
Laboratory Measurements of Hormones and Related Biomarkers: Technologies, Quality Management and Validation

Martin W. Elmlinger

Nycomed GmbH, Konstanz, Germany

Laboratory measurements have become an essential part of medical diagnostics. In pediatric endocrinology, measurements of hormones and related biomarkers like growth factors are practically indispensable. They initially allow the evaluation of the pathological stage and differential diagnosis of a patient. In the following, they are used to monitor and predict a therapeutic success and outcome. Thus, measurements of hormones and growth factors must be valid and reproducible in both technical and medical terms. The laboratory must warrant the technical validity and reproducibility of the hormone measurements through standardized laboratory procedures and adequate quality standards and quality management. It is, however, the responsibility of the treating physician, first to choose the right laboratory parameters to be examined, and second to ensure adequate test circumstances including the time of sampling, influence of nutrition and cyclic factors, co-medication and shipment of the sample. The physician/pediatrician is responsible for interpretation of the laboratory results and the use of this medical information in patient care.

During the last two decades, laboratory diagnostics in endocrinology and in general have undergone a technological revolution in several ways. The major trends are the following.

Multiplex Techniques and Miniaturization

Multiplex techniques represent a current megatrend in laboratory diagnostics with the aim of saving time and resources upon the release of the diagnostic result, and

of doing multiparametric analysis at one go. Multiplex panels can include up to ten-thousands of parameters. This goes along with miniaturization to reduce the sample volume needed for testing and thereby reduces invasiveness.

These techniques are based on either protein or antibody microarrays or chips, coated beads, glass fibers, or microcapillary discs. The immunological methods amongst them make use of the classical interaction of an antigen (analyte) with an immobilized antibody, or vice versa. Other than that, immobilized RNA, DNA, cDNA, aptamers, etc. are used. These methods do not play a significant role in endocrine routine diagnostics and are not discussed here.

Nevertheless, singleplex immunoassays like ELISA or RIA still play a central role in the endocrinology laboratory, due to their robustness and high sensitivity. One reason why they have not been replaced by multiplex techniques yet is the possible interference of different antibodies in multiplex, especially bead-based techniques. These interferences can lower the sensitivity and specificity of the methods. Thus, today multiplex and chip technologies are used in screening approaches in biomedical sciences and drug discovery, rather than in quantitative routine diagnostics.

Plasmon resonance as a principle of antigen detection is not used in clinical routine, but it can be helpful in the screening of binding properties of antibodies, such as binding strength and kinetics during the development of immunoassays.

Significant progress, however, has been achieved by making use of intelligent modifications in the detection systems and washing technology.

Automation, Integrated Care Systems and Theranostics

The pressure of reducing costs and laboratory staff and the necessity to speed up laboratory measurements, has fueled the development of fully automated diagnostic platforms. These are often assembled of modules, which can be used in different combinations, in order to increase the degree of automation and the capacity of measurements. They are used in large laboratories with a high daily throughput of samples. Midsized or small-sized automated platforms are also available for smaller laboratories. These systems are available for clinical chemistry, as well as immuno-diagnostic testing. The development of such high-technology platforms is, however, only possible in larger companies with an appropriate financial background. Sample and patient registration, sample injection from original vessels, dilution and the whole analytical procedure, including quality management, data handling and data transfer, as well as a customized standardized data output are managed automatically by the system. Automated data management can even include preliminary evaluation of data on the basis of relevant reference ranges.

The next step of innovation includes the implementation of integrated care systems, as driven by the big players in diagnostic industry (e.g. Siemens, Abbott, Philips, Roche), who combine imaging and laboratory diagnostics and link them

with an integrated concept of medical care. A typical feature of those integrated systems is a seamless data flow from bedside to the laboratory and back to the treating physician. The whole analytic procedure, but also the data management and transfer can be performed in a fully regulated environment (see 'Quality management', below).

Another kind of development is reflected by theranostics, which means the co-development of a drug with a specific diagnostic to pre-test the individual patient for responsiveness and thereby to optimize the therapy.

Biomarkers in Personalized Medicine

Personalized medicine makes use of biomarkers in several ways with the aim to find and optimize drug therapies for the individual patient. Generally spoken, the goal is to treat the right patient population with the right drug at the right dose regimen, in order to achieve a maximum response. Before a patient is treated, stratification biomarkers can be measured, if appropriate, to choose the right therapy or to adjust a dosage. Genetic predisposition like mutations of the 21-hydroxylase gene, which goes along with a high 17OH-progesterone level, is evaluated in the molecular diagnostics of adrenogenital syndrome (AGS). Other typical stratification biomarkers are individual expression levels or polymorphisms of drug transporters or metabolizers (e.g. cytochrome P450). Disease biomarkers are markers which reflect a relevant pathway of the respective pathomechanism. If this marker or set of markers responds adequately to a drug therapy, it can be also used as a pharmacodynamic biomarker. The serum level of insulin-like growth factor-1 (IGF-1) and of its major circulating binding protein IGFBP-3 are types of these of biomarkers. They are used to diagnose and assess the degree of the human growth hormone (hGH)-deficient state, and to measure responsiveness and predict the outcome of hGH therapy of a patient. They are also used in the differential diagnosis of growth disorders in order to elucidate the etiology of the disease. In addition, IGF-1 and IGFBP-3, i.e. their ratio, can serve as safety biomarkers of hGH therapy.

Point of Care Diagnostics

Cost pressure in the health care sector, the development of diagnostic techniques which can also be operated by nursing staff, and the demand for immediate measuring of results (e.g. in intensive care) are all factors contributing to the increasing use of point of care testing (POCT) equipment. While comparable with conventional laboratory methods in terms of accuracy and specificity, these instruments save time and reduce laboratory staff costs. POCT instruments are expected to find still wider use in the future, as they are available even in pocket-sized formate to be used anywhere.

Measuring blood glucose, HbA1c, and insulin in outpatient facilities are examples. Another approach to be used in outpatient facilities is standardized analysis strips to measure, for example, IL-6 in neonatal sepsis, critical biomarkers in stroke/cardiac disease management, e.g. brain natriuretic protein, luteinizing hormone and parameters of infection serology. Some strips like hCG strips for pregnancy testing are means of self-observation, which can be used even at home. Most of these technologies are based on antibody techniques.

Thereby, the markers can be measured from different fluids (blood, urine, etc.) mostly without any sample preparation and consuming a minimum of sample. During the last years POCT methods have become more reliable from a technical point of view and are comparable to classical laboratory techniques, including the integration in quality management systems [book recommendation: Luppä PB, Schlebusch H: Patientennahe Labordiagnostik, Heidelberg, Springer, 2008].

New Assays and Compartments

Tremendous progress has been achieved in the number and quality of parameters that can be measured from body fluids. Commercial providers constantly try to implement new assays on automated platforms, thereby massively increasing their analytical menus. However, the quality of singleplex assays is also continuously being improved, e.g. by switching from RIA to ELISA, and by increasing the sensitivity through technical improvement, e.g. high-sensitivity cortisol. It is also an important goal of today's diagnostics development to identify body fluids that can be gained less invasively, e.g. saliva to measure cortisol.

Quality Management

As the major part of the diagnostic and therapeutic decisions in pediatric endocrinology relies on hormone measurements, the latter must be valid and reproducible in both technical and medical terms. This request can be fulfilled by implementation of standardized quality management systems (QMS). A QMS is defined as the whole of all measures suitable for achieving the quality of a product or a service (analysis) that has been agreed upon or required for a certain purpose. An efficient and adequate QMS can be proven if the laboratory accepts an accreditation procedure, i.e. a proof of competence by an independent third party. The practice of licensing of diagnostic laboratories by national governments is handled differently in different countries. The use and management of internal and external quality controls (e.g. Ringversuche) represents a minimum of QMS in the clinical laboratory in Germany. With the goal of international harmonization, international standards like CLIA, (<http://www.cms.hhs.gov/clia/>), ISO9000, EN45000 and/or GLP (Good Laboratory

Practice) have been set up to ensure the quality of laboratory diagnostics across borders. Coordinating institutions like the WELAC (Western Europe Laboratory Accreditation Cooperation) are the drivers of implementation of the quality standards in diagnostic laboratories.

Immunoassay Methods to Measure Hormones and Growth Factors

Bioassays

Bioassays were the first available methods of determining hormone concentrations. A historical example is the measurement of human growth hormone (hGH) concentration using the growth of the proximal epiphyseal cartilage of the femur of hypophysectomized rats [1].

Today, *in vivo* and *in vitro* bioassays are only used in two areas:

Firstly, in the standardization of hormone preparations; examples are establishing the international reference preparation (IRP) for recombinant human growth hormone [2] and for human leptin [3] by reference laboratories of the World Health Organization. Here, bioassays complement physical and biochemical methods such as gravimetry, high-performance liquid chromatography (HPLC), protein microsequencing, mass spectrometry and immunoassays. The best approximation of the real IRP is attained by averaging the concentrations gained by the various methods.

Secondly, bioassays are sometimes used to relate the biological activity of a hormone in the serum of patients, e.g. of TSH [4] or of hGH to its measured concentration. Through the introduction of an automatic cell counter the measurement of the biological activity of hGH using the proliferation assay on rat Nb2 lymphoma cells [5] has reached a precision comparable to that of immunoassays. Since the introduction of the stable GH receptor transgenic cell line BA/F3-GHR [6], the GH bioactivity can even be measured directly with GH receptor-mediated cell proliferation, instead of indirectly by prolactin receptor-mediated cell proliferation as with Nb2 cells. However, bioassays are not used in routine clinical diagnostics.

Radioreceptor Assays

Radioreceptor assays (RRA) make use of the specific binding of the respective hormone to a membrane (peptide hormone) receptor or cytoplasmatic (steroid or thyroid hormone) receptor. A major difference to bioassays is that here only the binding and not the biological activity of the hormone is examined. The quantitative evaluation is thus comparable to that of immunoassays. Lymphocytes are mostly used as target cells as they are usually easy to handle [7]. As receptor binding is generally weaker than binding to specific antibodies, RRA is very prone to disturbances. For example, the presence of soluble GH-binding proteins is a serious hindrance for the GH RRA [8]. RRA also plays a role in testing the receptor binding potential of synthetic steroid variants or analogs [9].

Immunofunctional Assays

Immunofunctional assays are variants that involve binding of the ligand to be measured to the binding site of its natural receptor, and subsequent recognition of the ligand bound by a specific monoclonal antibody that serves for quantitative detection. This kind of assay has been developed, for example, for leptin, GH and IGFBP-3 [10–12].

In situ Methods

The use of labeled antibodies and antibody sandwiches for detecting immobilized antigens in microtome sections of tissues (immunohistochemistry), in cells (cytochemistry) and after transfer (blotting) onto carrier membranes in an electric field (dotblot, Western immunoblot), generally belongs to basic research areas but is also used in infection serology. The antibodies can be labeled with an enzyme, a dye or a fluorophor.

Immunoassays

Antibody Production

An important part of the humoral immune answer to foreign immunogens (cells, antigens, haptens) is the production and secretion of specific antibodies by B lymphocytes. Antibodies can thus be produced by repeated immunization (boostering) of animals with pure antigen or hapten solutions using Freund's adjuvans. The antiserum can be directly used or is purified and enriched by affinity-chromatography using a protein-A or protein-G column. The IgG fraction obtained contains polyclonal antibodies directed against various epitopes (fig. 1).

The antigen-binding $F(ab)_2$ and the F_c region are the functionally important areas of the IgG molecule (150 kDa). The former consists of a light (V_L) and a heavy chain (V_H), which together form the antigen-binding slot like a keyhole. Due to the enormous number of possible antigens (10^8), the composition of the 110 amino acids in the V area is highly variable, especially in the three hypervariable regions (V_H). In contrast, the amino acid composition in the C-regions is largely constant as these are responsible for complement and macrophage activation.

The method of producing monoclonal antibodies devised by Köhler and Milstein (Nobel Prize, 1984) is based on the following principle: after immunizing mice with a particular antigen, antibody-producing B cells are gained from their spleen and immortalized into hybridoma through fusion with a mouse myeloma cell line. The hybridoma cell line thus unites the potential immortality of the myeloma cells and the ability of the B cells to produce a single kind of specific antibodies. Through separating the individual hybridoma, re-cloning and testing

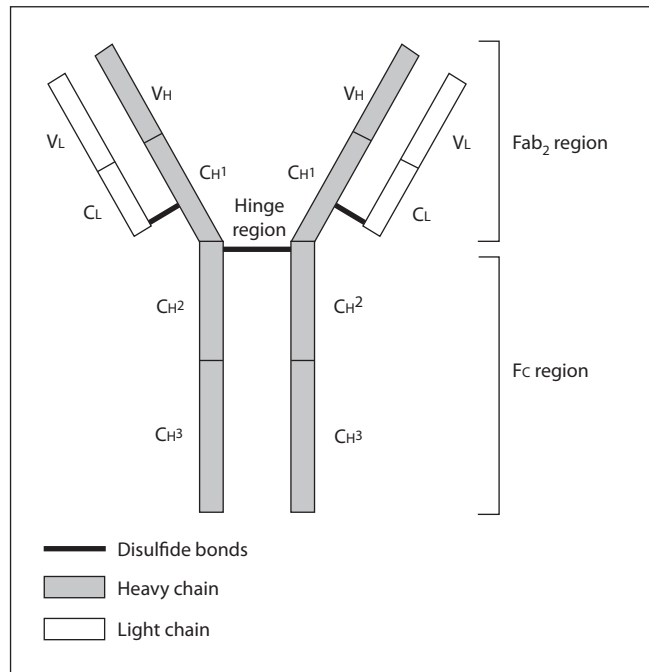


Fig. 1. Antibody structure (IgG).

for specificity, particular cell clones can be gained which produce IgG for specific epitopes. These can then be used to produce large quantities of specific antibodies in bioreactors or in mouse ascites (in Germany, the latter is forbidden by the Law for Protection of Animals).

Today, genetic engineering enables the cDNAs which encode the bond-specific Fab parts of IgGs to be cloned and recombinantly expressed. By cloning, targeted modifications can be brought into the protein expressed. Nowadays one often only uses the specific F(ab)₂ or Fab fragments instead of the whole IgG molecule. As long as stocks of the genetically programmed organisms (*Escherichia coli*, eucaryotic cells) are preserved, monoclonal antibodies are producible in nearly unlimited quantities and constant high quality, in contrast to polyclonal antibodies.

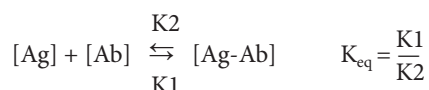
Antigen Production

Pure antigens can either be produced by sequential chromatographic steps from biomaterials (body fluids, cell lysates) or, nowadays much more commonly, from lysates after recombinant expression of cloned cDNA sequences in *E. coli* or in eucaryotic (insect, mouse, monkey) cells. Recombinant proteins can be produced in unlimited amounts and constant quality. The choice of the expression system depends on the respective protein. *E. coli* can easily express large amounts of relatively small peptide hormones, but is often unable to produce complicated high-molecular-weight proteins with disulfide bonds in correct folding. In this case,

eucaryotic expression in mammalian or insect cell systems with specific vectors have to be used.

Principles

Immunoassays are an *in vitro* implementation of the specific, high-affinity bonding between antibodies (IgG) and their antigens. According to the law of mass action, a steady state develops between the concentrations of the bound and free antigens [Ag] and antibodies [Ab]. An important step in the immunoassay is therefore the subsequent removal of the unbound component in order to evaluate the amount of bound antigens or antibodies:



Berson and Yalow's [13] discovery of the radioimmunoassay (RIA) for measuring insulin in 1959 opened a new chapter in the history of hormone measurements and replaced the hitherto used bioassay and the precipitation tests (radial double diffusion, immunoelectrophoresis, rocket immunoelectrophoresis) which were based on making the precipitates directly visible at the point of equivalence. The rapid success of immunoassays was promoted by improved techniques for producing antibodies and for their chromatographic purification. Later, recombinant expression of antigens led to further improvements. Further, various labeling and detection techniques (radionuclides, enzymes, coupling through biotine/streptavidine, dyes, fluorophores, chromophores, etc.) have been developed and intelligent methods of phase separation (adsorption techniques; fixed phases: microplates, coated tubes, beads) have been implemented. Depending on whether reagents are used in limited or in excess amounts, one speaks of competitive or noncompetitive immunoassays. The usability of the RIA is limited by the half-life of the radionuclide used (e.g. ^{125}I , $t_{1/2} = 60$ days) and by the problem of radioactive waste disposal. Nevertheless, the RIA is a robust and inexpensive technique.

Competitive versus Noncompetitive Assays

The competitive radioimmunoassay (RIA) is based on the competitive reaction between a determined amount of radioactive ^{125}I -antigen [Ag*] and an unknown amount of 'cold' antigen [Ag] with a limited amount of antibodies [Ab] in the sample.

The separation of the unbound components is mostly done by precipitation and subsequent centrifugation of the antigen-antibody complex with the help of a secondary antibody and addition of a precipitation catalyst (such as polyethylene glycol). In the competitive assay (fig. 2, 3) the initial amount of antigen in the sample is inversely proportional to the bound radioactivity. As in practice neither K_{eq} nor [Ab] are known, a standard dilution series is used for calibration: bound and free [Ag*] (% bound) are plotted against the concentration of the used antigen in a standard graph.

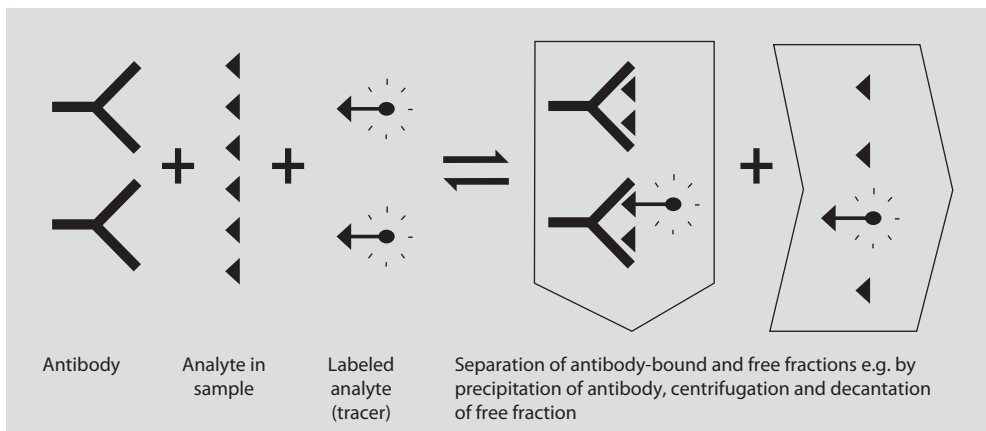


Fig. 2. Principle of the competitive RIA.

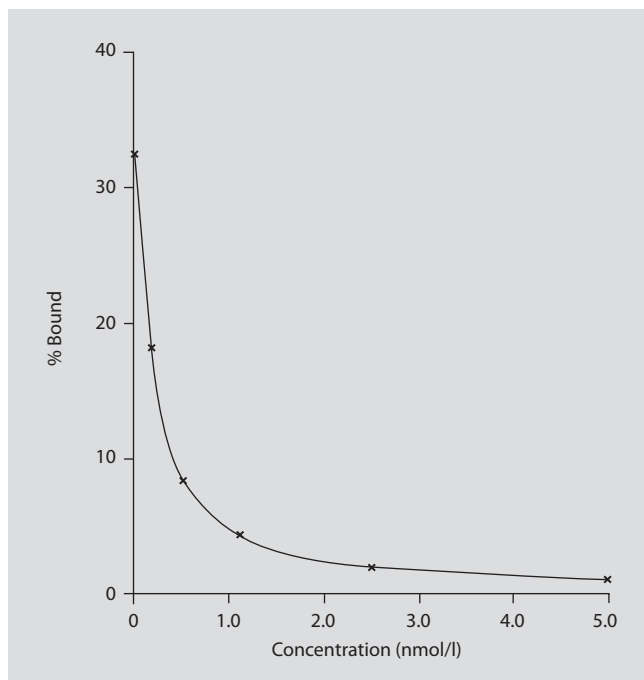


Fig. 3. Calibration (dose-response) graph of the competitive RIA.

Since only the availability of one of the components is limited, noncompetitive assays are less labile than competitive ones with respect to the precision of the used amounts and the purity of labeled antigen and antibody. They are, however, more prone to disturbances through cross-reactions and nonspecific binding.

In 1968, Miles and Hales [14] achieved a further step by developing noncompetitive (single-site) immunoradiometric assays (IRMA) using labeled antibodies.

In the IRMA a marginal surplus of ^{125}I -[Ab*] reacts with the antigen in the sample. All unsaturated ^{125}I -[Ab*] is fully bound to the antigens fixed to the solid phase and subsequently measured. Two-site immunometric or sandwich assays (fig. 4, 15) make use of the fact that antigens generally have several epitopes for antigen bonding. The antigen initially binds the immobilized antibody [Ab1], is then separated and a labeled second antibody [Ab2] is added for detection. The advantages are an approximately ten times higher sensitivity and wider range than RIA. A disadvantage is the so-called high-dose hook effect which sometimes occurs in the binding curve. This effect depends on the quality of the antibodies. A variant of this assay is the highly sensitive coated-tube assay in which the primary antibody is directly fixed to the reaction tube. This simplifies and abbreviates the separation steps and thus reduces costs. However, higher amount of tracer is needed in this type of assay. Here, the bound radioactivity is proportional to the amount of antigens in the sample.

Labeling of Assay Components and Detection Systems

The chemical reaction conditions attending the labeling of antigens or antibodies with ^{125}I (e.g. by the chloramine-T method) or with enzymes or biotin or streptavidin must first be optimized and standardized for each individual antigen. Today, labeling with ^3H , ^{131}I or ^{57}Co has largely fallen out of fashion. Label addition or subtraction should be controlled with a view to avoiding conformational changes of the molecules. Furthermore, radiolysis resulting from the use of radioactive markers can cause serious damage of the labeled protein. The specific activity (mCi/ μg protein) of the tracer must be known. Bound radioactivity is measured in counts per minute (cpm) using a gamma (or beta) counter with inbuilt correction for counting efficiency.

The following drawbacks spoke in favor of developing nonradioactive markers for assay components along with the requisite alternative detection systems: (1) health risks, (2) stringent regulations for radiation protection and removal of radioactive waste and hence higher costs for isotope laboratories, (3) limited automation and speed limitations due to long incubation periods and long counting periods (counting error = $\sqrt{\text{accumulated counts}}$), and (4) limited shelf-life of labeled components due to radioactive decay.

The fact that radioimmunoassays are still in use today can be attributed to the following advantages: (1) assays are robust, (2) their analytical sensitivity can be manually increased by increasing the amounts of specimen or antibodies, e.g. for hormones with very low serum concentration such as sex steroids, (3) sparing in the consumption of limited polyclonal antibodies, and (4) commercial assays are relatively inexpensive, a particularly significant aspect in less-affluent countries.

Nonradioactive methods (table 1) involve labeling of either the antigen or the antibody, depending on the format of the assay. This involves either direct conjugation of the respective component with an enzyme, fluorophore or chromophore, or indirect conjugation through labeling with biotin or streptavidin/avidin. Due to the catalytic

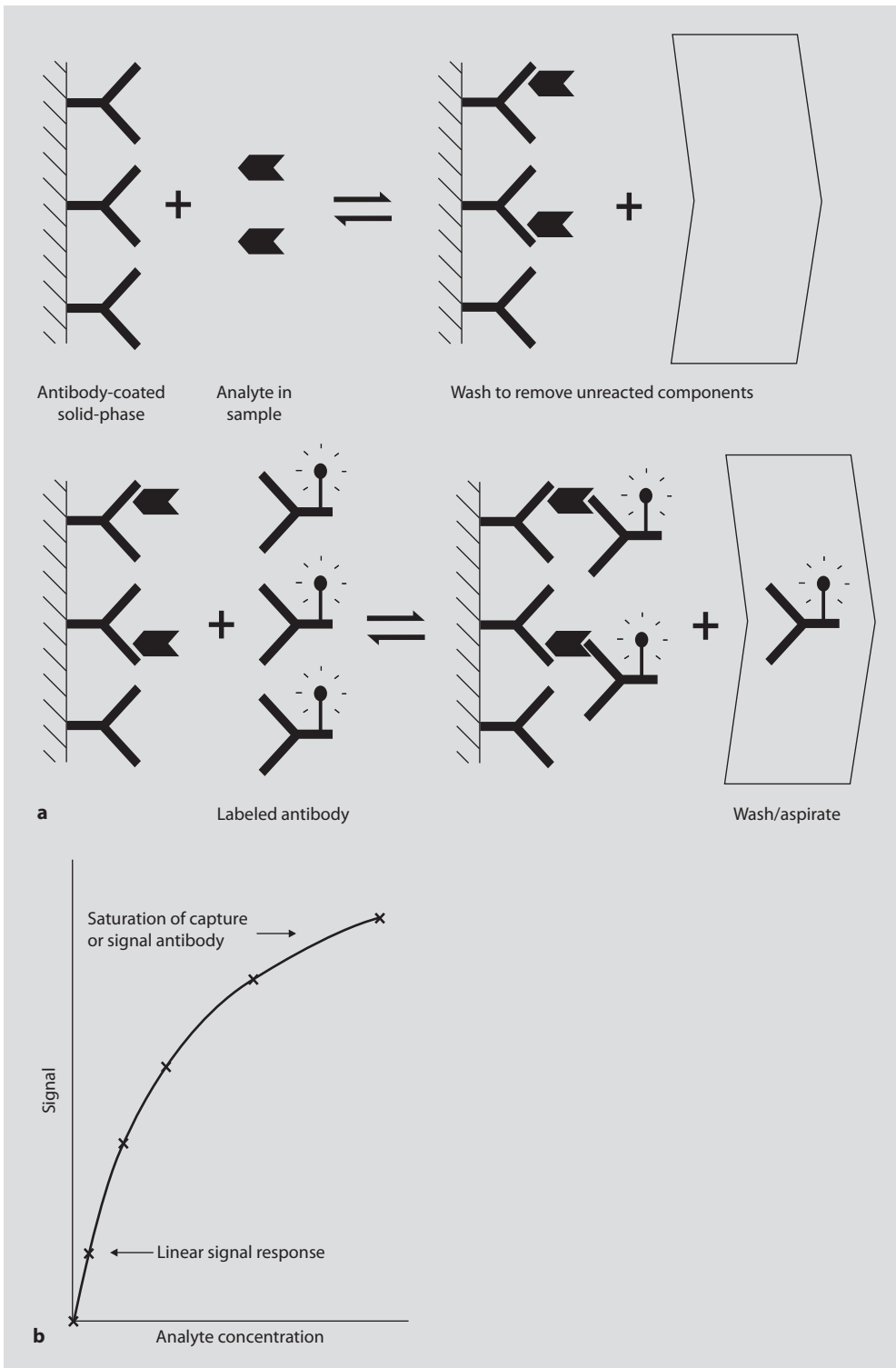


Fig. 4. **a** Principle of two-site immunometric assay (reagent excess). **b** Respective dose-response curve.

Table 1. The most common nonradioactive detection techniques

	Colorimetry	Direct fluorometry	Chemiluminescence	Time-resolved fluorescence
Principal	enzyme-bound antigen competes for the Ab-binding site end-point spectrophotometry of color strength.	comparable to RIA. fluorophores (e.g. fluorescein, rhodamine) act directly as labels or are enzymatically produced. Fluorescence is measured	light emission in the course of a chemical reaction catalyzed by an enzyme. Variant: electrochemiluminescence (ECL)	delayed measurement of the emitted light from the excited fluorophore use of lanthanides with extremely large Stokes' shift, decay times and quantum yields (e.g. Europium)
Assay format and type	manual or automated ELISA 96-well microtiter plates alternative: tab-ELISA	automated assays, EIA, ELISA	automated assays	DELFLIA system as automated assay or manual microtiter plate assay
Substrates	2,2'-azino-bis(ethylbenzothiazoline-6-sulfonate) (ABTS), o-phenylenediamine OPD), 3,3',5,5'-tetramethylbenzidine (TMB)	4-methylumbelliferyl phosphate (4-MUP) stimulation with light of 365 nm wavelength, emission at 448 nm (measurement of the Stokes' shift)	adamantyl 1,2-dioxetane or in ECL: simultaneous oxidation of tripropylamine and ruthenium-II, chelate label of the antigen at the anode, excited Ru-II emits light	None
Sensitivity/range	range of spectrophotometer: - 2.0 OD peroxidase-based assays more sensitive	higher sensitivity than colorimetry through repeated excitation and photomultiplying	high sensitivity wide range	high sensitivity wide range
Advantages	robust simple	sensitivity as high as isotope techniques	minimal background quick signal generation, high specific activity	delayed measurement minimizes interference from fluorescent molecules in the sample
Drawbacks	narrow range, but can be extended by choosing a photometry wavelength next to the absorption peak	interference of the fluorescence with background fluorescence and quenching	only for automated systems	not known

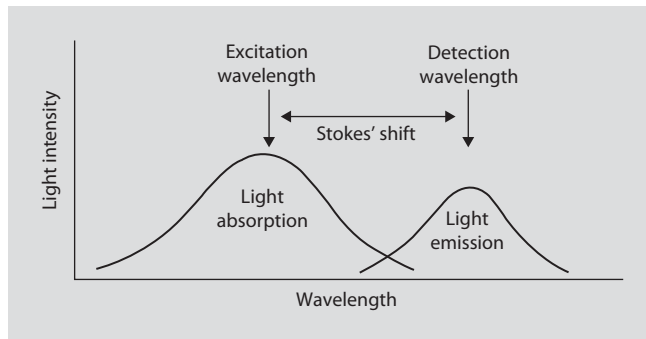


Fig. 5. Principle of fluorescent measurement (Stoke's shift).

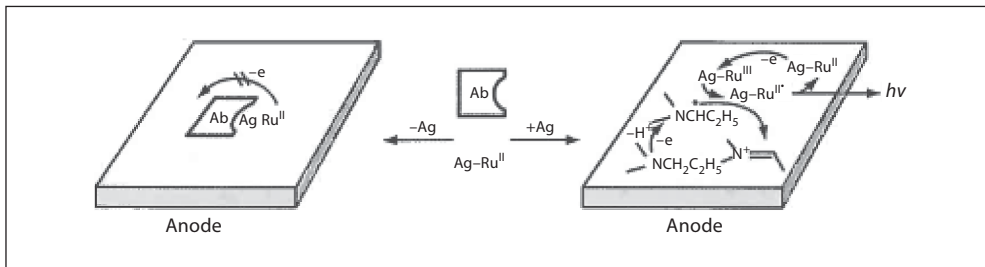


Fig. 6. Homogeneous immunoassay electrochemiluminescence (ECL) principle. (From Wild: Immunoassay Handbook, fig. 12.17.)

effect in the enzyme immunoassay (EIA), it is possible for small amounts of bound enzyme to metabolize large quantities of substrate, depending on reaction time which results in an amplification of the signal generated. EIAs are therefore highly sensitive, but also prone to interference.

Most Commonly Used Enzyme Immunoassays

Alkaline phosphatase, hydrolyses phosphate esters of primary alcohols, phenols and amines. Horseradish peroxidase, which initially oxidizes by reacting with the substrate H_2O_2 , is then in turn able to oxidize other substrates, causing a change in their absorption spectra (fig. 5) and thus forming a colored, fluorescent or luminescent derivative.

Another widely used immunoassay principle is electrochemiluminescence (ECL) (fig. 6). The measurement of ECL, which allows homogeneous immunoassays with great sensitivity, is nowadays widely used in automated immunoassay platforms. It is based on the simultaneous oxidation of a tripropylamine and a ruthenium-II, $Ru(II)$, chelate label of the antigen at the anode. As an intermediate product, an excited state of $Ru(II)$ is generated, which subsequently emits light that is detected and quantified. The immunoassay is carried out by causing the ruthenium label to bind to antibodies

on suspendable beads, or sometimes magnetic beads. The ECL system is used in large immunoassay systems (e.g. Roche's Elecsys®).

Small antigens often undergo steric changes through conjugation with enzymes. To avoid this, such antigens can be conjugated with a spacer like biotin, a vitamin of only 244 Da, which couples the enzyme by way of the linkers avidin or streptavidin. The egg protein avidin and the bacterial (*Staphylococcus aureus*) streptavidin have an enormous affinity (10^{15} l/mol) to biotin.

Separation Techniques of Solid from Fluid Phase in Heterogeneous Assays

The separation of the antibody-bound antigen from the free antigen fraction is a decisive step in heterogeneous immunoassay techniques, having a major influence on the accuracy and reproducibility of the assay. A distinction is made between fluid-phase and solid-phase separation techniques. Which method is used depends on the analyte, amongst other factors. Assays in which one component is coupled to a solid phase are easier and quicker to perform and are therefore by far the most frequently used.

In fluid-phase assays, especially in RIA, the bound and unbound hormone fractions are separated by precipitation techniques using polyethylene glycol or secondary antibody or a combination of the two. The high-molecular-weight complexes formed hereby can easily be pelleted by centrifugation and the fluid phase can subsequently be decanted. The technically demanding and time-consuming gel filtration and electrophoresis techniques are hardly used in clinical practice. The same applies to absorption on dextran-coated charcoal for steroid measurements.

In solid-phase assays, either antigen or antibody is bound to glass, plastic or magnetic beads (microparticle-capture enzyme immunoassay, MEIA) or large spheres. Another widespread technique is immobilization on polystyrene tubes (coated tube assay) or 96-well microtitration plates (ELISA) (fig. 7, 8). A variant in terms of performance is available in the tab ELISA which uses an anti-IgG-coated tab to bind the Ag-Ab complex. Automated processes sometimes make use of sophisticated phase separation techniques, e.g. evasion of the fluid phase through rotation of the test tube (fig. 9), for instance in the Siemens Immulite® system [16].

Solid-phase assays are generally more sensitive and less dependent on the affinity of the antibody than fluid-phase assays since a stationary layer forms above the solid phase in which the probability of a reassociation between antigen and antibody is greater than in the fluid phase, even in cases of antibodies with low affinity.

Calibration Curve Fitting

A calibration curve is used to determine the analyte concentration in samples from the strength of the signal produced in an immunoassay. It represents a plot of calibrator

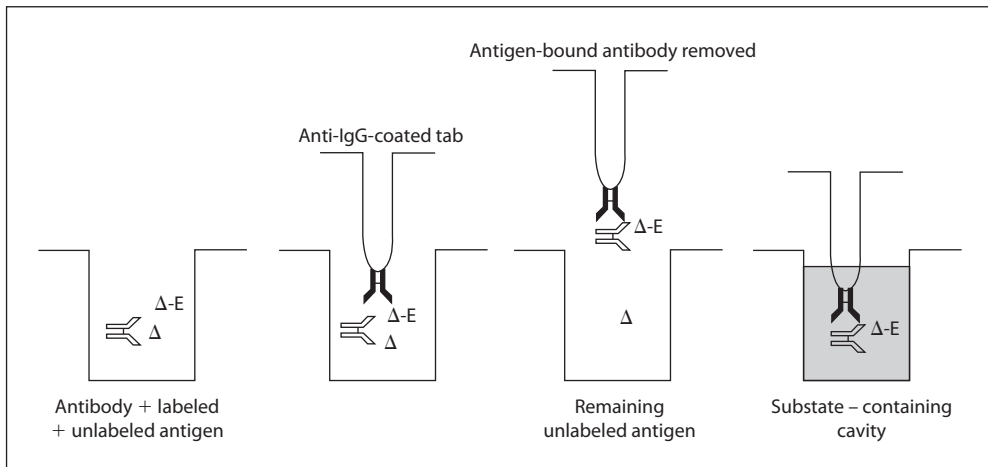
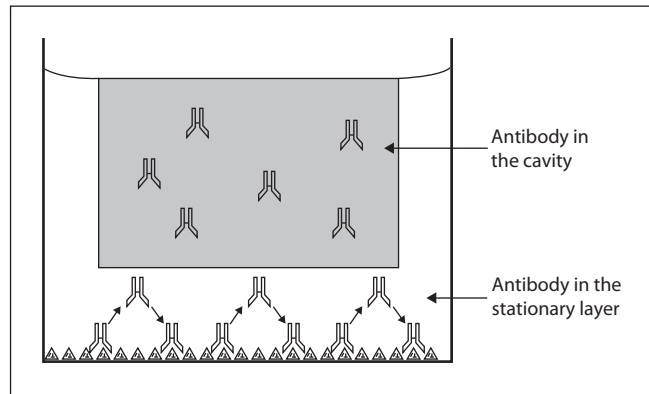


Fig. 7. Principle of microtiter tab-ELISA.

Fig. 8. Increased probability of a reassociation in an assay between bound antigen and antibody in the stationary layer above the solid phase.



concentration against signal level. Nowadays, the calculation process is commonly computerized. The computer calculates a mathematical algorithm, which has to match the calibration curve as closely as possible across the concentration range (fig. 10).

Instead of the two variables (analyte concentration and signal level) dose-response metameters are often used, which are functions of these variables. For example, the logarithm of the analyte concentration is plotted against the percent bound. If it is possible functions linearize the plot in order to enable the use of a simple linear regression analysis. This function, however, alters the error structure. Thus, ideally weighted regression should be used. In principal, the various curve fitting techniques are categorized in two ways [17]: Empirical or interpolatory methods, and regression methods.

Frequently used curve fitting methods are the linear interpolation, spline fitting, polynomial regression, and the logit-log transformation method.

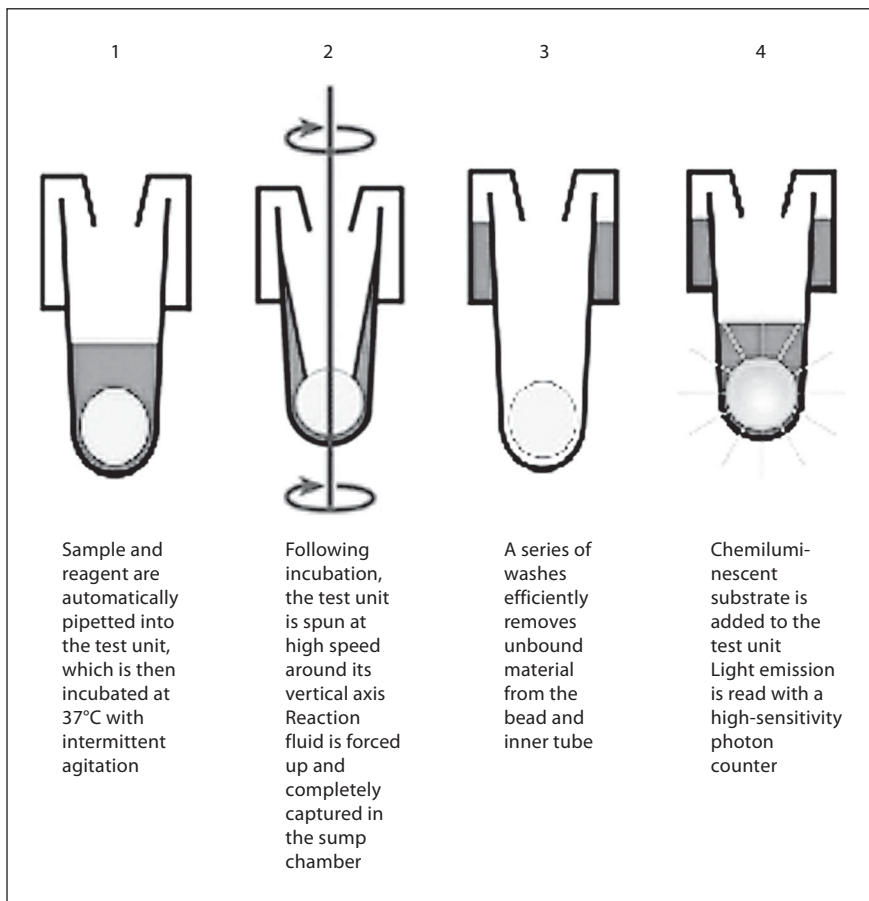


Fig. 9. Specific spinning technology to remove fluids in the Siemens IMMULITE system.

Homogeneous Immunological Techniques

Agglutination assays, which nowadays are rarely used, require no separation step and are therefore suited for automation. They are based on the immobilization of the antibody on cells (e.g. sheep erythrocytes) or latex particles. Upon contact with a ligand agglutination takes place. Nephelometry is likewise based on the formation of high-molecular complexes, with light scattering serving as a quantitative measure.

Technical Trends in Immunoassays

As mentioned above miniaturization, automation, new detection systems, multiplexing on beads (e.g. Luminex® technology), chips or arrays are the major trends of making immunoassays more efficient, more rapid and cheaper. They are mainly based on the classical antigen-antibody interaction.

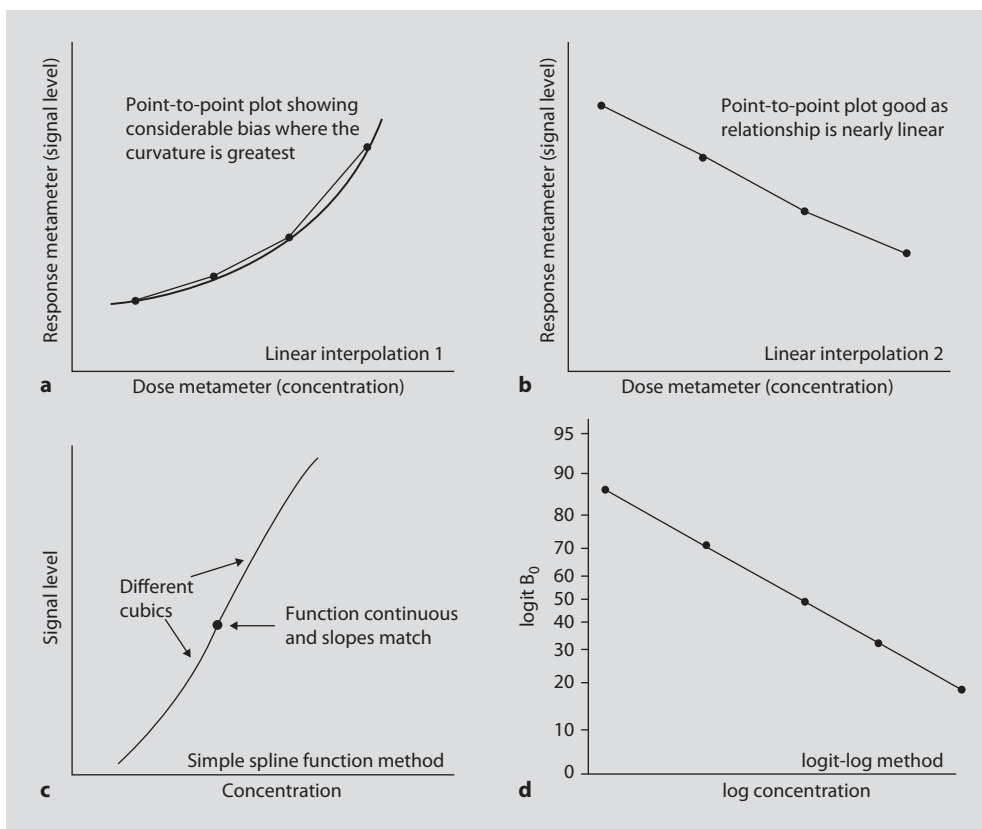


Fig. 10. Types of calibration curve fittings. **a** Linear interpolation 1. **b** Linear interpolation 2. **c** Simple spline function. **d** logit-log method.

An elegant variant of multiplexing, miniaturized (nanoliter scale!) immunoassay systems is the Gyros® system, which is based on a disc designed in the format of a compact disc (CD). The CD contains microcapillary segments with integrated microcolumns filled with streptavidin-coated beads to couple biotinylated proteins.

The Gyros system (fig. 11) is an open, parallel nanoliter microfluidic analysis system [18] which allows the development of new assays if suitable biotinylated antigens are available. The streptavidin-coated beads are activated with biotinylated capture reagents as the first step in the immunoassay. Capillary action is used to draw liquids into a distribution channel, filling a volume definition chamber. A hydrophobic barrier prevents the liquid from moving further into the microstructure. Spinning the CD generates centrifugal force causing the distribution channel to empty, leaving behind a precisely defined liquid volume. A second spin, at higher speed, creates a g force sufficient to drive the liquid over the hydrophobic barrier and through the

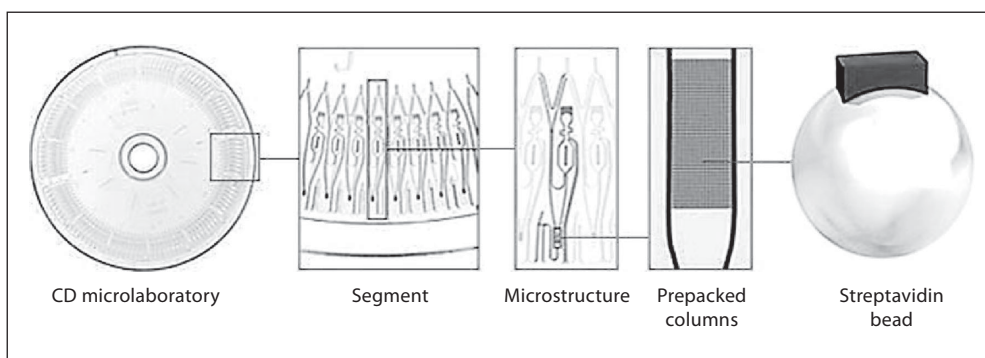


Fig. 11. Gyros disc formate microcapillary system.

capture column. Predetermined spin speeds generate the exact flow rate required for each step, i.e. to capture the maximum amount of protein while minimizing reaction times. As the CD spins, samples are processed in parallel under uniform conditions. After automatic sample processing, the captured protein is quantified using a highly sensitive, laser-induced fluorescence detector.

Protein and antibody arrays, which are powerful techniques in terms of throughput, flexibility, speed and consumption of low amounts of sample may replace the singleplex technologies in the future. Either antigens or antibodies are immobilized. In open systems, specifically customized sets of analytes can be assembled to measure whole biomarker panels of parameters of a certain medical indication. However, today they are mainly used in research and screening approaches, e.g. in the screening of IgE-binding proteins in allergy [19].

The use of biosensorics like the principle of plasmon resonance [20], which generate a physical signal from any kind of highly specific interaction between biomolecules (antibody-antigen, DNA-DNA, DNA-RNA, DNA-protein or receptor and ligand) are still restricted to research use.

For small biomolecules, e.g. steroids or peptides, advanced physical methods of mass spectrometry that are highly specific and sensitive are often used in combination with liquid chromatography (LC/MS; MALDI). These methods are discussed in detail later in this book.

Performance and Validation of Immunoassays: Preanalytical Phase

The preanalysis phase begins with a decision to carry out a laboratory test on a patient and ends with the delivery of the required specimen to the laboratory.

Technical Validation of Immunoassays

To be able to assess the measuring results the pathologist or clinical chemist requires knowledge of certain technical characteristics of the assay employed, as will be

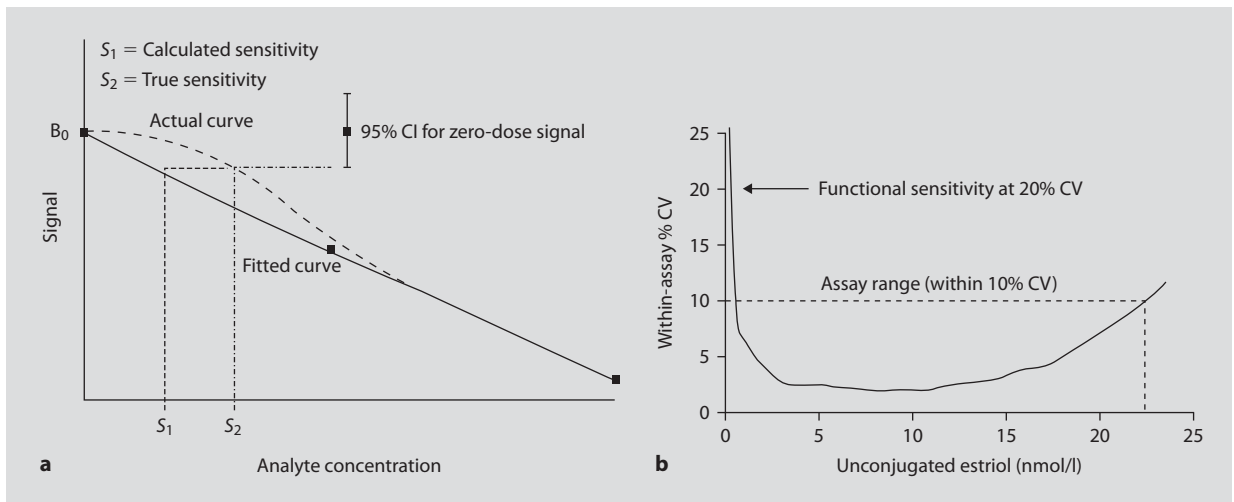


Fig. 12. Analytical (a), and functional (b) sensitivity of immunoassays.

described in the following. In making a diagnosis, the clinician needs to have maximum confidence in the technical reliability of the results.

The analytical sensitivity (fig. 12a) of an assay is generally determined by repeated tests on a blank sample. It is defined as three times the standard deviation from the mean analyte concentration found in the blank test series. In practice there is often a discrepancy between the fitted calibration curve and the dose-response curve in the measuring range below the calibrator, and calculated sensitivity (S_1) and true sensitivity (S_2) therefore often differ quite significantly.

A more practicable approach is therefore to determine the functional sensitivity (fig. 12b), which is defined as the smallest concentration for which the coefficient of variation (CV) is below a set figure, usually 20%. For this purpose, a precision profile must be prepared showing precision (y-axis) as a function of concentration.

Precision (fig. 13) is a dimensionless quantity which gives the degree of agreement between at least two successive measurements; it is limited by random errors. Its opposite is imprecision, which is expressed as the standard deviation or coefficient of variation (CV) of the results of a series of at least 10 successive measurements. A distinction is made between the percental intra-assay CV of successive measurements within a test series and the percental interassay CV of the results of at least 10 different test series, laboratories or measuring days: $\%CV = (SD/mean) \times 100 = \text{relative SD}$.

While double measurements as they are commonly used in laboratories lead to no statistically significant improvement of precision, they make it easier to recognize outliers, especially in the case of manual methods. Therefore, automated methods often perform single measurements only. The assay drift describes systematic errors which only become apparent over time such as trends in the results of measurements


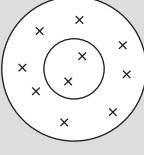
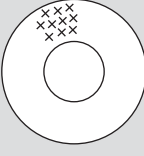
Accuracy	Precision		Error type
Optimum	Optimum		-
Good	Bad		Coincidental error
Bad	Good		Systematic error
-	-	Wrong target used	Gross error

Fig. 13. Illustration of accuracy and precision (target model).

performed on standards or calibrators over months or years (see chapter: quality assurance).

Accuracy (fig. 13) is defined as the degree of agreement between the best estimate of a quantity and its true value (result of an immunoassay as compared with the value obtained with a reference preparation. Accuracy is limited by systematic errors. The following table illustrates the nature of accuracy and precision.

The detection limit is the smallest concentration of an analyte which produces a measuring signal that can be distinguished from that of a blank sample with a defined statistical reliability. Most assays perform reliably in a measuring range beginning 10 blank-value SDs above the detection threshold (fig. 12).

The specificity of an immunoassay is its ability to selectively determine the concentration of a single component (e.g. thyroxin) within a mixture of chemically related substances (e.g. serum). Most assays are multistage, the specificity of which is at best as good as their most selective step. Specificity is limited by antibody cross-reactivity (fig. 14) which can be determined by spiking tests on chemically related substances. Interfering companion substances may also detract from the specificity of an assay.

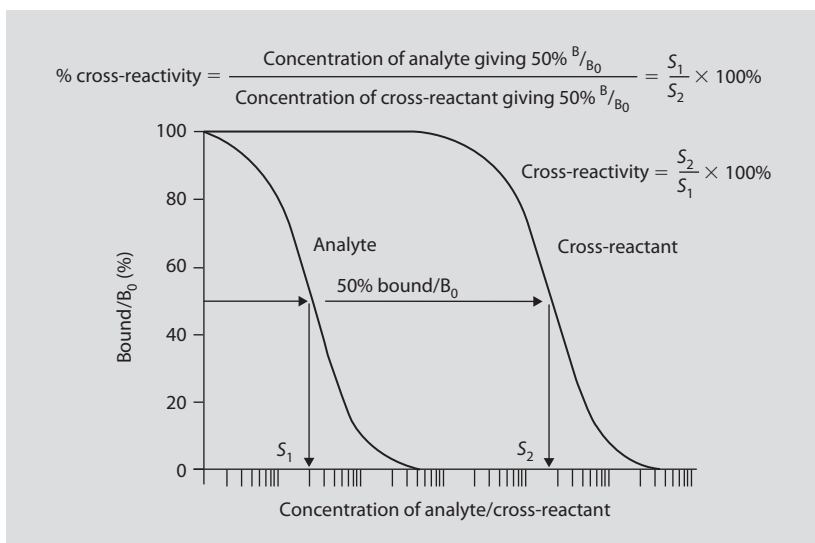


Fig. 14. Crossreactivity of an immunoassay.

Recovery (expressed as a percentage) expresses the proportion of added analyte which can still be detected by the measurement. It is defined as the ratio of the incremental increase in measured analyte concentration to the predicted increase after addition of a known amount of analyte to the sample. It is usually determined by spiking blank samples or samples containing three different low-range concentrations of analyte. The recovery gives an indication of the quality of calibration of an assay and also reveals matrix effects. Studies on the parallelism of dilution are based on precisely defined standards (e.g. an IRP). They provide information on the affinity of the antibody employed and also on matrix effects, since the matrix is being diluted as well. It is recommended to use fractional (100, 90, ..., 10) rather than doubling dilutions (100, 50, ..., 3, 125, ...) to avoid pipetting errors.

In order to compare two assays designed to measure the same hormone or growth factor one has to apply both to as many standards as possible, covering the entire working range in the process. Then the results of the two methods are plotted against each other and analyzed graphically and mathematically (fig. 15). The Pearson product correlation coefficient, which can be inferred from the scatter plot, is prone to interference and does not give a true picture of the data distribution. The scatter produced by most assays is dependent on concentration. Thus, the difference plot presented by Bland and Altman [22] yields proportional information on the reliability of an assay also in the low concentration range. Which regression method is most suitable for calculating the intercept and slope must be decided case by case. Linnet [23] has written a good review article on this issue.

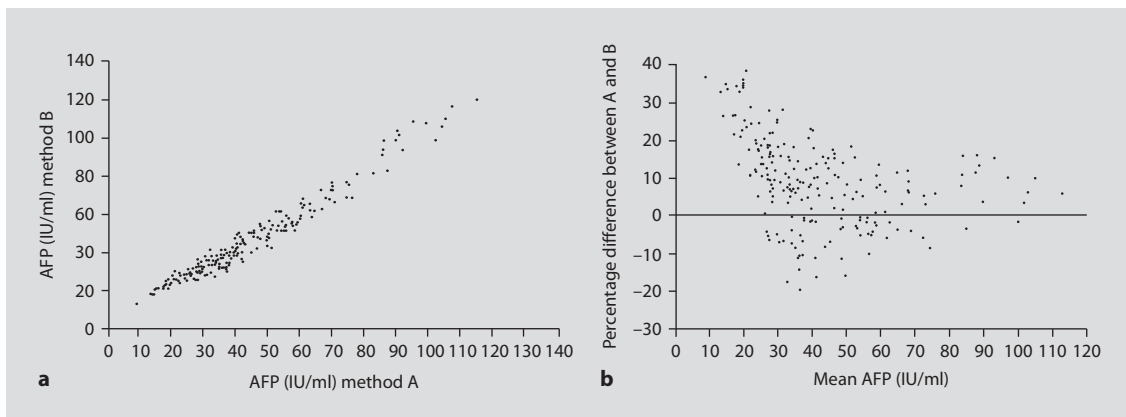


Fig. 15. Scatter plot (a) and Difference plot (b) for comparison of two assays.

Immunoassay Quality Management and Quality Assurance

The aim is for clinical laboratories to only supply results which permit meaningful estimates of the true value and show minimum scatter. In addition to the permanent internal quality control inherent to every assay, there are also regular external quality inspections which provide information on the precision and accuracy of measurements as compared with those of other laboratories. All measures taken must be accurately documented in suitable computer programs and must also be understandable to outsider observers such as calibration or licensing authorities.

Internal quality control includes the regular maintenance and operational testing of pipettes and instruments, consistency in the use of standard operation procedures and in reagent quality, and of course the use of specially trained staff. The best way to ensure the certifiability of operations in a medical laboratory is to have the laboratory accredited according to an international standard for quality management (DIN ISO 9001, ISO 15189, ISO 17025, CLIA, or GLP).

However, the laboratory staff has virtually no influence on the preanalysis phase (table 2). In the event of deviations the troubleshooting will initially be limited to variable parameters (e.g. quality of commercial assays or of the control preparation). Internal quality control procedures are usually based on measurements of commercial control preparations followed by longitudinal evaluation. Many laboratories additionally have in-house quality controls such as a standardized serum pool. If this is used over a long period of time, the analyte must be guaranteed to be sufficiently stable, which is checked by repeated testing. In contrast to commercial preparations which come in successive charges, a serum pool will often be available over years or even decades. Commercial control preparations usually have a matrix which differs from those of samples submitted to the laboratory so that matrix-related systematic errors therefore may often go undetected.

Table 2. Preanalytical steps and responsibilities

Steps	Remarks	Supervised by
Appointment and patient preparation	time of specimen collection nutritional status of patient co- medication requirements for laboratory parameters	physician
Choice of the primary specimen, e.g. blood, 24-hour urine, spontaneous urine, saliva, feces, liquor	suitable specimen storage (e.g. with anticoagulants, such as EDTA, citrate, LiCl)	physician
Unequivocal labeling and documentation of specimen	type of sample, date of collection, name (code), age, sex, tentative diagnosis	physician
Preparation of secondary specimen (e.g. serum)	e.g. centrifugation, portioning	physician /laboratory
Intermediate storage and shipment to laboratory	laboratory must be competent to deal with the parameter in question; cooling/freezing, knowledge of possible mechanism of decay of the analyte	physician /laboratory
Specimen processing	documentation of specimen and patient data on a computer in a regulated environment (e.g. GLP), immediate portioning, preservation if necessary and submission for laboratory testing storage of residual specimen as backup	laboratory

As a general rule, each assay should be controlled at three different concentration levels (fig. 16). Assays often also come with technical calibrators, but these should be used in addition and should not replace control preparations. Results from control preparations should be evaluated immediately after completion of the assay, since the outcome decides on whether the patient's data can be released. The results are documented in control charts, which today are available in special quality control computer programs or included in the software of laboratory information managing systems.

Here, time, expressed in terms of days, is plotted against a concentration interval measured. The central line represents the target value for the sample, while the upper and lower warning limits represent two, and the upper and lower control limits three standard deviations from the target value. These analyte specific limits are often specified by the manufacturer, but they must nevertheless be verified by evaluation over a preliminary testing period of at least 20 assay days and corrected as necessary. Subsequently measured values should then oscillate within these limits. As a rule, the sum of three standard deviations should not exceed 15% of the mean value. In the event of values

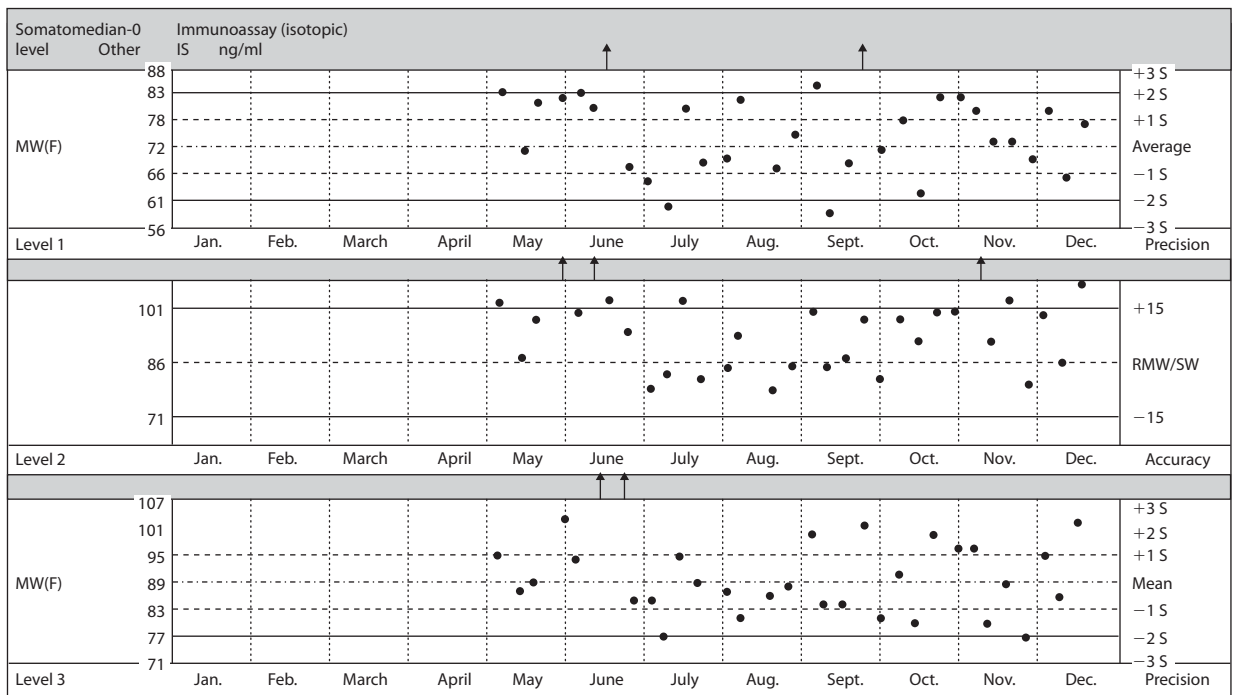


Fig. 16. Example for a control chart of IGF-1.

exceeding the 3S (SDs) limit, seven successive values showing a continuous upward or downward trend, or of seven successive values coming to lie on one side of the central line there is probably a systematic error. Computer programs in such cases will give an optical or acoustic alert signal and advise to reject the sample and eliminate the error.

External quality control requires spot-check measurements of one sample pair per analyte several times a year. It also involves an assessment of the measuring techniques employed by the laboratory and a comparison with other measuring techniques commonly used for the parameters in question as well as with the results of numerous other laboratories participating in the quality control program (inter-laboratory testing). Participation in such programs is not mandatory in all countries. The control preparation is made available to the participating laboratories by state-certified centers (in Germany, e.g. by DGKL, Instand e.V.), and the results are sent back to the center for evaluation. The outcome is usually communicated to the participants in the form of a Youden plot (fig. 17) (see also <http://www.medcalc.be/manual/youdenplot.php>). If the results found by the laboratory are within a predefined tolerance range around the target value, the participant is presented with a certificate confirming the accuracy of his measurements.

Value pairs are read as coordinates and entered into the plot as dots each representing a laboratory. All results which fall within a predefined square are considered to

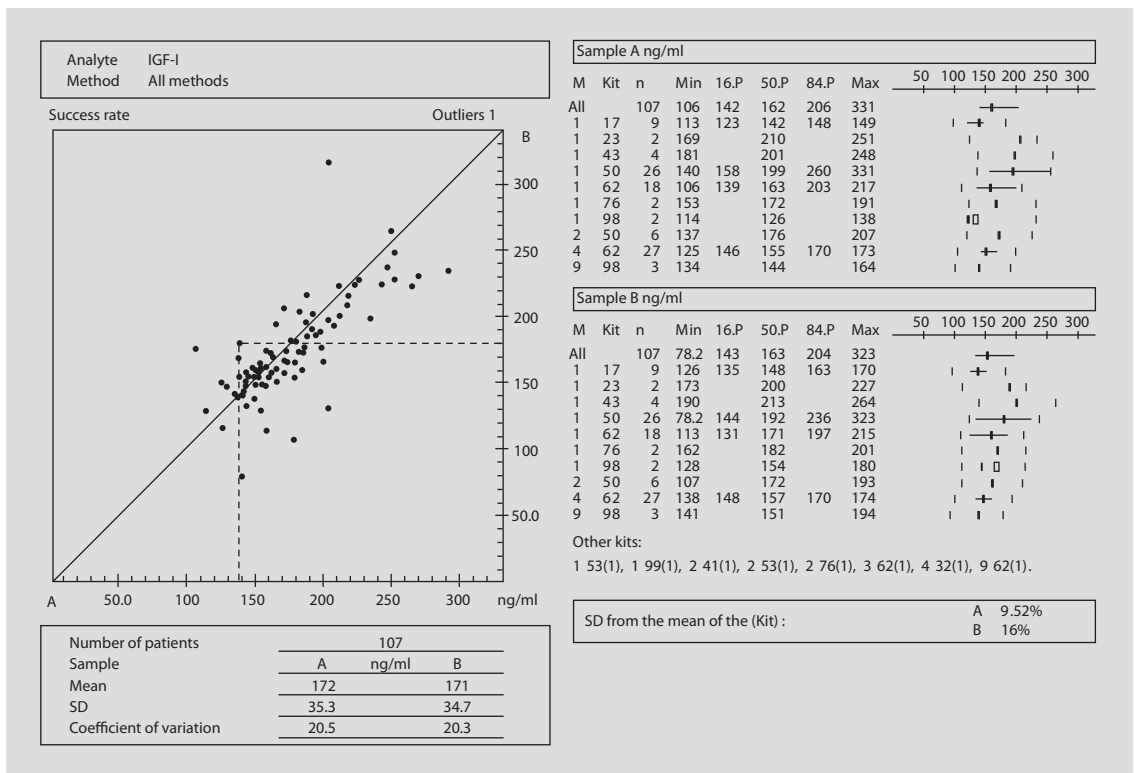


Fig. 17. Youden plot for external quality control.

have met the requirements. Values located outside the square but close to the diagonal are indicative of systematic errors. Values far from the diagonal are grossly imprecise. With some programs, the participants are informed on the number and types of methods employed by all participants. Technical and method-specific calibrators are also used for the validation of hormone assays.

Clinical Validation of Hormone Measurements and Diagnosis

Immunoassays are used for diagnosis, quantification (example: TSH in screening of the newborn) and monitoring (example: IGF-1 measurement before and after growth hormone therapy) of endocrine diseases. Before the physician can make a clinical interpretation of the results they must have been fully validated by the pathologist, clinical chemist and laboratory staff. Diagnoses can only be made on the basis of accurate and precise measurement data. Laboratory data can only be interpreted properly if due consideration is given to other clinical parameters of the patient as well as to factors that might have influenced the assay (e.g. medication, nutrition, weight, cachexia, infections, circadian rhythms etc.). Some laboratory parameters such as a basal growth hormone value (pulsatile secretion) are of little informative value when considered alone.

Moreover, every measuring method must be clinically validated. First and foremost, laboratories must determine their own reference intervals for a certain method from which limits for determining whether a given value is normal or probably pathological can be derived statistically. Furthermore, to be able to verify tentative diagnoses (e.g. thyroxin deficiency), it must be known when and how reliably an abnormally low T4 value can predict hypothyreosis. The above-mentioned clinical aspects are discussed in detail in this book.

Determining Reference Intervals

Prerequisites: reference intervals give the statistical mean and range of scatter of the concentration of a parameter (e.g. a hormone), i.e. its interindividual variability within a reference population. Only by comparing the measured value with a suitable reference interval can pathological deviations be identified. The reference population needs not necessarily be healthy, as is demonstrated by reference curves for bodily growth in girls with Turner syndrome [24]. An appropriate number of samples from totally healthy subjects is often hard to obtain in particular from very young children, for ethical and practical (only low amount of blood can be drawn) reasons. Furthermore, environmental factors may also influence the concentrations of the parameters that are measured. The concept of 'normal values' can therefore be regarded as obsolete. However, it is necessary for reference populations to be composed of precisely defined reference individuals with due consideration to age, sex, body weight, pregnancy, state of health and all other factors that might have an influence on the parameter in question. In the time from birth to adulthood, the hormone balance undergoes drastic changes as the body grows and the gonads develop. Reference intervals for hormones in children must therefore be defined for small age intervals and according to different stages of sexual development. Reference intervals also depend on ethnic and cultural factors (e.g. dietary habits and physical activity). For example, it is not permissible to apply reference intervals defined for Japanese children directly to Central European children.

Reference values are also assay-specific (fig. 18). Ideally, every laboratory should determine its own reference intervals. The minimum requirement is that results should always be assessed on the basis of reference values obtained with the same type of assay and for a representative reference population.

Statistical Methods

The most common approach is to perform a parametric calculation of the reference intervals of a normal distribution of either the values of reference individuals directly or of the output of a logarithmic normalizing transformation performed on

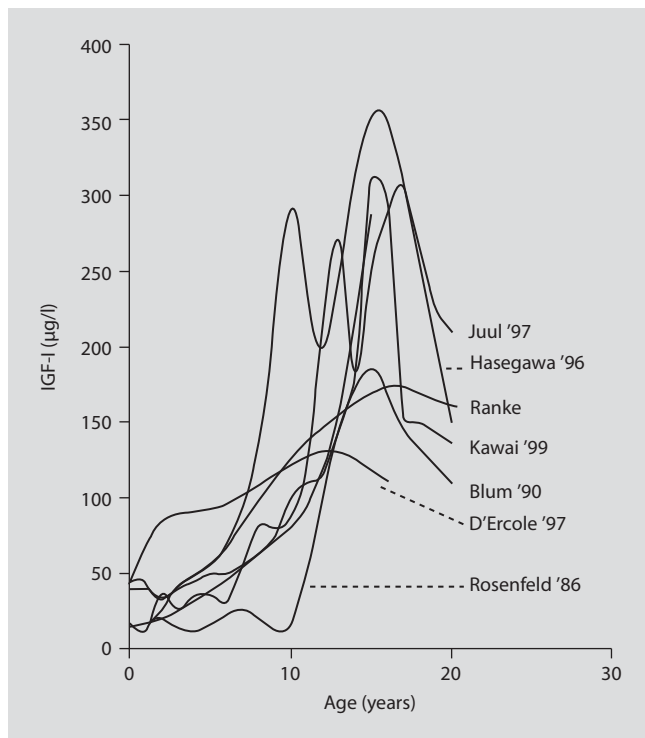


Fig. 18. Comparison of 50th percentiles for IGF-1 evaluated by different commercial and in-house assays.

those values. The exact procedure is set forth as a proposal in the Guidelines of the International Federation of Clinical Chemistry (IFCC) [25].

If there is a sufficient number of individuals per group, the central 95% CI is used (2.5th to 97.5th percentile), corresponding to the mean value of ± 1.96 SD. This method is relatively safe when it comes to assessing individual values, and it excludes extreme outliers. Another robust method makes use of the median to estimate the arithmetic mean and defines the SD as the difference between the 90th and 10th percentile divided by 2.563. The age curves for the arithmetic mean and standard deviation are commonly smoothed by means of a spline-fitting algorithm. Depending on which cut-off is postulated for the assessment of patient data, one can then calculate different percentiles and depict them graphically.

Interpretation of Laboratory Test Results

As already mentioned, an assessment of laboratory data by the physician presupposes that the conditions at the time of sampling (time of day, preanalysis), the patient's status (dietary habits, medication, if any) as well as possible interfering factors are known and verifiable. Furthermore, it is the physician's responsibility to ensure that

Table 3. Scenario to demonstrate the basic tools to perform clinical interpretation of laboratory data

	Disease status		Total
	diseased (n)	healthy (n)	
Test positive	95 (a)	2 (b)	97 (a+b)
Test negative	5 (c)	98 (d)	103 (c+d)
Total	100 (a+c)	100 (b+d)	–

the parameters indicated on the laboratory sheet suffice to verify the tentative diagnosis. For financial (and sometimes ethical) reasons, the range of parameters should be limited to what is diagnostically necessary.

When evaluating laboratory findings, the pathologist and clinical chemist must use the information on the patient that is available to them to decide whether values are to be regarded as ‘normal’, ‘pathological’, ‘critical’ or ‘ambiguous’. In deciding they must also consider the reference intervals and corresponding cut-off limits distinguishing pathological from normal values. In addition to this information, the pathologist or clinical chemist will often also provide the physician with recommendations for further tests, and in some cases even for an appropriate therapy.

Diagnostic procedures involving single measurements of hormones or simulation tests must be examined by the clinician in terms of their clinical sensitivity, specificity, relevance and predictive value.

This will be illustrated in the following scenario (table 3) by comparing a cohort of 100 patients with confirmed endocrine disease with a cohort of 100 healthy patients. All patients are immunoassayed for a defined hormone parameter:

The clinical sensitivity measures how well the test detects those patients with the endocrine disease. It is expressed as the proportion of true positives that are correctly identified out of the 100 persons with the disease:

$$\text{Sensitivity} = a/(a + c); 95\% \text{ in the example.}$$

The clinical specificity measures how well the test correctly identifies those patients who do not have the disease. It is expressed as the proportion of true negatives that are correctly identified out of the 100 healthy patients:

$$\text{Specificity} = d/(b + d); 98\% \text{ in the example.}$$

The false-positive rate is the proportion of demonstrably unaffected individuals falsely identified as having the disease:

$$\text{False-positive rate} = 1 - \text{specificity.}$$

The decisive criterion in clinical practice, however, is how reliably a positive result can predict that the endocrine disease really is present. This quantity gives the

likelihood in the test population of a positive test result being true. The positive predictive value (PPV) is the likelihood that a patient with a positive test result actually has the endocrine disease, i.e. the proportion of patients with positive test results who are correctly diagnosed:

$PPV = a/(a + b)$; 97.5% in the example.

The negative predictive value (NPV) is the likelihood that a patient with a negative test result does not have the endocrine disease:

$NPV = d/(c + d)$; 95.1% in the example.

In contrast to the incidence, which gives the proportion of new cases of a disease (e.g. diabetes mellitus) within a population, the prevalence gives the proportion of persons in that population who already have the disease at a given point in time. In the above example the prevalence of the disease would be:

$P = \text{diseased persons} / (\text{diseased persons} + \text{healthy persons}) = 50\%$.

This value does probably not represent the true prevalence of the disease in the population, which is normally far lower. The clinical utility of a diagnostic test, and hence of PPV and NPV, can be severely restricted by the prevalence of the disease in question.

On the Critical Use of Laboratory Data in Clinical Diagnostics

The number of laboratory parameters that can be measured and for which reimbursement is provided has increased considerably over the past years, but is now limited by the financial limitations of the public health care systems. This makes it increasingly difficult for physicians to select the right parameters for diagnosis.

In matters of quality control, the physician must be able to trust in the reliability of the laboratory. Quality management systems are used in every diagnostic laboratory today. Nevertheless, he should not trust laboratory results blindly, since errors can also go undetected in the laboratory. In cases of doubt, improbable measuring results must be checked, or it may even be necessary to generate a new sample. This presupposes a continuous dialogue between the physician and the laboratory. At the same time, for cost reasons, laboratory diagnostics should be reduced to the minimum of what is necessary. To avoid false-positive, or worse still, false-negative results, physicians must command a wide experience and continuously train their knowledge.

Last but not least, it should be pointed out that a clinical diagnosis can be accomplished and refined, rather than be replaced by measuring meaningful laboratory parameters.

References

- 1 Evans HM, Simpson ME, Marx W, Kibrick E: Bioassay of pituitary growth hormone: width of the proximal epiphyseal cartilage of the tibia in hypophysectomized rats. *Endocrinology* 1943;32:13–16.
- 2 Bristow AF, Jespersen AM: The Second International Standard for somatropin (recombinant DNA-derived human growth hormone): preparation and calibration in an international collaborative study. *Biologicals* 2001;29:97–106.
- 3 Robinson CJ, Gaines Das R, Woollacott D: The first international standard for human leptin and the first international standard for mouse leptin: comparison of candidate preparations by in vitro bioassays and immunoassays. *J Mol Endocrinol* 2001;27:69–76.
- 4 Persani L, Ferretti E, Borgato S, Faglia G, Beck-Peccoz P: Circulating thyrotropin bioactivity in sporadic central hypothyroidism. *J Clin Endocrinol Metab* 2000;85:3631–3635.
- 5 Binder G, Benz MR, Elmlinger M, Pflaum CD, Strasburger CJ, Ranke MB: Reduced human growth hormone (hGH) bioactivity without a defect of the GH-1 gene in three patients with rhGH responsive growth failure. *Clin Endocrinol (Oxf)* 1999;51:89–95.
- 6 Ishikawa M, Nimura A, Horikawa R, Katsumata N, Arisaka O, Wada M, Honjo M, Tanaka T: A novel specific bioassay for serum human growth hormone. *J Clin Endocrinol Metab* 2000;85:4274–4279.
- 7 Kiess W, Butenandt O: Regulation of GH binding to specific cellular receptors in vitro—a new model of growth regulation in vivo. *Horm Metab Res* 1987;19:171–176.
- 8 Baumann G: Growth hormone binding proteins and various forms of growth hormone: implications for measurements. *Acta Paediatr Scand* 1990(suppl); 370:2–80.
- 9 Asai D, Shimohigashi Y: The assessment of xenosteroids by competitive receptor binding assays. *Nippon Rinsho* 2000;58:2486–2490.
- 10 Strasburger CJ, Wu Z, Pflaum CD, Dressendorfer RA: Immunofunctional assay of human growth hormone (hGH) in serum: a possible consensus for quantitative hGH measurement. *J Clin Endocrinol Metab* 1996;81:2613–2620.
- 11 Lammert A, Kiess W, Bottner A, Glasow A, Kratzsch J: Soluble leptin receptor represents the main leptin binding activity in human blood. *Biochem Biophys Res Commun* 2001;283:982–988.
- 12 Lassarre C, Duron F, Binoux M: Use of the ligand immunofunctional assay for human insulin-like growth factor (IGF) binding protein-3 (IGFBP-3) to analyze IGFBP-3 proteolysis and IGF-I bioavailability in healthy adults, GH-deficient and acromegalic patients, and diabetics. *J Clin Endocrinol Metab* 2001;86:1942–1952.
- 13 Berson SA, Yalow RS: Assay of plasma insulin in human subjects by immunological methods. *Nature* 1959;184:1984.
- 14 Miles LEM, Hales CN: Labelled antibodies and immunological assay systems. *Nature* 1968;219:186–189.
- 15 Addison GM, Hales CN: The immunoradiometric assay; in Kirkham KE, Hunter WM (eds): *Radioimmunoassay Methods*. Edinburgh, Churchill-Livingstone, 1970, pp 447–461.
- 16 Babson AL, Olson DR, Palmieri T, Ross AF, Becker DM, Mulqueen PJ: The IMMULITE assay tube: a new approach to heterogeneous ligand assay. *Clin Chem* 1991;37:1521–1522.
- 17 Box FEP, Hunter WG: A useful method for model-building. *Technometrics* 1962;4:301–318.
- 18 Andersson P, Jesson G, Kylberg G, Ekstrand G, Thorsén G: Parallel nanoliter microfluidic analysis system. *Anal Chem* 2007;79:4022–4030.
- 19 Templin MF, Stoll D, Schwenk JM, Pötz O, Kramer S, Joos TO: Protein microarrays: promising tools for proteomic research. *Proteomics* 2003;3:2155–2166.
- 20 Schultz JS: Biosensors. *Scientific American*, August 1991;48–55.
- 21 Gordon Malan P: *Immunological Biosensors*; in Wild D (ed): *The Immunoassay Handbook*. London, Macmillan Press, 1994, p 131 (UK).
- 22 Bland JM, Altman DG: Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;ii:307–310.
- 23 Linnet K: Evaluation of regression procedures for method comparison studies. *Clin Chem* 1993;39:424–432.

- 24 Rongen-Westerlaken C, Corel L, van den Broeck J, Massa G, Karlberg J, Albertsson-Wikland K, Naeraa RW, Wit JM: Reference values for height, height velocity and weight in Turner's syndrome. Swedish Study Group for GH treatment. *Acta Paediatr* 1997; 86:937-942.
- 25 International Federation of Clinical Chemistry Scientific Committee, Clinical Section: Expert Panel on Theory of Reference Values and International Committee for Standardization in Haematology. 5. Statistical treatment on the theory of reference values: determinations of reference limits. *J Clin Chem Clin Biochem* 1987;25:645-656.

Prof. Martin W. Elmlinger
Nycomed International Management GmbH
Thurgauerstrasse 130
CH-8152 Glattpark-Opfikon, Zurich (Switzerland)
Tel. +41 505 44 555 1228, Fax +41 505 44 555 1455, E-Mail martin.elmlinger@nycomed.com