA); the center chromatogram is that of a urine sample from a horse pretreated with procaine HCl and similarly treated; the left-hand chromatogram is that of urine from a control horse similarly derivatized.[9]

Now the question of how does one know what each peak the pen writes represents, and for that matter, how can one know that a peak does not contain two or more drugs, obviously arises. The answer to this question is the same as before: you guess at what a peak may be, then test your guess by running a pure (authentic) sample of the drug you suspect. If the pure sample and the unknown come off the column at exactly the same time, and the peaks look exactly the same shape, you may have guessed right. On the other hand, if the pure standard and the suspect peak came off at different times, you know for sure that you are wrong and will have to try another guess.

With respect to the problem of specificity, exactly the same principles apply to GC as apply to TLC. On a 200 m capillary GC column, one can separate about 400 compounds in a four-hour experiment. However, most drug detection laboratories, and especially prerace laboratories, do not have this kind of time to play with, so they use short columns (1 to 2 m) and short retention times. With these systems, one cannot hope to separate more than about 100 compounds. Since there are about 800,000 compounds which can be volatilized, there are 8,000 possible candidates for each peak. The usual procedure at this point is to further test one's guess by changing the column tem-

perature once or twice to see whether or not the material continues to co-chromatograph with the analyst's best guess. Then the pair (suspect and authentic) are usually run on a different GC column at three different temperatures. If the unknown and suspect continue to chromatograph together, many analysts will conclude that they have identified a drug.

A helpful test rarely used is to mix an equivalent amount of the suspect and authentic material and chromatograph them together. If they are indeed the same substance, the suspect peak should double in size, remain symmetrical, and show no tendency to "split," no matter what the chromatographic conditions. This is a particularly useful test, because in its absence one is relying on stopwatch timings of peaks by a person who is paid to match unknown peaks in urine samples with drug peaks. Personally, therefore, I would very much like to see combined samples in every GC run where an unknown is being compared with an authentic for forensic identification.

Again, because GC only produces a small number of data points at best indirectly related to overall chemical structure, one cannot conclude from GC that one has identified a drug. All that one can say is that the unknown and the authentic compound are indistinguishable in the systems in which they were compared. While GC is unquestionably more accurate than TLC, just how accurate it is is not clear. To my knowledge, there is no study on the resolving power of GC to compare with the Sunshine[7] study on TLC. Some analysts suggest a 10% error rate in drug identification based on TLC and GC, while other authorities hold that co-chromatography on three distinct column systems is very good evidence for identity,[a] and that the most important factor is the skill and ability of the analyst.

The Gas Chromatograph—Mass Spectrometer (GC—MS)

In the gas chromatography-mass spectrometer, the unknown materials forming the peaks which come off the gas chromatograph are fed directly into a mass spectrometer, which functions as a detector for the gas chromatograph. However, as well as simply detecting the peak, the mass spectrometer can measure its molecular weight and determine its fragmentation pattern. This gives what is sometimes called a molecular "fingerprint" for each drug, and a good mass spectrum is considered among the best evidence available as to the identity of a drug.[6]

[a] Personal communication from Dr. H. W. Dorough, Department of Entomology, University of Kentucky, Lexington, KY.

The principle of mass spectrometry is quite straightforward. The unknown drug peak from the gas chromatograph is introduced into a vacuum chamber where it is bombarded with an electron beam. The electron beam positively charges the drug molecules and, depending on its energy, may fragment the molecules. These charged particles are then accelerated through a magnetic or electrical field which separates them on the basis of their mass/charge (m/e) ratio. At the end of the analysis tube, the impact of the ions is recorded on an ion detector, and the number of each mass is tallied. Changing the strength of the magnetic field changes the mass/charge ratio of the ions hitting the detector and in less than a second yields a mass spectrum for a drug.

A pair of such mass spectra are shown in Figure 5, one of which (B) is authentic dipyrone and one of which is a racetrack sample identified as containing dipyrone (A). In these experiments, dipyrone (molecular weight 351.4) broke down to yield a fragment of molecular weight 217.2, present in quite small amounts and a series of smaller fragments, all of which are represented essentially equally in both spectra. It is this kind of detailed evidence of similar chemical structures which makes mass spectrometry such a useful tool for identifying drugs. Since each known chemical, and thus its fragmentation pattern, is unique, this type of information yields a virtual "fingerprint" of the chemical in question. In addition, the mass spectrome-

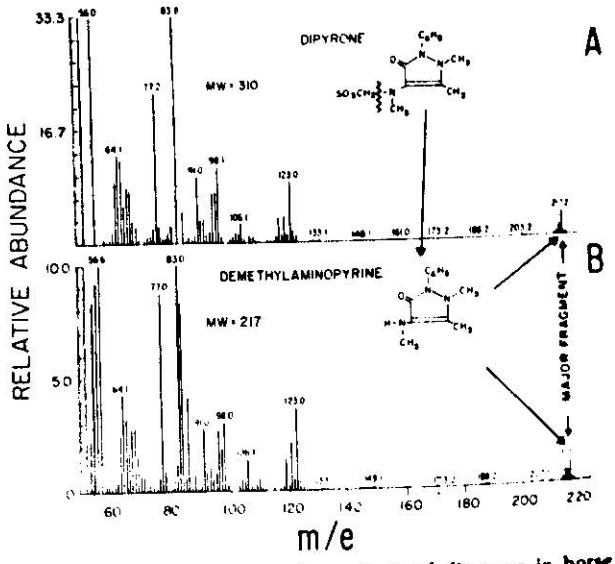

Figure 5. Mass spectrometric analysis of dipyrone in horse urine. Panel A shows the mass spectrum of a dipyrone standard, while panel B shows the spectrum of an unknown recovered from horse urine. The largest fragment present in both spectra is of mass 217.2, consistent with breakdown of dipyrone to demethylaminopyrine in both samples. Many other ions of similar m/e are present in similar abundances, suggesting that sample B does indeed contain dipyrone.

ter is able to detect nanogram quantities of drugs and as such is sufficiently sensitive to be useful for drug detection in body fluids of horses.[4]

High Performance Liquid Chromatography (HPLC)

High performance liquid chromatography is broadly similar to GC except that the mobile phase is a liquid under very high pressure. Because solvents for most chemical compounds are available, the potential field of application of HPLC is much wider than that for GC which requires volatile compounds. However, at this time HPLC is only beginning to be used in routine drug testing.

Radioimmunoassay (RIA)

In RIA, a specific antibody to a drug molecule is used to bind both a radiolabeled drug and whatever unlabeled drug may be added with the unknown plasma or urine sample. If a significant amount of unlabeled drug is added to the system, sufficient radiolabeled drug is displaced for the reduction of radioactivity in the system to be measurable. Given a supply of the specific antibody, RIA is rapid, inexpensive, sensitive, and sufficiently specific to make a good screening test. However, since only a poorly defined portion of the drug binds to and interacts with the antibody, RIA can never be a chemically specific test.

"Calling" a "Positive"

Having briefly reviewed the analytical methods available to the analyst, we can now look over his shoulder as he examines his data. The average analyst will have on hand about 200 standards with which he will try to match any suspect peaks or spots in his test systems. Because the 200 standards cover many of the drugs used routinely on the track, he will be able to match them up successfully much more often than odds of 200 knowns to 4,000, 63,000 or 4.2 million unknowns might suggest. If the analyst has GC-MS with a computerized library, he will have access to about 40,000 chemical spectra, which improves the scope of his search 200-fold.

At some point in this matching procedure, the analyst may decide to call a "positive." He will do this when he is satisfied that he has attained certain analytical criteria for the presence of a prohibited drug. In some jurisdictions these criteria are explicitly stated. For example, to call a "positive" in Canadian racing, the analyst must present evidence of identity in three analytical systems, one of which must be GC-MS. Most United States racing jurisdictions, however, do not have any explicit statements on what the analytical requirements for a "positive" are. Therefore, in these jurisdictions the criteria on which positives are called are the analyst's own. While this system allows the analyst considerable flexibility, it can also leave him open to pressure to call "positives" against his better judgement.

A very real problem arises in "calling positives" when GC-MS data is not available for either economic or technical reasons. If MS data is not available, positives called on TLC, GC, or other empirical methods are "positives" called on evidence with a distinct probability of error. Data from the Cornell University Drug Testing Program, which has the longest experience with GC-MS confirmation of analytical data, suggests that up to 10% of drug identifications obtained with TLC and GC may turn out to be incorrect when tested by MS.[e] Analysts who call positives in the absence of MS data should keep such estimated error levels in mind and only "call positives" on the very best of empirical evidence.

Calling positives on empirical or any other kind of evidence means that the analyst should be prepared to present all his evidence for examination if required. This should include all chart records such as presented in Figures 2, 4, and 5, and all actual physical evidence such as thin layer plates and microcrystals, or good quality color photographs of these if the plates cannot be preserved. The physical evidence should be presented with accompanying evidence from suitable controls matched as closely as possible with the test samples and run *simultaneously* with the tests in question.[6] Just producing typed statments that "UV spectra typical of drug X" were observed is good evidence of an ability to type but totally useless as chemical evidence for the presence of drug "X." Though affidavits and sworn statements carry considerable weight in law, they carry no weight in science, where physical evidence and ability to reproduce the evidence or events are the only criteria. Therefore, the remains of the original duplicate sample should be preserved at all costs for independent analysis. Only in this way can the potential for error inherent in any human process be reduced and all concerned be assured that the matching process on which drug detection depends has been as good as possible.

[e] Personal communication from Dr. Jack Henion, New York Racing and Wagering Board, Drug Testing and Research Program, Cornell University, Ithaca, NY.

References

1. Atti Della Conferenza Internazionale sul Doping. *Stampa Degli Atti a Cura Dello Studio Ego Congressi*, Viale Tiziano, Roma, Italia: 146-147.

2. Blake, J.W. and Tobin, T.: The Gas Liquid Chromatograph and the Electron Capture Detector in Equine Drug Testing. *Bnt J Sports Med.* **10,** (1976): 129-132.
3. Maugh, T.H., III: Speaking of Science. Chemicals: How Many Are There? *Science,* **199,** (13), (1978): 162.
4. Maylin, G.A.: Pre-Race Testing. *Com Vet,* **64,** (1974): 325-334.
5. Shapiro, R.H.: Chemical Defenses in Drug Cases. *J Crim Def,* **2,** (1976): 131-144.
6. Stein, B., Laessig, R.H., and Indriksons, A.: An Evaluation of Drug Testing Procedures Used by Forensic Laboratories and the Qualifications of their Analysts. *Wis Law Rev,* **3,** (1973): 726-789.
7. Sunshine, I., Fike, W. F., and Landesman, H.: Identification of Therapeutically Significant Organic Bases by Thin Layer Chromatography. *J Forensic Sci,* **11,** (1966): 428-439.
8. Tobin, T.: A Review of the Pharmacology of Reserpine in the Horse. *J Eq Med Surg,* **10,** (1978): 433-438.
9. Tobin, T., Blake, J.W., Sturma, L., Arnett, S., and Truelove, J.: Pharmacology of Procaine in the Horse: Pharmacokinetics and Behavioral Effects. *Am J Vet Res,* **38,** (1977): 637-647.
10. Tobin, T., Roberts, B. L., and Miller, J.R.: The Pharmacology of Furosemide in the Horse. I. Effects on the Disposition of Procaine, Methylphenidate, Phenylbutazone and Pentazocine. *J Eq Med Surg,* **1,** (1977): 402-409.
11. Tobin, T., Maylin, G.A., Henion, J., Woodward, C., Ray, R., Johnston, J., and Blake, J.W.: An Evaluation of Pre- and Post-Race Testing and Blood and Urine Testing in Horses. *J Eq Med Surg,* (1979): in press.

Publication No. 44 from the Kentucky Equine Drug Research and Testing Programs, Department of Veterinary Science, College of Agriculture, University of Kentucky, Lexington, KY. Published as Kentucky Agricultural Experiment Station Article No. 78-4-186 with the permission of the Dean and Director, College of Agriculture, Agricultural Experiment Station. The assistance of Mr. Brian L. Roberts in preparing this article is gratefully acknowledged.