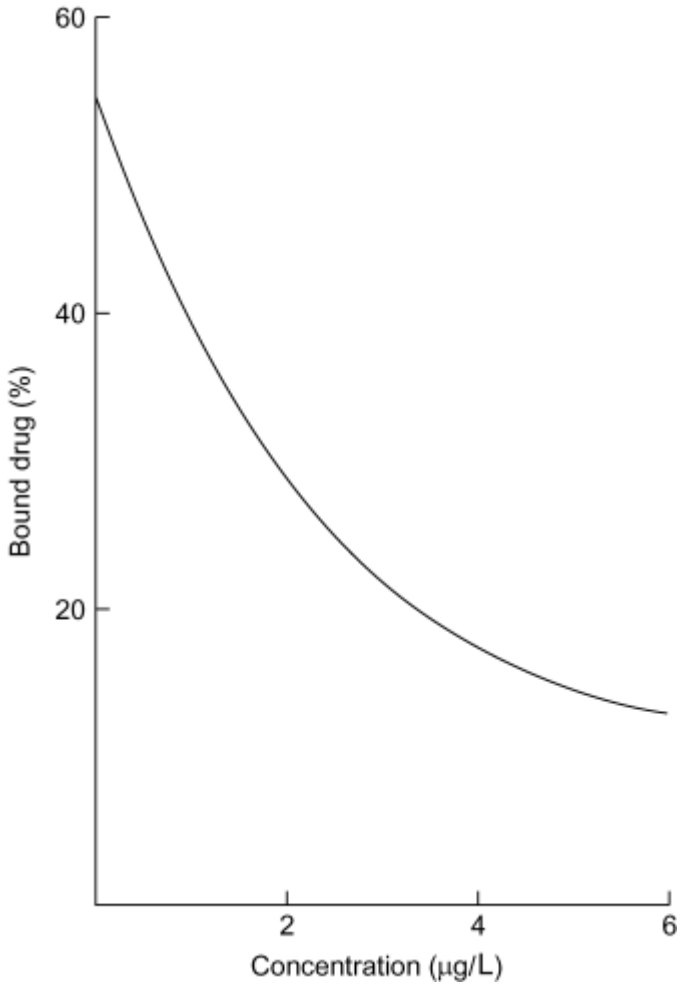
Immunoassays have a firm place among routine methods for the analysis of drugs in biological fluids and other matrices. The technique may be used by the smallest or largest of laboratories with methods that range from single–use point–of–care tests for the analysis of a single sample to fully automated systems capable of analysing thousands of samples per day. Regardless of format, all immunoassays are based on the interaction of a target molecule (antigen) with a corresponding antibody. When applied to drug testing, the immunoassay technique uses an antibody specific for the drug or drug class being assayed, and a labelled form of the same drug or a labelled form of the antibody to generate a measurable signal. Radioimmunoassay (RIA) is popular because of its high sensitivity, its ability to analyse large numbers of samples rapidly and also because preliminary extraction stages are not required.

The drug immunoassay involves setting up a competition for binding to the antibody between antigen (drug) in the sample and a fixed amount of antigen added as part of the test system. Labelling of this added antigen or antibody with a suitable marker and measuring the signal generated by a suitable analytical measurement enables results to be compared with a calibration curve made from a set of standards with known amounts of added drug. Alternatively, the result from a sample or set of samples can be compared with a single–point 'cut–off' calibrator to give a simple positive or negative qualitative result. The technique is capable of high sensitivity, with µg/L quantities easily detectable and ng/L possible with specialist assays.

The basic technique consists of placing a fixed quantity of an antibody in a tube together with a fixed quantity of the labelled drug, and the test sample that contains the drug to be assayed. The specific binding sites on the antibody bind both the labelled drug molecules and the unlabelled drug molecules present in the test sample. The proportion of labelled drug molecules bound is inversely proportional to the number of unlabelled drug molecules.

Immunoassays are divided into two types depending on whether, after the antibody binding reaction has taken place, there is a need to separate the bound and free fractions of the assay before measuring the signal: in RIAs, the radioactivity is measured. However, there is no difference between the signals produced by bound and free labelled drug so that it is necessary to separate the two before measurement. These assays are referred to as heterogeneous immunoassays. Since the amount of radio–labelled drug that remains in the solution, or is bound to the antibody, depends on the concentration of unlabelled drug, measurement of the radio–labelled drug in either the bound or the free form gives an estimate of the original concentration of unlabelled drug. Heterogeneous immunoassays

(see below) using non–isotopic labels have widely replaced RIA, but in this case it is the bound fraction that is measured because a large excess of enzyme is present in the free fraction. With RIA, a calibration curve is constructed in which the percentage of labelled drug bound to the antibody is plotted against the concentration of drug.



Calibration curve for an immunoassay.

For assays in which the drug is labelled with an enzyme or a fluorescent substance, the measurement is of an optically detected change, such as ultraviolet absorption, fluorescence or luminescence. In certain types of optical immunoassays, there is a difference between the signals generated by the free and the bound labelled–drug. Thus, no separation step is necessary and these are referred to as homogeneous immunoassays. The difference between the signals may arise because the signal can be suppressed on binding, produced on binding or altered on binding.

Immunoassays for drugs can therefore be separated into two main groups:

- Homogeneous immunoassays do not require antibody–bound antigen and free (unbound) antigen to be separated before measurement of the signal.
- Heterogeneous immunoassays require the separation of antibody–bound and free antigen before measurement of the signal.

Both types of assay have wide application in the drug–testing field, each with its advantages and disadvantages. Detailed descriptions of different types of assay within each group are given below.

Immunoassays offer a flexible approach to the analysis of various biological fluids for the presence or absence of drugs. They provide a rapid and convenient method to screen large numbers of samples in a variety of matrices and they allow the differentiation of negative specimens and thus avoid the need for further, more complicated and expensive processing or investigation. Where specific assays are available they can be used to quantify accurately and precisely the concentration of drug in a sample as required for therapeutic drug monitoring (TDM) in a plasma, serum or blood sample.

Drug immunoassays have been used successfully in many clinical research applications. Moore et al. (1984) studied the pharmacokinetics of plasma morphine after an intravenous dose, while a series of immunoassays were applied to the study of morphine and its metabolites by Hand et al. (1987). More recently, Parish et al. (1995) used a cotinine RIA to establish a relationship between cotinine and the number of cigarettes smoked.

Immunoassay has become the method of choice for TDM and the principles and considerations described here for immunoassay are applicable to toxicology and drug screening. The main differences between these applications is the degree of quantification required (TDM requires fully quantified results), and the selection of antibodies with the required cross–reactivity profiles. For example, TDM assays may require highly specific antibodies that are reactive only with a single chemical entity, whereas immunoassays used in toxicology or drug–abuse screening may be designed to be broadly cross–reactive with a whole drug group.

Labels used in immunoassay include radioisotopes, enzymes, chemiluminescent molecules, particles (such as colloidal gold or latex beads) and fluorescent molecules. Immunoassay formats vary from systems designed for rapid point–of care testing of a single sample through to fully automated systems that can process thousands of samples per day.

The first RIA method, published by Yalow and Berson (1959), was for the analysis of insulin.

This heralded the beginning of a field of analysis that is now used routinely across the disciplines of clinical chemistry, endocrinology, pharmacology and toxicology. Antibodies raised against morphine by Spector and Parker (1970) were used in the development of an immunoassay using a radioactive tracer (RIA) and many variants have followed since. From an end–user's viewpoint, the vast majority of immunoassays used today are in the form of commercially available, easy to use, automated, semi–automated or simple point–of–care tests. However, the immunochemistry used to produce these test kits involves many intricate interactions of biochemicals with the sample under test. This chapter describes the principles of immunoassay as they specifically relate to testing for the presence of drugs and discusses the major types of immunoassay in common use.

Sommaire

Antibody production

All immunoassays require the use of antibodies; it is this element of the system and the way in which it is produced that is the key to the performance of the assay.

Antibodies are proteins – immunoglobulins (Ig) – produced by beta-lymphocytes in response to an immunogen. There are many different immunoglobulins (IgG, IgA, IgM, IgE and IgD), but mostly IgG is used in immunoassay design. The IgG molecule can be represented simply by the letter 'Y' with two antigen binding fragments, 'Fab' at the top and the 'Fc' at the base.

Drug molecules are too small to provoke an immune response, so drug immunogens are created by conjugation of a drug or drug derivative to a larger carrier protein to give an immunogen through a process called haptenisation. Host animals for antibody production are usually rabbits, sheep or goats for polyclonal antisera, or mice for the production of monoclonal antibodies.

Suitable carrier proteins include bovine serum albumin (BSA), immunoglobulins or keyhole limpet haemocyanin (KLH). The aim of the haptenisation process is to conjugate multiple derivatised drug molecules (the hapten) to the carrier protein. This conjugation is usually via a –COOH or $–NH_2$ moiety on the drug or via a similar group introduced by derivatisation.

The position on the drug molecule used for haptenisation to the protein determines the specificity of the resultant antibodies. Knowledge of the metabolism of the drug and information on structurally related compounds are important when beginning the antibody synthesis. For example, Fig. 2 shows the molecular structure of morphine in which the

carbon atoms at positions 3, 6, 2 and the N-group are all readily amenable to derivatisation and thus enable conjugation to a carrier protein. The position through which the drug molecule is conjugated to the protein is 'hidden' from the immune system and so changes to the target molecule at the conjugation position do not usually affect the binding of the antibody.

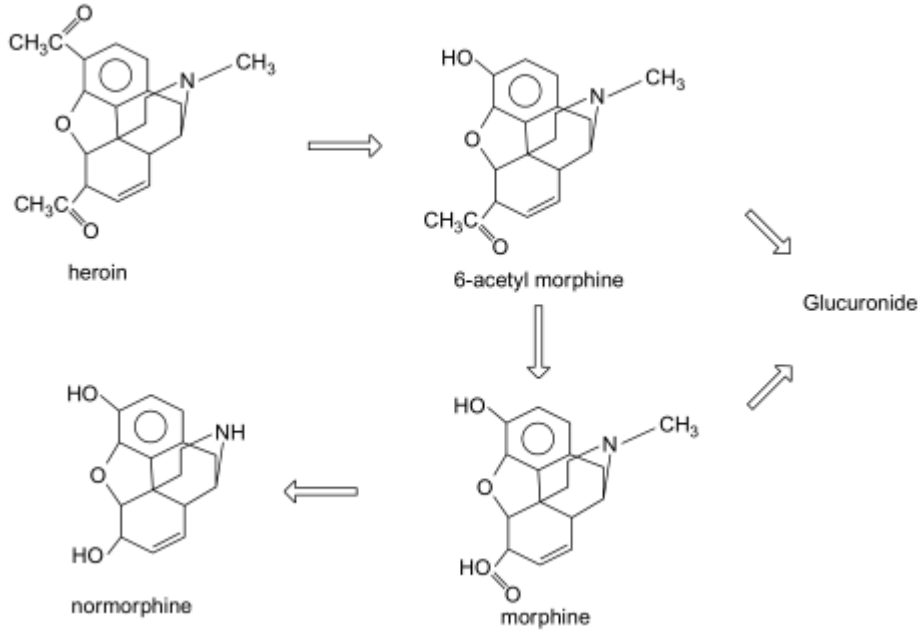


Figure 2. Structure of diamorphine (heroin) and its major metabolites.

The example of morphine is a good illustration of how this can be used to the advantage of the immunoassay developer or the analyst using the assay. If an immunogen is produced via the 3–position it would no longer be a specific determinant against which the antibodies are raised. Hence the resultant antibody is likely to display cross–reactivity with the major metabolite of both morphine and diacetylmorphine (heroin), namely morphine–3–glucuronide. It would also recognise codeine (3-*O*-methylmorphine). The production of morphine antibodies using 3–position haptenisation is used commonly to produce broad cross–reacting 'anti–opiates' antibodies. The cross–reactivity to different opiates varies from one antiserum or antibody to another. It is important that each antiserum or antibody be characterised fully by the assay developer.

Production of morphine antibodies via the 6-position gives better specificity to morphine relative to codeine and morphine–3–glucuronide, and is expected to produce antibodies that display good cross–reactivity to 6–monoacetylmorphine and the active metabolite

morphine–6–glucuronide. Both of these cross–reactions may be desirable, depending on the purpose to which the antibody is being put.

To produce a more specific assay for morphine, derivatisation and conjugation via the nitrogen group can be utilised. This leaves the 3- and 6-positions as the antigenic determinants and therefore produces antibodies that are more likely to be specific for morphine without cross–reactivity to codeine or dihydrocodeine, for example. An *N*-linked antiserum will, however, give potential for cross–reactivity with molecules such as normorphine, which have an alteration at the *N*-position.

Another consideration during antibody production is the use of a 'bridge' molecule. This is employed as a chemical spacer between the hapten and carrier protein to allow better access to the drug for potential antibodies. It is important to remember that the immune response in the host animal may also produce antibodies to the whole immunogen – not just the drug molecule – which includes the bridge molecule. Interference from 'bridge antibodies' can be avoided or minimised by using a different chemical spacer to that used in immunogen synthesis when conjugating the drug–signal reagent.

The measure of the strength of the binding between an antigen and an antibody is described by the affinity constant. This binding is non–covalent, reversible and reaches equilibrium. High affinity antibodies bind faster than low affinity antibodies and perform better in immunochemical methods.

Monoclonal antibodies versus polyclonal antisera

Polyclonal antisera contain a complex mixture of antibodies raised during the immunisation process. By comparison, a monoclonal antibody is a single entity that results from the isolation of a single antibody–producing cell. In drug immunoassays, the higher affinity antibodies produced by polyclonal antisera can sometimes be preferable to the lower affinity antibodies produced by a monoclonal system.

Polyclonal antisera

Immunisation of a rabbit or sheep usually proceeds with 20 to 100 µg of the protein–conjugated hapten, which can be mixed with an adjuvant to stimulate the immune system. This produces an immune response that consists mainly of IgM followed by IgG. This response is further stimulated by additional immunisations of a similar or lower dose at

regular intervals – typically every 3 to 4 weeks for 20 weeks or longer. Test bleed samples of approximately 1 to 2 mL are taken at 10 to 14 days post–injection, at which point the immune response (and therefore antibody titre in the serum) is at its highest.

After analysis of the test bleeds, a larger antiserum sample can be drawn. This can be as much as several hundred millilitres when using larger host animals (e.g. sheep). Antisera yield can vary from hundreds to many thousands of tests per millilitre depending on the success of the immunisation programme and the type of assay in which they are employed. It is not the absolute volume of serum that is important, but rather the amount and quality of antibodies contained therein.

Using standard separation methods, such as ammonium sulfate precipitation, Protein A separation and/or ion–exchange chromatography, antisera can be purified for further use. Affinity chromatography can be beneficial in certain circumstances and involves passing the antiserum over a column that contains an immobilised form of the drug of interest. In this way it is possible to fractionate a complex polyclonal antiserum. In some cases purification is not necessary and the antiserum (i.e. the serum itself) can be used with simple dilution.

Monoclonal antibodies

Monoclonal antibodies offer the advantage of a continuous supply of antibodies with the same characteristics, so once a good antibody is selected it can be used indefinitely.

After immunisation and successful test bleeds, monoclonal antibodies are made by the fusion of mouse lymphocytes (or lymphocytes from other species) from the spleen with myeloma cells. The resultant hybridoma cells are separated by limiting dilution to give single cells that secrete single monoclonal antibodies. This technique was first described in 1975 and is now in routine use (Kohler and Milstein 1975).

Monoclonal antibodies generally have less affinity than polyclonal equivalents, which can lead to less sensitive assays. Monoclonal antibodies are not more specific than polyclonal antisera, but once a specific antibody is selected, the cell line can be stored and the antibody produced indefinitely. Note that in drug testing, it is possible for an antibody to be too specific, as it may be desirable to have broad cross–reactivity to a drug family (such as benzodiazepines) or to a single drug and its metabolites (such as buprenorphine).

Monoclonal antibodies offer the advantages of purity and homology, which is useful for circumstances in which the antibody is being labelled or conjugated as part of the

immunoassay set up – for example when being labelled with an enzyme or coated with colloidal gold.

Other molecular biology and recombinant techniques, such as phage display (Chiswell and McCafferty 1992), in which the genetic code is harnessed to produce antibodies give an exciting additional source of this important immunoassay component.

Antibody dilution curves

Once an immunisation is underway, the quality of the antiserum is assessed by means of an antiserum dilution curve. This demonstrates the binding of the antibody to the target drug and is an indication of the antibody's affinity for the antigen. Using a heterogeneous enzyme immunoassay (EIA) as an example, a test bleed of rabbit anti–cotinine was coated onto a microplate at different dilutions in bicarbonate buffer (pH 9) using incubation overnight at room temperature. The resultant binding to different titres of horseradish peroxidase–labelled cotinine. Dilutions of the enzyme–drug conjugate for this type of experiment can be made in simple phosphate buffer (pH 7.4) with a small amount of protein (e.g. BSA at 0.05% w/v) to prevent non–specific binding (NSB; i.e. non–specific binding) of materials to the assay tube. The hook (or apparent peak) at high antiserum concentrations is caused by a combination of steric hindrance and saturation of coating antibody to the microplate.

Once the concentration of antibody (titre) that produces the desired response is selected, it is necessary to check that the antibody also responds as required with the target drug (i.e. that the target drug successfully competes with the labelled drug for binding to the antibody). The improved performance with subsequent test bleeds for samples taken from a rabbit immunised with a buprenorphine-protein immunogen. It is possible to gain useful information by combining both the experiments described above and analysing both a positive and negative sample at each antibody dilution. This way, binding and displacement can be seen at each antibody titre.

Careful titering of the labelled drug derivative and antibody dilution can improve the assay characteristics, and the assay can be optimised further by the addition of other proteins, surfactants and stabilisers to the assay buffer.

An immunisation programme usually involves the injection of between three and six animals with the same antigen. If suitable antibodies are not produced after several immunisations, it may be necessary to start the programme again with different animals and possibly a

different immunogen.

Analytical specificity

The analytical specificity or cross–reactivity of an immunoassay provides an indication of how the assay responds to other drugs relative to the drug used as a calibrator or standard. The cross–reactivity profile of the immunoassay is important when assessing if it is suitable for a particular task. For example, determination of cocaine use by the analysis of a urine specimen requires the immunoassay to exhibit good cross–reactivity to the urinary metabolite benzoylecgonine. By contrast, the assay needs to demonstrate good cross–reactivity to cocaine if the samples to be analysed are saliva or hair extracts, since high concentrations of the parent drug relative to metabolites are found in these specimens.

With classic competitive immunoassays (such as RIA and EIA), cross–reactivity is often calculated relative to the amount of drug that displaces 50% of the antibody–bound signal reagent. In the drug–testing field it is common to add a known amount of drug to the assay and record what response is given. Percentage cross–reactivity is then calculated as:

$$\text{Cross reactivity (\%)} = \frac{\text{Apparent concentration of target drug}}{\text{Concentration of added drug}} \times 100$$

For example, if in an opiate assay (calibrated with morphine standards) to which a solution of 1000 ng/mL of codeine is added as a test sample, the result from the assay read against the morphine calibration curve is 100 ng/mL, the cross–reactivity of codeine in the assay relative to morphine is 10%, i.e.

$$\text{Cross reactivity (\%)} = \frac{100\ \mu\text{g/L apparent morphine}}{1000\ \ \mu\text{g/L added codeine}} \times 100 = 10\,\%$$

However, for many drug immunoassays it is not sufficient to test cross–reactants at a single concentration. Several concentrations must be prepared in the sample matrix under test and the effect on the assay examined, effectively making a calibration or standard curve for each potential cross reactant.

Non–parallel cross–reactivity can be seen within a drug group for both parent drug(s) and metabolite(s). Isomeric forms can also exhibit different cross–reactivity profiles and it is common to see different reactivity between (+)-amfetamine and (–)- amfetamine.

Drugs tested for cross–reactivity indicate to the analyst which compounds might be appropriate for subsequent confirmatory analysis after a positive immunoassay result. Drugs that show little or no cross–reactivity *in vitro* can give strong immunoassay positives from people who take the drug *in vivo*. For example, benzfetamine is converted into metamfetamine and amfetamine, but the parent drug would not show cross–reactivity in many metamfetamine or amfetamine immunoassays.

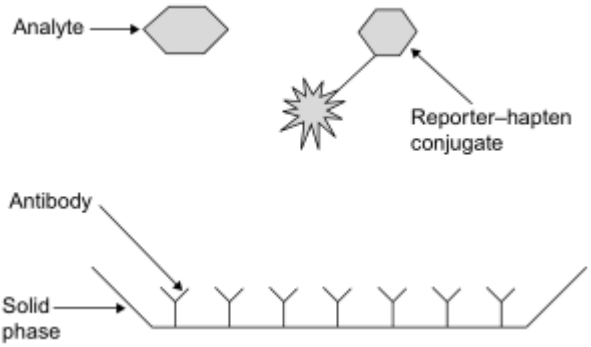
The important factor with regard to antibodies from an analyst's perspective is to know the characteristics of the antibody used in the immunoassay. Knowledge of the type of immunoassay used is important when considering the next stage of the analytical process, for example gas chromatography–mass spectrometry (GC–MS). If an immunoassay is used to demonstrate the use of a drug by measuring its metabolites in urine, the confirmation method should be configured appropriately. This can be achieved by the introduction of hydrolysis steps so that, for example, glucuronide metabolites, which are detected by the immunoassay screen, can then be analysed as the parent drug compound by GC–MS. Other metabolites may exist to which the immunoassay may be directed, as with the buprenorphine metabolite norbuprenorphine, and consideration must be given to the most appropriate substance to target by GC–MS. It is normal practice to set a lower cut–off for confirmatory analysis than that used for the initial screening method.

The above assay design features have important consequences when using immunoassay for samples other than urine (e.g. saliva, hair and blood). For example, if the drug in the sample is largely present in parent form in saliva or hair, the antibody and immunoassay must be directed to the parent molecule rather than to the metabolites.

Heterogeneous immunoassays

Heterogeneous immunoassays require the separation of bound and free antigen before measurement of the signal. This is because of a lack of difference in the signal generated from the antibody–bound drug or free-label. The separation step of heterogeneous assays gives two distinct advantages over homogeneous assays. Firstly, potential endogenous interference from the sample produced by whole blood or highly discoloured urine is removed at the wash stage prior to signal development. This has the benefit that preliminary sample–extraction steps are not required. Secondly, the reaction has lower limits of detection compared to homogeneous assays. This difference is illustrated by the ability to use whole blood directly in a heterogeneous test kit, whereas extraction of the blood sample is generally required before analysis in a homogeneous method (Perrigo and

Joynt 1995) and in the detection of potent drugs such as LSD (Cassells et al. 1996; Kerrigan and Brooks 1999). Other applications for heterogeneous assays include screening of hair and saliva, in which drug concentrations are significantly lower than in urine.



Basics of an immunoassay.

Enzyme immunoassays

The terms enzyme immunoassay (EIA) and enzyme–linked immunosorbent assay (ELISA) are often used interchangeably to outline non–isotopic assays and, more loosely, to describe heterogeneous and homogeneous assays.

Within the context of heterogeneous methods, EIA and ELISA have widely replaced RIA as the assay system because of the advantages already outlined. EIA methods are based upon antibody capture using labelled antigen systems, while ELISA methods are based on labelled antibodies. Both types of immunoassays are competitive systems.

Antibody–capture systems

The system uses anti–drug antibodies coated onto a microplate well by passive absorption. A microplate is a tray of 96 wells (each having a volume of approximately 400 µL) that are generally arranged as 12 × 8–well strips. The coating can be achieved by simply adding a solution of the antiserum or purified antibody in a bicarbonate buffer at pH 9 using high–precision pipettes. A suitable incubation period (e.g. overnight at room temperature) allows the antibodies to bind to the plastic microplate well. Following this incubation, the solution is washed from the solid phase and dried to leave an antibody–coated well. In some cases, a secondary coating of a non–relevant protein is performed to increase stability of the solid–phase antibody and to prevent NSB of assay components to the plastic well of the microplate. This process forms part of the reagent manufacturing process of commercial

kits provided as a dry microplate with the antibody pre–coated.

In running the assay, a sample (10 µL to 50 µL) is added to the microplate followed by a buffered solution (typically 100 µL) that contains a fixed amount of drug labelled with an enzyme such as horseradish peroxidase or alkaline phosphatase. The plate is then left for sufficient time to allow the horseradish peroxidase–labelled drug and any drug present in the sample to compete for binding to the solid–phase antibody. This is referred to as the incubation period. In the absence of drug, the maximum binding of the enzyme label occurs. Increased amounts of drug in the sample result in decreased amounts of antibody–bound enzyme label. After the incubation period, the microplate is washed with a suitable buffer (the separation step) to remove all traces of unbound enzyme label. Antibody–bound enzyme conjugate (and antibody–bound drug) is left immobilised to the wall of the antibody–coated microplate.

After the wash step, a substrate reagent is added to the microplate well and left to develop colour. The substrate of choice for horseradish peroxidase is dilute hydrogen peroxide with tetramethylbenzidine (TMB) as chromagen, which is available as a liquid–stable, ready–to–use product. TMB provides a blue coloured signal that can be measured at 630 nm using a standard laboratory microplate reader. In routine use, the colour development is stopped, after it has developed over a 20 to 30 min incubation period, by the addition of 1 M dilute sulfuric acid. This causes a shift in absorbance that is read at 450 nm. The whole assay process described here can be fully automated or performed with the simplest of laboratory equipment.

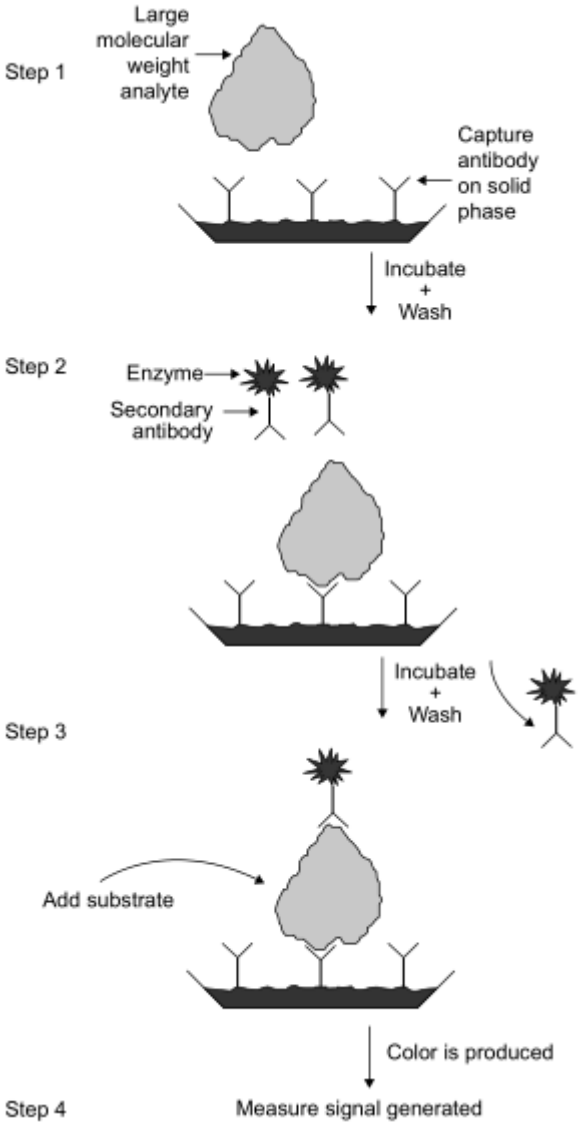
The amount of enzyme–labelled drug bound to the solid–phase antibody, and therefore available for colour development, is inversely proportional to the amount of drug in the specimen. The colour produced at the final stage of the EIA is therefore inversely proportional to the amount of drug in the specimen, which gives a calibration curve similar to that found with RIA .

Enzyme–linked immunosorbent assay, antibody labelled systems

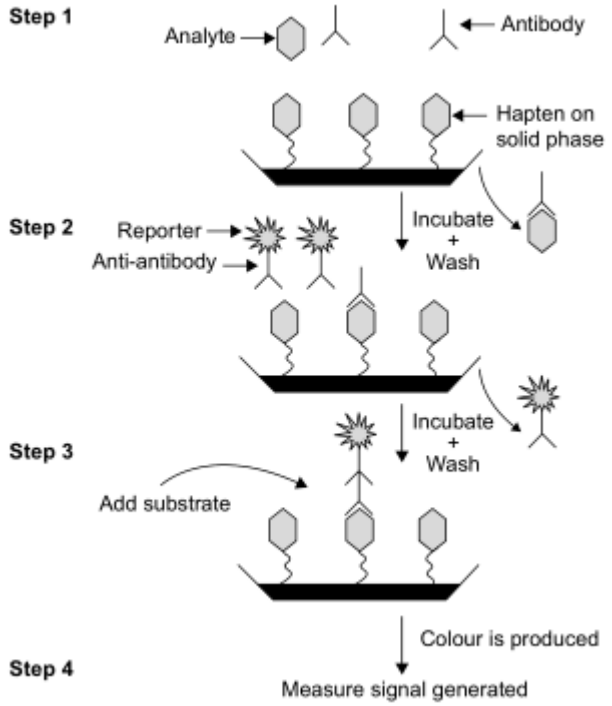
Competitive ELISA procedures are similar to the antibody–coated microplate EIA described above. The difference is that the anti–drug antibody is enzyme labelled rather than the drug. A drug derivative is coated onto the plastic well, which serves as a way to separate bound and free fractions.

The drug is conjugated to a protein using a process similar to that used to prepare an

immunogen. This must be a different protein from the one used as the carrier protein to make the antibody in order to prevent binding of anti–protein antibodies made during the immunisation process. The protein–drug conjugate can be coated to the microplate well in the same manner used for coating antibodies in EIA (see above). During the assay, using similar procedures to those described for EIA, competition occurs between the drug in the sample and the immobilised drug for binding to an enzyme labelled anti–drug antibody. Following incubation, a wash step is used to separate bound and free fractions, and leave labelled antibody immobilised to the solid–phase drug derivative on the microplate well wall. The amount of labelled antibody that is bound, and responsible for the signal generation, is inversely proportional to concentration of drug in the sample.



The basic scheme for an ELISA assay. Note the wash between steps. This is the distinctive feature of a heterogeneous assay.



The scheme shows a typical competitive solid-phase immunoassay commonly used to detect drugs of abuse. It is a heterogeneous format using washing steps to remove materials that did not bind immunologically. Signal is inversely proportional to the concentration of free drug in the sample.

Radioimmunoassay

RIA was the forerunner of heterogeneous immunoassays. Difficulties associated with the handling and storage of radioactivity, disposal of radioactive waste and the half–life of the radioactive labels have resulted in RIA being replaced largely by non–isotopic EIA methods.

Two main types of RIA have been used for drug testing based on the isotope employed. An isotopic drug label is termed a tracer. Use of tritium (^{3}H)- or ^{14}C-labelled drug allows an

identical molecular structure of tracer to the drug being tested, although counting the beta (β) emissions involves an organic–based scintillation fluid and a sophisticated beta scintillation counter. Use of the gamma–emitting ^{125}I as tracer allows faster and more efficient counting and a number of research (Hand et al. 1986) and commercial RIAs were based on this isotope.

The immunoassay principle for RIA is the same as that for EIA described above. Antibody–coated tubes have been used to good effect with iodinated tracers. Simple decanting easily separates the bound and free fractions, and allows quantification of the bound fraction using a gamma (γ) counter. Other ways to separate bound and free fractions include the use of a second antibody precipitation step. This requires a centrifuge to form a pellet of bound material that can be counted after the free fraction has been decanted or aspirated. Magnetic beads and other particles coated with a second antibody have also been used to aid this separation of the bound and free fractions. Dextran–coated charcoal has been used to absorb free tracer from solution, and thereby allow the antibody–bound fraction in the supernatant to be counted (Bartlett et al. 1980).

Chemiluminescence immunoassays

Chemiluminescence offers the potential for increased sensitivity by an order of magnitude greater than RIA. However, such increased sensitivity would result in detection limits far greater than those of current confirmation techniques and so chemiluminescence assays have not gained popularity in either urine screening or toxicology.

The signal generated in a chemiluminescent immunoassay is caused by compounds that emit light during a chemical reaction. The flash of light can be extended by the addition of enhancers to the system. Enhanced luminescence, as seen with luminol in the presence of horseradish peroxidase and substituted phenols, causes a prolonged ‘glow’ of light that is easier to measure (Whitehead et al. 1983). Substrates that provide chemiluminescent end–points for other enzyme labels are available, such as adamantyl dioxetane phosphate, as used with alkaline phosphatase in the DPC Immulite system (Hand 1994).

Fluorescent labels

Fluorescent end–points for heterogeneous immunoassays are based on a fluorescent label or fluorophore. The principle of fluorescence relies on the ‘Stokes’ shift’ or movement of wavelength from that used to excite the fluorophore to the detection wavelength at which

light is emitted from the fluorophore. The labels can be used in an immunoassay in the same way as a radioactive label or by using an enzyme label to generate a fluorescent molecule.

An example of this type of signal system utilises an alkaline phosphatase enzyme label and 4–methylumbelliferyl phosphate substrate. The substrate is converted into the fluorescent 4–methylumbelliferone, which can be detected by a suitable fluorimeter. When using the assay procedure described for the heterogeneous EIA above, the fluorescence obtained is inversely proportional to the amount of drug in the sample.

The use of fluorescent labels in toxicology testing has not gained wide popularity because of the interference from endogenous sample components.

Lateral flow methods

There is a growing trend in the immunoassay sector to develop point–of–care tests. The Syva Emit system (see Homogeneous methods) on small analysers, such as the ETS, has allowed drug testing outside the laboratory for many years (Centofanti 1994). The use of inexpensive, easy–to–use single disposable cartridges or slides for a variety of drugs has accelerated this trend.

Single–use cartridges are basically variants of a competitive ELISA assay run on a solid–phase strip rather than inside a well or tube. These tests use particles as the means of signal generation and are mostly used with antibodies that have been labelled with colloidal gold or coloured latex spheres. This technique was established for urine drug testing and can also be applied to serum, whole blood (using a suitable filter pad to separate red cells) and saliva (employing a suitable collection method).

As with ELISA, an antigen–protein derivative is fixed to a solid–phase. In this case it is bound to a nitrocellulose membrane as a line or series of dots at a defined position . Labelled antibody (typically labelled with colloidal gold or coloured latex) is located in a pad that overlaps with the nitrocellulose membrane.

When the pad is wetted with sample, the labelled antibody re–hydrates and flows from the pad trun “indexing onto the membrane by capillary action. As the sample and re–hydrated gold–labelled antibody flow along the membrane, any drug in the sample binds to the labelled antibody and inhibits it from binding to the immobilised drug on the membrane. Separation of bound and free drug continues, as part of the assay process, with the movement of liquid across the membrane leaving either bound material (a visible line)

or free material to travel to an absorbent end pad. As with ELISA, the amount of colour developed (or line formed at the site of the immobilised drug) is inversely proportional to the amount of drug present in the sample.

The addition of a control line on the membrane provides an indication that sufficient fluid has passed through the length of the test strip and thus an indication that the test has been performed correctly. Control lines typically use anti-IgG antibodies to capture excess drug–specific labelled antibody.

The addition of a number of drug–derivative lines in different positions on the nitrocellulose membrane together with suitable matching labelled antibodies allows the analysis of multiple drugs at the same time from a single sample.

The majority of these tests are interpreted visually, and the reagents are titrated and optimised such that there is no colour or line present when the ‘cut off’ concentration of analyte is present. Recent developments have overcome the problems of visual interpretation associated with lateral flow tests. This has been achieved by developing a small portable reader based on digital imaging. The benefits are that the line can be determined more accurately, the result can be quantified if required and the subjectivity of visual interpretation is removed (Spiehler et al. 2000).

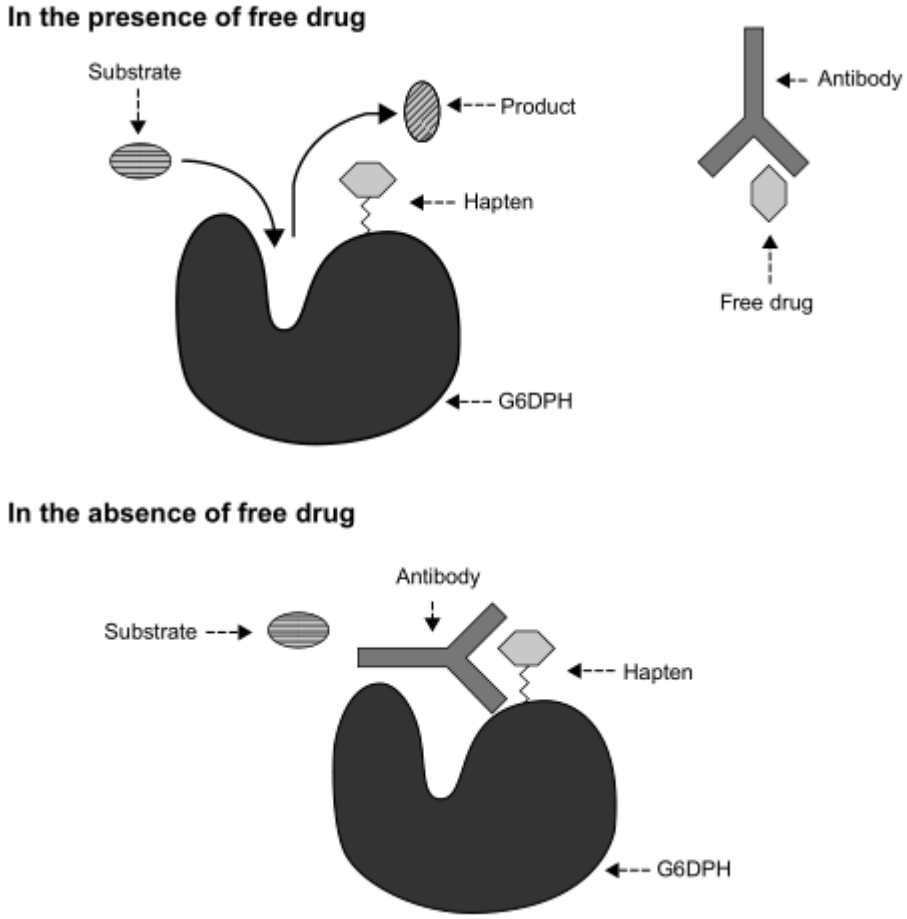
Homogeneous immunoassays

Homogeneous immunoassays do not require the separation of bound and free fractions, and as a result have been automated successfully in a number of different systems.

The most widely used and most cost–effective immunoassay used to screen large numbers of urine samples when there is no need for high sensitivity is the homogeneous EIA marketed by Syva (Dade-Behring) for many years as EMIT (enzyme multiplied immunoassay technique). This method, for use for the analysis of morphine, was first published in 1972 (Rubenstein et al. 1972).

The EMIT technique is based on the enzyme activity of a drug–labelled enzyme, glucose–6–phosphate dehydrogenase (PDH), modulated by antibodies raised against the drug. Enzyme activity (in the presence of glucose–6–phosphate) results in the conversion of nicotine adenine dinucleotide (NAD) to the reduced form NADH and the subsequent increase in absorbance at 340 nm is monitored spectrophotometrically.

Addition of the anti–drug antibody results in binding to the drug–labelled enzyme, which effectively reduces the enzyme activity. Any drug in the sample competes with the drug–labelled enzyme in binding to the antibody, which allows unbound enzyme to become active and thereby increases absorbance at 340 nm. Therefore increasing amounts of drug in the sample produce increased enzyme activity and hence an increase in the rate of change of absorbance at 340 nm.



EMIT assay scheme; signal is directly proportional to the concentration of analyte.

The use of bacterial PDH, which uses NAD as coenzyme, avoids interference from endogenous PDH, which uses nicotine adenine dinucleotide phosphate (NADP). Earlier examples of the technique used lysozyme (Rubenstein et al. 1972) or malate dehydrogenase (Ullman and Maggio 1980; Kabakoff and Greenwood 1981).

CEDIA, cloned enzyme donor immunoassay

A more recent variant on the homogeneous enzyme immunoassay is the CEDIA system. This system, like EMIT, uses the binding of an antibody to influence the activity of an enzyme. Competition for binding by drug present in the sample results in an increase in enzyme activity.

The CEDIA principle (Henderson et al. 1986) relies on the spontaneous complementation of two genetically engineered fragments of beta-galactosidase from Escherichia coli. The polypeptide chain of the enzyme is presented as two inactive fragments, the large fragment, termed enzyme acceptor (EA), and a small fragment, enzyme donor (ED), which contains the 5% of enzyme missing from the EA portion. By conjugating hapten to the ED fragment, the addition of antibodies that bind the hapten can prevent the formation of an intact and therefore active enzyme. Any drug present in the sample competes for antibody–binding sites, such that an increase in drug concentrations yields less binding of antibody to the ED fragment and therefore results in more enzyme activity which can be monitored spectrophotometrically.

Fluorescence polarisation immunoassay

Fluorescence polarisation immunoassay (FPIA) is another example of a homogenous immunoassay, and uses fluorescein–labelled antigen as the tracer molecule (Dandliker et al. 1973). It has been widely used as the basis of assay systems to analyse drugs in urine (Colbert et al. 1985).

The fluorescein–labelled antigen rotates rapidly when not bound by an antibody. When bound by the antibody, the rotation is slowed dramatically compared to the unbound molecule. To generate the assay signal, a fluorimeter shines light, at the excitation wavelength for fluorescein, through a vertical polarising filter. Rapidly rotating unbound fluorescein molecules emit light in a different plane to the incident light, whereas the relatively stationary antibody–bound fluorescein returns light in a similar plane, which is detected via the polarising filter.

Drug added via the sample competes for binding to the antibody with the fluorescein–labelled hapten, thereby reducing the amount of fluorescein bound to the antibody, which results in less emitted fluorescence being detected via the polarised filter. The background fluorescence found with many biological samples means it is usual to take a

blank reading of the sample and reagents before the addition of the fluorescent tracer to the mixture. This method is the basis of the Abbott ADx system.

Micro–particle methods

Agglutination of micro–particles has been used as a signal system in homogeneous based methods. Agglutination reactions for drugs use particles, generally latex, that have drug conjugated onto them. Addition of antibody (which has two binding sites) to the drug–coated particles causes bridging between antibodies and the particles to produce agglutinates. The drug present in a sample, added to the mixture of antibody and drug–conjugated particles, competes for antibody–binding sites and therefore prevents agglutination. In this type of assay the degree of agglutination is inversely proportional to the amount of drug in the sample.

The basis of the Roche Abuscreen Online system termed KIMS (kinetic interaction of micro–particles in solution) involves monitoring the rate of agglutination via spectrophotometric means.

Automation of immunoassay

The ability of an immunoassay to screen large numbers of samples has lead to the development of a number of different automated solutions. These can be tailored to meet the requirements of laboratories using either semi–automated systems through to full automation with walk–away capability. Automation of a heterogeneous assay sequence up to the separation step is referred to as the ‘front–end’, while those steps post separation are termed ‘back–end’. Full automation of all stages requires a system able to carry out the separation stage. Front–end automation involves an accurate pipetting mechanism for samples and reagents. This is achieved by employing a robotic sample processor that facilitates the automatic transfer of samples (with bar code recognition) from primary tubes to reaction wells (or tubes). This transfer can be done in batch mode (i.e. all samples are pipetted from one destination to another) or by random access (i.e. the selective transfer of sample to one or more different reaction wells) and can be read directly from an imported work–list.

Sampling can be done using either disposable pipette tips or a fixed polytetrafluoroethylene (PTFE)-coated probe coupled with a high–precision syringe drive. Use of disposable pipette tips removes the possibility of carry–over, while fixed probe instruments require extensive

washing to reduce carry–over (see Troubleshooting, later). As this increases the processing times, multiple probe instruments have been developed to increase throughput.

Back–end automation can also use pipettes for the reagents (e.g. substrate) and, as with front–end systems, can have the ability to time incubation stages and maintain elevated temperatures if required. Both systems require a mechanism that allows multiple batches to be run at the same time. Back–end systems also include a reader and suitable software for data reduction, and can be interfaced with laboratory mainframe computers.

Full automation is essentially a combination of front- and back–end systems that incorporates a suitable mechanism to complete the separation phase. For microplates this involves a simple ‘on–board’ washer that automatically dispenses and aspirates wash fluid. Similar systems have been developed for tube washing, although the size constraint means these are stand–alone modules. The automation process of heterogeneous enzyme immunoassays is well developed and is widely used in blood–bank laboratories in which many thousands of tests can be processed per day. Automation of homogeneous assays is less complex as they do not require a separation step. The most widely used systems, such as the EMIT, CEDIA and OnLine methods, are readily automated on modern clinical chemistry analysers. These analysers are able to pipette accurately, control temperature and measure rates of absorbance. The throughput of the instrument can range from a hundred up to a thousand tests per hour.

Analysis of alternative samples to urine

Immunoassays for drugs have been widely employed in the analysis of urine samples. Other sample matrices, such as whole blood, serum, hair extracts, saliva and sweat, can be used successfully. Heterogeneous assays are ideally suited to these alternative matrices. Their inherent sensitivity provides the lower limits of detection required for forensic use and the wash step endows the ability to deal with difficult matrix effects. Correct selection of antibody allows the assay system to be targetted towards specific drugs or metabolites.

Immunoassays are generally designed for a particular use and sample type. With testing for drugs of abuse the assay is often designed for urine specimens, whereas therapeutic drug assays are generally designed for plasma or serum. Situations can arise in which the drug of interest is in a different matrix to that for which the kit has been validated, so the onus is on the user to test and validate this new application. Urine and serum immunoassay kits are widely available and assays for other matrices, which include blood, cerebrospinal fluid, fingernails, hair, meconium, saliva, stomach contents, and sweat, have been developed

successfully.

Many assays can be adapted to work with different matrices, although it is advisable to consult the kit manufacturer prior to making kit adaptations. After discussion and advice from the developer or manufacturer of an assay, the next step is to obtain similar samples to the test case; it is essential to try and match these as closely as possible, with special reference to the type of preservative used. Some samples (e.g. bottled beer) are readily available and it should be possible to obtain the exact brand under test, so that the alcohol and other ingredients are matched identically. In cases such as aged, putrefied whole blood, to obtain and select a sample matrix can prove more difficult

Once a suitable sample matrix has been obtained, known amounts of the drug of interest are added to the sample matrix. The concentrations are usually chosen to fall within the range of the kit. The high sensitivity of immunoassays means it is often possible to introduce a dilution factor without compromising the limit of detection in the sample. For example, if the limit of detection of the confirmatory method is 100 µg/L, use of an EIA with a cut–off of 10 µg/L allows the use of a 1:10 dilution factor. Such a dilution reduces potential interference without compromising overall sensitivity.

Heterogeneous immunoassays are more robust and tolerant of different types of sample matrices than homogeneous assays because of the separation stage of the assay.

A common finding, often described as a ‘matrix effect’, is that the newly created calibration curve is not useable through adverse interference. This can arise from an inherent interference from the sample itself with the signal system or from interference with the antibody–antigen reaction. Homogeneous immunoassays are more prone to interference since the sample is present during the signal–generation stage. Reducing the sample load by dilution in a suitable buffer or using a smaller test volume should reduce the effect. Unfortunately, the trade–off is that the limit of detection and curve shape might be sacrificed. Liquid- or solid–phase extraction of the sample can be used when direct application of the sample fails.

Quality control, calibration, standardisation and curve fitting

It is common practice to perform laboratory–based immunoassays as a batch or a single run that contains calibration and control material together with unknown samples. From the

calibrators (or standards), a calibration curve can be derived and a concentration for the unknown sample can be determined. Inclusion of samples with known drug concentration (controls) at regular intervals allows an assessment of reproducibility within the batch. Duplication and control of reproducibility forms the basis of 'within run' precision assessment.

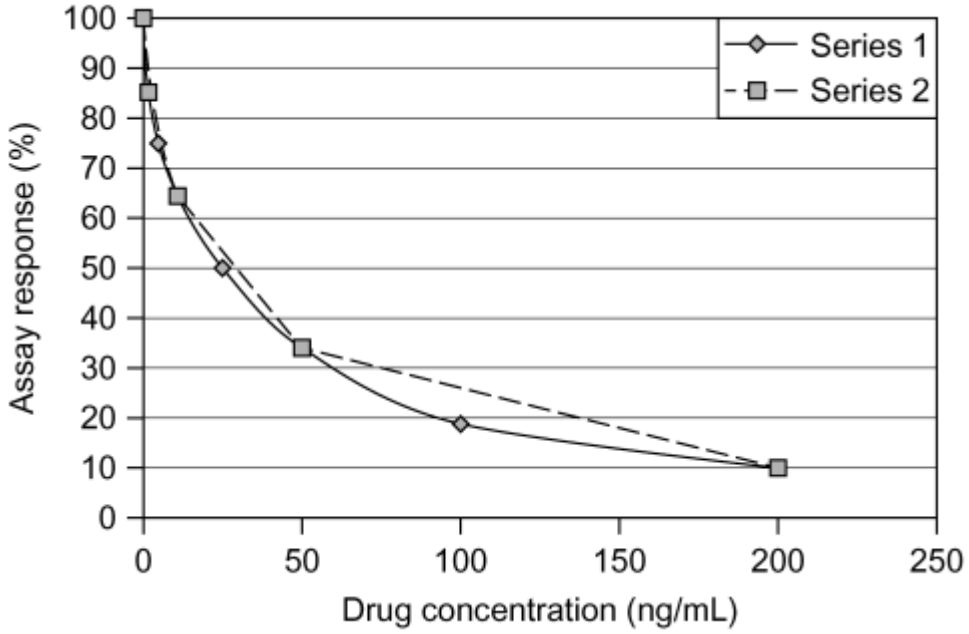
Inclusion of the same quality–control material in consecutive runs gives an indication of run–to–run reproducibility. This can be supplemented with the inclusion of previously analysed patient or test samples. Together with monitoring curve parameters, these components allow an opinion to be formed as to the acceptability of the batch–to–batch precision when compared with predefined parameters.

Drug immunoassays can be run in a number of ways, depending on user requirements. Urine drug screening may only require a simple qualitative (positive or negative) result. Fully quantitative results are required for diagnostic purposes or TDM. In forensic applications, an assay run in a semi–quantitative mode (with a small number of calibrators) provides an estimated concentration. This estimated concentration is useful when progressing to a confirmation stage (e.g. GC–MS).

For quantitative assays, the curve fit between data points is important and a number of curve–fitting options are available. Some manufacturers supply 'closed' curve–fitting programs that determine the fit, and calibrator ranges are fixed on all instruments of that particular make. Open systems allow the user to select a suitable fit and transformation for the particular application. Care should be taken with this selection, as it is possible to mismatch the curve fit and immunoassay. Common problems include automated systems that attempt to fit a straight line to a curve or poorly fit the points of a curve. There is no substitute for the operator checking the raw data from an assay.

For toxicological analysis, it is often appropriate to use a point–to–point curve fit (while accepting that the curve when far from a point is less accurate). Other examples of curve fits include spline and four–parameter fits. The overriding consideration is to ensure, by validating the method, that the chosen fit connects the data points in the most suitable way. A simple example of how the same assay can produce different results because of data handling is shown in Fig. 9. Both curves are plotted using the same results, but in Series 1 extra calibrator points are plotted. The curves plotted for Series 2 work perfectly satisfactorily as a 'semi–quantitative' assay, but give inaccurate numerical results at points where the point–to–point curve fit is not suitable to provide a quantitative answer (illustrated best in Fig. 9 at the 100 µg/L region). The introduction of extra calibrators provides a more robust curve more suited to quantification of drug concentration. The

addition of extra calibrators also ensures that a single calibrator point does not unduly bias the curve. In Fig.9, using Series 2 any small error or change to the 200 µg/L calibrator alters the curve significantly, but with Series 1 the curve is held by the 100 µg/L calibrator point.



Example to illustrate how different curve–fitting programs can influence the results of an immunoassay.

Troubleshooting

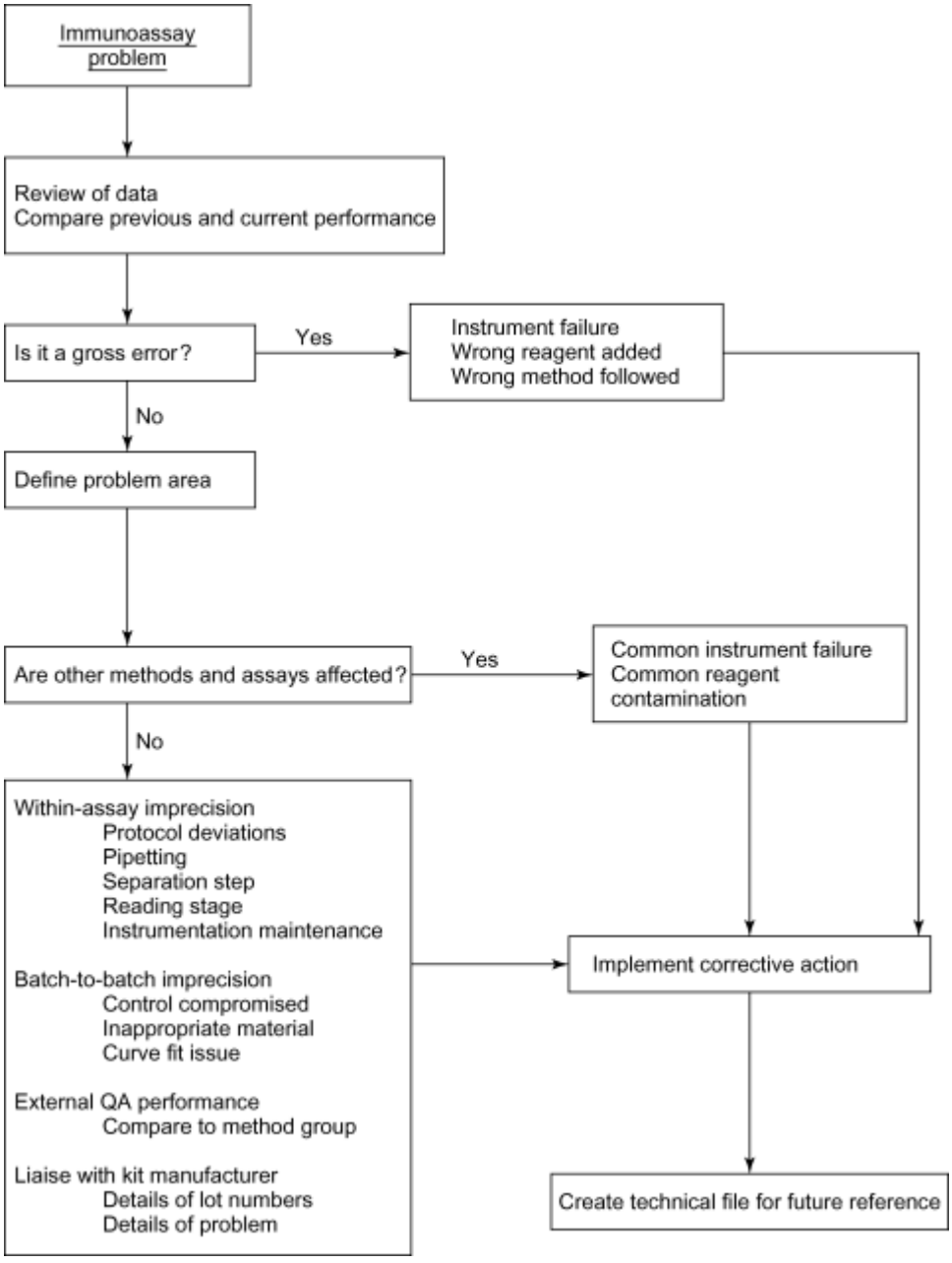
Problems can, and will, arise in all immunoassay tests, whether these are methods developed in–house or are commercial assays. It is beyond the scope of this section to cover troubleshooting of all methods as their diversity is too vast. It may be necessary to undertake a series of experiments to establish the cause of the problem. This largely depends on the nature of the problem, but the golden rule is to change one variable at a time. This ultimately saves time and reagents.

When the problem has been corrected, details of the investigation should be logged in a technical file for each immunoassay method. Ideally, this file should be initiated at the method development or validation stage.

Figure. 10 offers an approach to troubleshooting. With the exception of a complete assay

failure, which usually results from a gross error, most problems require careful and detailed analysis. This involves comparing the current poor performance of an assay against past data. Historical reference material can take the form of data accumulated while validating or developing the method, internal and external quality control (QC) data and possibly analysis of qualified samples (e.g. those with known GC–MS status).

By assessing this information it should be possible to establish if an assay kit or method is performing outside the set parameters and to gain an insight into the cause and ultimately to resolve the issue. In defining the problem area, most issues can be categorised under the headings of within–assay imprecision, batch–to–batch imprecision and external quality assessment performance.



Flow chart to assist with troubleshooting problems encountered with immunoassays.

Within-assay imprecision

All immunoassays have an inherent imprecision and for this reason it has been common practice to perform replicate analyses. For a particular immunoassay, the agreement between replicates is a consistent feature that can be assessed for each assay. Poor

replication is easily identifiable and is often associated with a pattern. Those laboratories that only analyse singletons lack this early warning system.

A number of factors can cause within–assay imprecision and it is good practice to employ preventative testing regimes rather than look for curative problems after the event. Problems with sample and reagent pipetting are commonplace. Apart from poor pipetting technique, more subtle problems can arise, particularly when robotic sample handling is involved. These employ either one or more fixed sample probes or probe(s) with disposable tips. Most are based on liquid displacement and do not perform optimally if there is air in the tubing or a liquid–leaking valve or syringe tip.

Carry-over of sample material can be a particular problem with fixed–probe instruments. Often, manufacturers quote figures between 0.01 and 0.001% for carry-over. In drug testing, it is common to find grossly elevated samples that contaminate the samples following in a run. For example, a sample that contains 500 mg/L transfers 50 µg/L at 0.01% carry-over. Extensive washing of the probe on both its external and internal surfaces reduces the carry-over to acceptable levels, although this has the detrimental effect of slowing the sample pipetting. Reagents can also be contaminated or diluted with each other in a similar way, so it is good practice not to add them with a fixed probe instrument. Disposable–tip instruments are far superior in this respect, since they have zero carry-over.

As with any instrument, regular maintenance and performance testing is essential for reliable results. A simple coloured dye or radioactivity pipetting test is satisfactory to demonstrate pipetting accuracy and precision, provided it mimics the actual immunoassay and sample as closely as possible. This seems an obvious statement, but most immunoassays utilise relatively small volumes (typically <100 µL), while many instrument manufacturers recommend testing at higher volumes (500 to 1000 µL).

A particular problem for chemistry analysers utilising homogenous assays is bubble formation or froth in the sample cup. This is sufficient to activate the liquid level sensor and results in air being aspirated instead of the sample. Aspiration of reagents can be affected similarly. Bent or damaged probes can also cause similar effects that can lead to poor replication, false negatives and positives.

The progressive displacement of control values in one direction during a run can be described as assay drift. This can be manifested as either an increase or a decrease in concentration through a run. The cause of this problem is an alteration to the assay kinetics within the batch, and arises because of two factors. Firstly, the incubation time for many

solid–phase assays is relatively short (to allow for high throughput) and timely addition of samples and reagents is required. Secondly, given that the reaction rate of an assay generally increases with temperature, care should be taken to allow reagents to reach the operating temperature. Reagents pipetted directly into the first tubes or wells from 2 to 8° storage have longer to warm up than the later tubes or wells and thus have different reaction kinetics.

The washing (separation stage) of a heterogeneous enzyme assay is a critical step, since most assays employ a large excess of enzyme, most of which is removed prior to signal generation. Wash volumes should be set to fill the well or assay tube, and aspiration of the waste should be set to be as complete as possible. The soak time is of equal importance and should be set for between 30 s and 2 min. For microplate assays, it is usual to wash either by column across the plate or by row down the plate using an eight- or 12–way wash head, respectively. Partial blockage of a probe can cause inefficient washing, which will lead to imprecision. In a plate of replicated samples using an eight–way wash head this gives a pattern of a poor replicate on every fourth pair throughout the run. Blockage of the aspiration ports or dispense probes can occur through a build–up of wash buffer salts and sample material, so regular maintenance is essential for good performance. Automatic flushing of water through the wash head is a useful feature in this respect.

Reading devices such as gamma counters and microplate readers can also cause within–assay imprecision. In an attempt to improve throughput, many instruments are fitted with multiple read devices (e.g. multi–head gamma counters). It is essential that these devices are matched to each other and checked regularly.

Non–validated changes to the immunoassay protocol can lead to within–assay imprecision. Commercial immunoassays are optimised to work with a particular sample volume and matrix. A number of parameters can be altered to meet user requirements, but such changes must be validated. Validation should involve extensive testing over a minimum of three different reagent lots, and for commercial kits should involve the manufacturer who should be able to help and advise.

Batch–to–batch imprecision

Any cause of within–assay imprecision has a detrimental effect on batch–to–batch precision. Provided the control material has been validated for use with the immunoassay in question, it will function as a means to assess batch–to–batch precision. As discussed earlier, such validation involves the assessment of a number of batches of both kit and control material. If

this validation has not been completed, problems can arise. Some examples of causes of problems derived from the control material itself include:

- Level of drug beyond the calibration range (above or below) of the assay.
- Matrix effect through the use of an animal–based control that behaves differently to the material (e.g. human) being tested.
- Use of a control in one matrix in an immunoassay designed for another matrix (e.g. a urine control used in a plasma assay).
- Inadequate reconstitution of control material and deterioration of the control material through long–term storage or repeated freeze–thaw cycles.

When choosing a commercial drug control, it is important to consider the cross–reactivity of the immunoassay. Controls can contain several congeners of the same drug group that could be recognised in the immunoassay and give higher than expected results. Conversely, if the control contains a mixture of isomers, lower than expected results can occur if the immunoassay has a higher cross–reactivity to one isomer compared with another.

Immunoassays can be used qualitatively to provide a positive or negative answer, semi–quantitatively to give approximate results, or quantitatively to yield the most accurate value. For simple screening, comparison to a cut–off calibrator is sufficient. Use of a semi–quantitative curve allows an estimate of drug concentration that is useful in certain circumstances, such as when confirming the result by another method. If a more accurate result is required a comprehensive calibration curve may be required (see Fig. 9).

A plethora of software and curve–fitting options are available that will transform both the axes and calibration curve for quantitative and semi–quantitative data reduction. However, many of these options are not compatible with some assays and lead to spurious results if implemented. Indeed, the same raw data applied to the same curve fit can give significantly different results. This test of robustness should ideally be carried out at the validation stage of the method evaluation.

External quality assessment performance

As discussed above, any cause of within–assay or batch–to–batch imprecision may be reflected in the external quality assurance (QA) performance. However, samples from these schemes represent a small fraction of the laboratory’s workload and should not be relied upon to indicate problems with internal QC.

External QA benefits the drug–testing community by allowing the same sample to be tested by a large number of laboratories using a variety of methods. It also allows unusual or difficult samples to be circulated. The objective of this type of scheme is to achieve consensus between the different testing locations and to look for any method–specific or location–specific variances. The individual laboratory is able to see that its performance matches that of others.

Sample adulteration

Specimens donors (i.e. patients, employees, detainees) may not want the immunoassay drug screen to perform correctly and may take steps to adulterate their specimen. This is largely a problem associated with screening of urine for drugs of abuse; hair, blood and saliva are less likely to be adulterated. Steps to counter this problem include observing the collection of the sample, together with physical and chemical tests to confirm the sample is voided freshly and from the individual being tested. Tests include measurement of sample temperature at the point of collection, on–site or laboratory tests of pH, relative density and urinary creatinine concentration. More specific tests can be performed for particular adulterants.

The effect of sample adulteration differs depending on the assay type. Some assays produce false–negative results, whereas others produce artificially elevated results.

The key to the successful application of immunoassay as part of a testing regimen is knowledge of the characteristics of the assay in use and knowledge of the limitations as well as the benefits of each of the many variants of the method. [1]Clarke's Analysis of Drugs and Poisons

Références

↑1 Clarke's Analysis of Drugs and Poisons