

MY APPROACH

My approach to internal quality control in a clinical immunology laboratory

R J Lock

J Clin Pathol 2006;**59**:681–684. doi: 10.1136/jcp.2005.032292

Clinical laboratory services strive to ensure that the right results are obtained in patients. Decisions on patient management are made on the basis of these results and the maintenance of high-quality output is essential. Precise and accurate results (by the processes of internal quality control (IQC) and external quality assessment (EQA)) and a timely and appropriate service (by means of a laboratory audit, clinical audit, laboratory accreditation and clinical governance) are generated by the delivery of a quality (defined as “a degree of excellence” in the Oxford English Dictionary) service in clinical immunology.

The major objective of quality assurance is to improve the quality of results for uniformity both within and between laboratories, so that an appropriate clinical interpretation can be made on the basis of that result. Within this objective, the role of IQC is to monitor the day-to-day precision and accuracy of a given assay. The role of EQA is broader in that it can compare and contrast different methods, thus also providing educational information.

The control of immunoassay systems can be especially challenging owing to the complexity of the methods used. Variability may be introduced at many levels, including the antigen source (whole tissue, cell extract, purified protein, recombinant protein), antibody detected (isotype, affinity, concentration), antibody detection system (polyclonal, monoclonal, affinity, conjugation (enzyme, fluorochrome) and methodological variations (incubation time, volume, choice of substrate). All these factors are taken into account when things go wrong.

This article suggests an approach to IQC, including the mathematical methods used to monitor precision and accuracy, appropriate reference materials and controls, and some internal checks that help in providing additional information.

Reference materials—the starting point for assured quality

Both IQC and EQA have underlying requirements for central reference materials (often incorrectly called standards) to enable direct comparisons to be made between laboratories. During the 1960s and 1970s, the World Health Organization and the International Standards Organisation outlined the requirements for the preparation of references. The major requirement was for a stable homogeneous material to be available in significant volume that would behave in the same way as the appropriate body fluid.¹ Several categories were defined according to the type of preparation—for example, whether

a numerical value was assigned or not. Table 1 lists some of the reference materials more useful to the clinical immunology laboratory. On a day-to-day basis, the laboratory is more likely to use a “secondary standard” derived from these, either internally or by a manufacturer.

We need to consider some points when using primary reference materials. Firstly, they are in relatively short supply and it is difficult to obtain large quantities for IQC purposes. The simplest, and probably best, way to use these materials is to include them as samples in the assay run. In quantitative assays, we would expect the obtained result to be the assigned level plus or minus an allowance for the precision of the assay. For qualitative assays, only the assigned reactivity should be detected. Here a second point of caution applies, especially with regard to antibodies to extractable nuclear antigens. Additional specificities may be found, particularly if the technology is different from that used in the definition of the standard (eg, enzyme-linked immunosorbent assay (ELISA) *v* immunoprecipitation assay). This often represents a shift in the balance between disease specificity and technological sensitivity.

Equally important are the materials used to monitor performance (quality-control materials). Often, the manufacturer of the assay under consideration will provide these materials, of which many are simply dilutions of the same material that is used as the standard. This is far from ideal, because if there is any deterioration in the standard, there will be a parallel deterioration in the control and the changed calibration may not become obvious for a considerable time. It is good practice to include a control from a separate source (“third-party control”—that is, material from another manufacturer), which, hopefully, will not deteriorate at the same rate. A recent example where this problem occurred was highlighted in the UK National External Quality Assessment Scheme for complement in 2004, where, over a period of about 6 months, it gradually became clear that one manufacturer had a calibration error of about 15% in the C4 secondary standard. It is normal practice to control the assay by using the manufacturer’s controls. Use of a third-party control, separately calibrated, may have identified the calibration error sooner.

Abbreviations: ELISA, enzyme-linked immunosorbent assay; EQA, external quality assessment; IQC, internal quality control

Correspondence to:
R J Lock, Immunology and
Immunogenetics, North
Bristol NHS Trust,
Southmead Hospital,
Bristol BS10 5NB, UK;
robert.lock@nbt.nhs.uk

Accepted for publication
22 December 2005

Table 1 Some reference materials for use in clinical immunology

Analyte	Reference material	Name	Source
ANA—homogeneous pattern	WHO1064	WHO 1st reference preparation for antinuclear antibody	ILBS
ANA—homogeneous pattern	66/233	Antinuclear factor serum, human	NIBSC
ANA—homogeneous pattern/ rim pattern—dsDNA	AF-CDC-ANA#1	ANA homogeneous pattern/rim pattern—dsDNA	CDC
ANA—nucleolar pattern	AF-CDC-ANA#6	ANA nucleolar pattern	CDC
ANA—nucleolar pattern	68/340	Anti-nucleolar factor plasma, human	NIBSC
Antibodies to ENA	AF-CDC-ANA#2	ANA speckled pattern—SSB/La	CDC
Antibodies to ENA	AF-CDC-ANA#4	Anti-U1-RNP	CDC
Antibodies to ENA	AF-CDC-ANA#5	Anti-Sm	CDC
Antibodies to ENA	AF-CDC-ANA#7	Anti-SSA/Ro	CDC
Antibodies to ENA	AF-CDC-ANA#9	Anti-Scl70	CDC
Antibodies to ENA	AF-CDC-ANA#10	Anti-Jo-1	CDC
Antibodies to ENA	WHO1063	WHO international reference human serum for anti-nuclear ribonuclear protein (nRNP) autoantibody	ILBS
Anti-dsDNA	WHO-W680	WHO 1st reference preparation for native (ds)DNA antibody (1985)	ILBS
Anti-centromere antibody	AF-CDC-ANA#8	ANA anti-centromere	CDC
Rheumatoid factor	64/2	1st British standard for rheumatoid arthritis serum	NIBSC
Rheumatoid factor	WHO1066	Human rheumatoid arthritis serum	ILBS
Anti-cardiolipin antibody	97/656	Cardiolipin antibody	PRU
IgG, IgA and IgM C3, C4, CRP	CRM470*	Serum reference preparation	IRMM
Anti-mitochondrial antibody	67/183	Primary biliary cirrhosis serum, human	NIBSC

AF-CDC, Arthritis Foundation; ANA, antinuclear antibody; CDC, Centers for Disease Control, Atlanta, USA; CRP, C-reactive protein; ENA, extractable nuclear antigen; ILBS, International Laboratory for Biological Standards, Netherlands Red Cross Blood Transfusion Service; IRMM, Institute for Reference Materials and Measurement, Belgium; NIBSC, National Institute of Biological Standards and Control, South Mimms, UK; PRU, Protein Reference Unit, Sheffield, UK.

*CRM470, although now the international reference material for many proteins, including immunoglobulins and complement, is not generally available to diagnostic laboratories. Many commercial preparations (secondary standards) are available, calibrated against this material.

Important concepts

1. Precision is the ability to obtain the same result on a given analyte each time a given material is analysed. For any given assay, the precision is lower as the level of the analyte approaches the lower level of sensitivity of the assay—that is, an assay is less precise at the bottom end of the measurable range. Mathematically, precision (or more correctly “imprecision”) is expressed as the coefficient of variation, where coefficient of variation = $SD/Mean \times 100\%$. Precision may worsen as a result of deterioration of reagents, instability of the apparatus or owing to operator problems. Table 2 shows suggested levels of acceptable precision. Whether these levels are achievable varies considerably with the method, analyte, operator experience and clinical requirement. For example, ELISAs for anti-cardiolipin antibodies are difficult to reproduce and often have coefficient of variation $>25\%$. Typically, clinically relevant antibody levels are high.² As a result of this, and to ensure clarity for the user, I suggest laboratories

consider reporting only qualitative results—that is, negative, weak, moderate or strong positive. This approach has achieved higher interlaboratory concordance than the numeric data in the UK National External Quality Assessment Scheme. These figures are based on the author’s experience and are not necessarily applicable to all assays.

2. Accuracy is the ability to measure the analyte correctly on any given occasion.

3. Monitoring analytical sensitivity is essential to prevent reporting false-negative results. This is best shown by considering examples of three different assay systems.

ELISA or solid-phase systems

Most kits provide positive and negative “quality” controls. In kits that provide only single-point calibration (often containing high levels of antibodies), it is advisable to include a control near the lower discrimination point, separating positive and negative (for qualitative assays) or low results and the normal reference range (for quantitative assays). This may be either an in-house control or a suitable dilution of the provided positive control. Ideally, it should sit in the steepest part of the standard curve—that is, the most sensitive part of the range covered. In some cases, it may be necessary to include further levels of control—for example, it may be considered important to include a control for the higher discriminator in quantitative assays.

Immunofluorescence

These are in-house assays that also use positive and negative sera to check the other materials used (tissue, conjugate, etc) and to provide reference patterns in these subjective assays. IQC samples at a dilution providing a weak positive result

Table 2 Suggested acceptable levels of precision

Method	Acceptable precision (between-batch coefficient of variation, %)
Turbidimetry	3–5
Nephelometry	3–5
ELISA	8–12
Radioimmunoassay	8–12
Radial immunodiffusion	10–20
Rocket electrophoresis	10–20

ELISA, enzyme-linked immunosorbent assay.

must also be used to check the sensitivity on a daily basis. In practice, serial dilutions of a positive control are used, which should be positive to the same titre in each batch.

Flow cytometry

IQC for cell surface marker analysis by using flow cytometry had proved to be problematic in the past, owing largely to the need to use fresh blood for the analyses. Recent developments that use stabilised whole-blood preparations are now available, however, and are stable for many of the major surface molecules (eg, CD3, CD4 and CD8). For immunochemical analyses, similar statistical analyses to those described below are applicable.

Shewhart and Cusum charts

The most widely used tool for IQC is the Shewhart chart³ (also called Levey-Jennings chart), an example of which is given in fig 1. This is prepared by first assaying the analyte on at least 20 separate occasions. From the data generated, the mean (SD) is calculated. In the case shown, this was 1.27 (0.035). The mean (2 SD) levels are plotted as shown. Further results are then plotted on the graph on a daily basis. Figure 1 shows 32 such entries and, if normally distributed, 95% of these should fall between these limits. In this example, the precision of the assay is given by coefficient of variation = 3.4%.

Westgard and colleagues⁴ proposed a series of rules to allow interpretation of the data shown in the Shewhart chart and these, or a modified version, are in use in many laboratories. The original proposals for some elements suggested looking for trends over nine batches. This may be inappropriate for less frequent analyses (eg, those carried out on a weekly basis), as this would equate to 2 months before action is considered. In such circumstances, using five consecutive points is more practical. The choice of five points is a compromise. As there is a 50% chance of a result lying above the mean each time the assay is carried out, there is a 2⁵ (3%) chance that five in a row will all lie above the mean, purely as a random event. This is, however, sufficiently unusual to act as a prompt. These rules give warnings of relatively rare events and are not absolute measures of assay failure.

Figure 1 shows several indicators that would alert the operator to potential problems with the assay. Each of these events occurs purely at random, <5% of the time. These are

1. two consecutive points outside 2 SD;
2. a single point outside 3 SD;
3. five consecutive points either rising or falling; and
4. five consecutive points either above or below the mean.

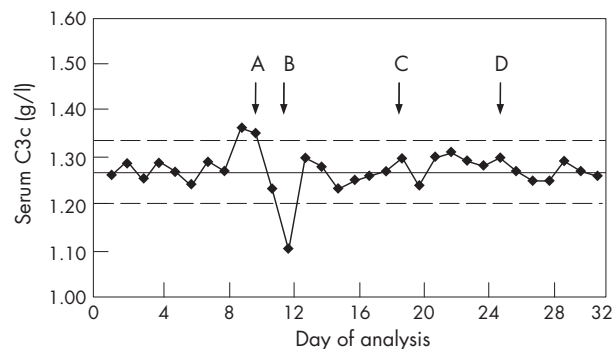


Figure 1 Shewhart chart for serum complement C3c. "Action points" are shown by arrows A-D. These indicate where the assay is potentially out of control (see text) and where remedial action should be considered.

Figure 1 shows each of these by means of arrows indicating the action points.

A better way to monitor changes in accuracy is to use the cumulative summation, or Cusum, chart (fig 2). It has not been popular in the past owing to the complexity of the data manipulation required, but many software packages now include this approach. In this example, the same data are used but the cumulative difference from the mean is plotted on the y axis. Table 3 shows a worked example with some of the data used to generate figs 1 and 2. A change in the slope indicates a change in accuracy and examples of this are indicated in fig 2 by arrows. Ideally, the graph should oscillate along the x axis, as in region "a". A rising line indicates a negative bias, and a falling line, a positive bias. The method is, perhaps, oversensitive to small inaccuracies in the set mean, which limits its utility.

It is, of course, insufficient merely to identify an action point. Use of quality-control charts must be incorporated into the standard operating procedures of the laboratories. The exact steps to be taken vary according to the method in question (eg, nephelometry v ELISA) and to the type of action point (eg, single point outside 3 SD v a trend of five rising points.) A one-off erroneous point suggests a sampling error, by either the analyser or the operator. Where a trend is seen—that is, a gradual rise or fall on the Shewhart chart—a change in accuracy, usually caused by deterioration of reagents or reference material, is indicated. In ELISA, for example, substrate, marker antibody, plate coating and calibrator must be considered. Simple checks, such as those for pH of buffer or formation of precipitates, should be performed. In practice, it is often easiest to discard all the reagents and start afresh.

Internal consistency

A panel of data to be assessed should be internally consistent—that is, things should add up. This concept is best illustrated with examples.

In immunophenotyping, the internal controls inherent to the samples being analysed allow us to check the quality of the antibodies used for leukaemia and lymphoma typing. Most samples have a population of cells that is residual normal, or malignant cells that are positive for a given marker. These can be used as a control for the reaction of the antibody—for example, in B cell leukaemia, the residual normal T cells can act as controls for the T cell markers used. The strength of a reaction (ie, the mean or median peak

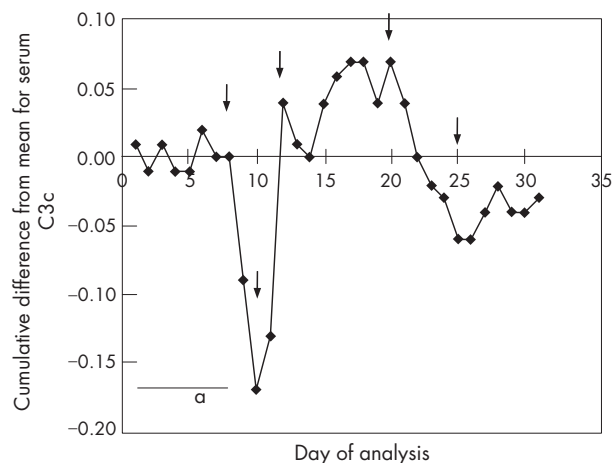


Figure 2 Cusum chart with the same serum complement C3c data as in fig 1. Area "a" shows an accurate assay in good control. Arrows indicate points where a change in accuracy has occurred.

Table 3 Worked example preparing data for plotting on a Cusum chart

Day	C3c (g/l)	Target–result	Cumulative sum
1	1.26	0.01	0.01
2	1.29	–0.02	–0.01
3	1.25	0.02	0.01
4	1.29	–0.02	–0.01
5	1.27	0.00	–0.01
6	1.24	0.03	0.02
7	1.29	–0.02	0.00
8	1.27	0.00	0.00
9	1.36	–0.09	–0.09
10	1.35	–0.08	–0.17
11	1.23	0.04	–0.13
12	1.10	0.17	0.04

The target from previous data is 1.27 g/l. Thus, on day 2, the measured result is 1.29. Target–result = 1.27–1.29 = –0.02.

Cumulative sum = previous cumulative sum + target–result = 0.01 + (–0.02) = –0.01.

height in normal samples), which is predictable in a given system, should be visually inspected to ensure adequate function of that marker. Antibodies should be inspected weekly for turbidity and, if found, turbid reagents should be discarded.

Other examples of the internal consistency of data come from an understanding of the biology. So, IgG1+IgG2+IgG3+IgG4 should equal the total IgG (SE 10% to allow for some error in the individual measurements). Similarly, in HIV monitoring, CD3-positive cells should equal CD4-positive and CD8-positive cells (SE 5%) and T cells+B cells+natural killer cells should equal 100% (SE 5%).

Another important check comes from an intimate familiarity with the method in question. For example, in our practice we use anti-tissue transglutaminase to screen for people with coeliac disease. Our assay generally finds 6–10% of all samples analysed to be positive. When a batch does not conform to the expected profile, great care should be taken to understand why (for instance, it may simply be that an extra gastroenterology clinic has taken place). This technique is similar in concept to the “patient daily means” technique, in which data for a batch are truncated to remove very low and very high results and the mean is calculated for the remainder. Changes in mean are considered to reflect bias in the analysis. As described above, the patient mix on a given day may cause large swings in the mean—for example, a myeloma follow-up clinic would potentially skew the mean for IgG, IgA and IgM analyses.

The final check of internal consistency is a process called “delta checking”, in which sequential data on patients are examined for changes (differential, hence Δ). Examples include paraprotein levels (in monoclonal gammopathies) and anti-double stranded DNA antibodies in systemic lupus erythematosus. Many of these remain stable for long periods, and unexpected fluctuations, especially when viewed as a batch as opposed to any one individual, may indicate calibration problems. It should be noted that to be able to apply this effectively, the operator must have a clear idea of the assay characteristics, including assay variability

Take-home messages

- The major objective of quality assurance is to improve the quality of results such that uniformity exists both within and between laboratories.
- Monitoring analytical sensitivity is essential to prevent the reporting of false-negative results.
- For quantitative data, Shewhart/Levey–Jennings charts are useful in monitoring precision.
- “Third-party” controls should be included wherever possible.
- The main benefits of internal quality control are realised with early recognition of problems and swift introduction of corrective action.

(coefficient of variation) and normal physiological variability of the analyte.

All the above criteria have led to a position whereby a decision can be made of whether the results provided are acceptable. If the IQC fails to meet the above criteria, the operator must consider what action to take. Some of the alternatives may include the following:

1. Accept the batch regardless, watch and wait. Possibly, this is a quirk of the statistics.
2. Rerun a new vial or batch of quality-control material. Some materials are less stable than others—for example, reconstituted freeze-dried controls.
3. Recalibrate and rerun.
4. Carry out preventive or corrective maintenance on the analyser and rerun.
5. Rebuild the assay from first principles, taking into account the variables described in the first section.

This list is by no means exhaustive, but the underlying principle is that if the IQC fails it cannot simply be ignored. Some response, and a record thereof, is essential.

IQC is clearly only one element in the laboratory’s quest for quality and should not be considered in isolation. Striving for internal consistency in terms of precision and accuracy, however, lays the foundation for the remainder. The main benefits of IQC are realised with early recognition of problems and swift introduction of corrective action.

REFERENCES

- 1 **World Health Organization.** *Guidelines for the preparation and establishment of reference materials for biological substances*, WHO/BS/626/78. Geneva: World Health Organization, 1978.
- 2 **Wilson WA, Gharavi AE, Koike T, et al.** International consensus statement on preliminary classification criteria for definite antiphospholipid syndrome. *Arthritis Rheum* 1999;**42**:1309–11.
- 3 **Shewhart WA.** *Economic control of quality of manufactured products*. New York: Van Nostrand, 1931.
- 4 **Westgard JO, Barry PL, Hunt MR.** Proposed selected method: a multi-rule Shewhart control chart for quality control in clinical chemistry. *Clin Chem* 1981;**27**:493–501.