Upsala J Med Sci 98: 221–240, 1993

4. Guidelines for Assessing Analytical Quality Requirements

Per Hyltoft Petersen,[1] Torgny Groth[2] and Carl-Henric de Verdier,[3]

[1]*Department of Clinical Chemistry, University Hospital, Odense Denmark,* [2]*Unit for Biomedical Systems Analysis, and* [3]*Department of Clinical Chemistry, University of Uppsala, Sweden*

In ref. (30) (Chapter 2) the main principles of analytical goal-setting and formulation of analytical quality specifications (AQSpecs) were outlined. From this document it is possible to learn how to assess quality specifications according to 'the state of the art approach' and 'the biological approach', as well as 'the clinical usefulness approach'. Assessment of specifications based on clinical strategies/situations, however, need more preparation, starting with a detailed analysis of the problem. There are no simple formulas, but there are several applicable models for the purpose, which will be described in this chapter.

4.1. Defining the Clinical Situation/Strategy/Problem

Before starting the process of assessing the analytical quality requirements, the overall situation for which the AQSpecs are to be formulated should be considered, without taking models and methods into account:

a. Is the purpose of the AQSpecs general in nature? *e.g.* without specifications of the clinical use of the laboratory investigation(s) in question?
b. Are the AQSpecs specific for a defined clinical situation, *e.g.* screening, diagnosis, other classification, intervention, regulation, monitoring, or more complicated situations?
c. Are the AQSpecs specific for a scientific investigation? *e.g.* estimation of biological/pathological response, clinical trial, or other defined investigation?

In the first case (a), the general nature of the purpose (or lack of concrete intentions) make the models and the solutions to the problem as general as the purpose for the

AQSpecs. The most relevant approach will be to apply a general strategy of 'not increasing the measured biological within-subject variation significantly' or of 'sharing common reference intervals', where analytical variation is generally included, *cp.* ref. (30).

In the second case (b) of a defined clinical situation, the following steps should be considered:

b. 1 Specify the clinical situation and define the outcome of the clinical decision making process in quantitative terms, *e.g.* number or fraction of misclassifications/misinterpretations, economical consequences *etc.* This could be the outcome of the total strategy/situation or of a well defined part of the process.

b. 2 Define the laboratory investigations, (generic quantities) (5), included in the clinical strategy. Are they determinative of the clinical outcome? If not or if the significance is doubtful, the assessment process may be difficult to perform or the AQSpecs may be uncertain. Therefore, the strategy should be reconsidered, aiming at a more clear definition of the impact of the results of laboratory investigations.

In the third case (c) of a scientific investigation, the problem must be described in detail and the outcome must be clearly defined. It may be easier to separate the problem into a number of subproblems, to be investigated one by one. As the scientific problems may be very different, a general formula cannot be given. The steps (b.1 - b.2 above) for a defined clinical situation may, however, serve as a guideline for a relevant assessment.

4.2. Selecting Methods and Tools for Statistical, Graphical, and Computer Analysis

The methods are often designated as 'statistical', 'graphical', or as 'computer' methods. Many of the methods are based on general statistics, and for these the three approaches will lead to similar specifications, when the same assumptions are made.
Data for statistical assessment of 'clinical goals' and 'AQSpecs' are generally collected according to well defined situations:

i) single point measuring (using a value at one time point originating from one or several samplings and one or several measurements) for comparison with one single reference distribution *(unimodal classification)*,

ii) single point measuring for comparison with two reference distributions *(bimodal classification)* ,

iii) two point monitoring (change compared with s_{Bw}),

iv) several point monitoring (time series),

v) other situations.

The word monitor is here used in the clinical sense *i.e.*: "to keep close and constant watch of a condition or function".

A 'statistical method' based on the manual use of tabulated statistical distributions (usually gaussian or log-gaussian) does not allow a combination of different types of distributions. The 'graphical method' may be used to illustrate any gaussian or non-gaussian distribution, but has got its limitations in combining gaussian analytical variation with non-gaussian biological distributions. The 'computer methods' have no limitations in mixing various statistics, graphics, and computational methods.

The choice between the three types of methods presented below, all involving strong elements of statistics, may be a question of available equipment or of individual taste and educational background, but it is also a matter of the complexity of the clinical situation studied.

Statistical (table-look-up) method

A Gaussian distribution describing a *measured* biological reference distribution has the following parameters: mean value = M_T and standard deviation = s_T (coefficient of variation = CV_T). *For explanation of abbreviations see ref. (7) = Chapter 11*

Any decision limit, DL (cut off point), can be expressed in terms of the parameter $z = (DL - M_T)/s_T$, and the fractions of the distribution above and below the DL can be read off from a statistical table for given z values.

The effect of analytical bias, B_A, is simply calculated from

$z\ (bias) = (DL - M_T + B_A)/s_T$.

Investigation of the effect of analytical imprecision (s_A) is a two step process. First the biological standard deviation, $s_B = (s_T^2 - s_A^2)^{½}$, and $z\ (ideal) = (DL - M_T)/s_B$ are calculated. Then, a different value of imprecision, s_A, is assumed and the total standard deviation, s_T, is calculated for this value of imprecision:

$s_T = [s_B^2 + s_A^2]^{½}$.

The combined effect of various values of bias and imprecision can be read off from statistical tables using z (combined) = $(DL - M_T + B_A)/s_T$. Statistical tables may be found in most books on statistics; see *e.g.* (3, 45).

Graphical (tail-area) method

One reference distribution ('unimodal classification').

A gaussian reference distribution can be delineated as a bell-shaped curve in a linear plot or, as a straight line in a probit-plot (Fig. 4.2.1.). The bell-shaped presentation may be easier to grasp, but more difficult to evaluate, whereas the probit-plot is, after some practice, relatively easy to evaluate [for the theory *cf. e.g.* Bliis (2) or Gowans *et al.* (18)].

A positive bias will move the distribution upwards resulting in a lower fraction below DL and a higher above DL (Fig. 4.2.2). In fact, a bias will have the same effect as moving the DL - but with the opposite sign.

Evaluation of the effect of imprecision needs a few calculations like those for the statistical table-look-up method. The ideal biological distribution will show up as a more narrow bell-shaped curve and in the probit-plot the slope of the line will become steeper (Fig. 4.2.3).

Assumed values of analytical imprecision, s_A, can be combined with the ideal biological standard deviation, s_B, to estimate the corresponding total distributions as illustrated in Fig. 4.2.4.

Non-gaussian reference distributions can sometimes be handled after transformation to gaussian, *e.g.* a log-gaussian distribution, where the logarithmic values are distributed as gaussian. Even non-parametric distributions that cannot be transformed in this simple way can be evaluated with regard to bias (Fig. 4.2.5), but not with regard to imprecision.

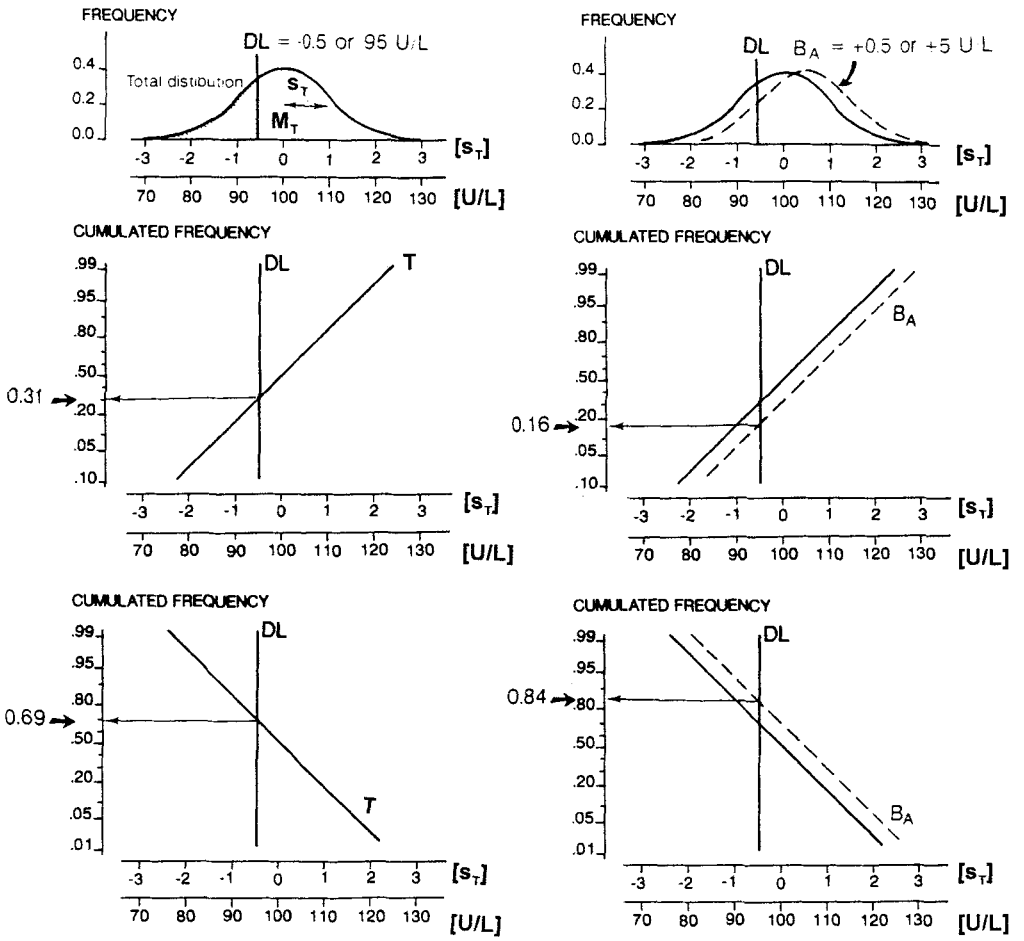


Fig. 4.2.1. Graphical presentations of a total *(measured)* biological reference distribution (T) with mean = M_T and standard deviation = s_T, in the example 100 U/L and 10 U/L, respec-tively. The decision limit (DL), at 95 U/L (= - 0.5 s_T), is also illustrated.

Top: The relative frequency as a function of the standard deviation which relates to the concentration. DL separates the distribution in values below and above DL and the two areas represent the corresponding fractions of the total distribution.

Middle: Probit transformation of the same distribution cumulated from low values. The ordinate indicates the area below any DL (in the example DL is - 0.5 s_T and the area or fraction is ≈ 0.31.

Bottom: Probit transformation of the same distribution cumulated from high values. The ordinate indicates the area above any DL (here SD - 0.5 s_T equal to 95 U/L and corresponding to ≈ 0.69).

Fig. 4.2.2. Same reference distributions as in Fig. 4.2.1., but with a positive bias (dashed lines). In the example B_A = + 0.5 (+ 5 U/L) resulting in a smaller fraction below DL (which now corresponds to - 1.0 s_T or ≈ 0,16) and a larger fraction above DL ≈ 0.84 .

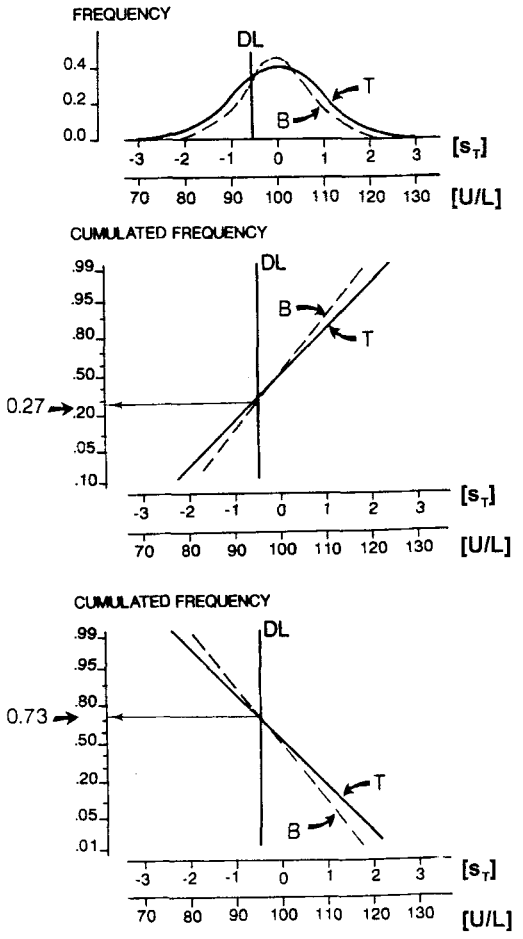


Fig. 4.2.3. Same reference distributions as in Fig 4.2.1. together with the pure biological distribution (dashed line). In the example the analytical s_A during measurements is estimated to 6 U/L, so $s_B = (10^2 - 6^2)^{½} = 8$ U/L (-0.625) and the fractions below and above DL become ≈ 0.27 and ≈ 0.73, respectively.

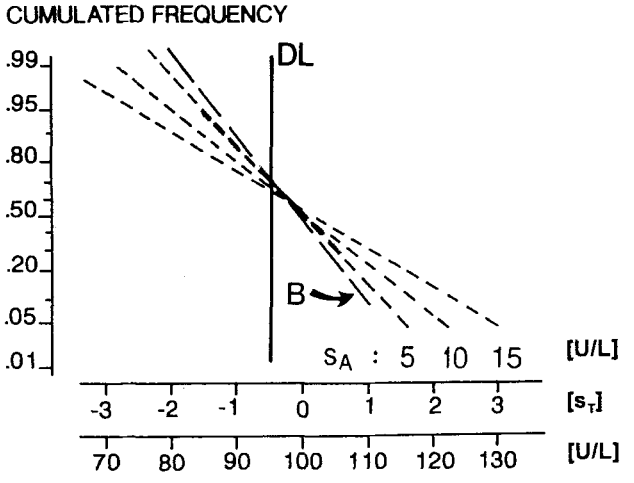


Fig. 4.2.4. The pure biological reference distribution (long dashed line) together with assumed values of imprecision s_A = 5, 10, and 15 U/L (short dashed lines). $s_T = (8^2 + s_A^2)^{½}$. Illustrated for the probit transformation with cumulation from the high values, only.

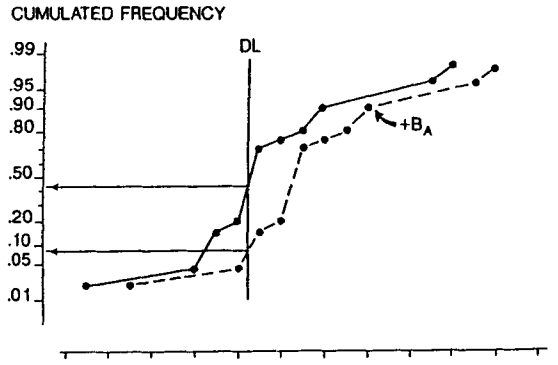


Fig. 4.2.5. Illustration of a non-parametric biological distribution and the effect of bias.

Probit paper is available from H. W. Peel and Co. Ltd., Jeymer Drive, Greenford, Middlesex, UBG 8 NX, England and A.G. Frisinette og Sønner Aps, Egsmark, 8400 Æbeltoft, Denmark (Sandsynlighedspapir no. 2110 (linear abscissa) and 2109 (log abscissa).

Two reference distributions (Bimodal classification).

This procedure is used for single point measuring when two biological reference distributions are assumed, representing a diseased (or pre-diseased) and a healthy reference sample group. The prevalence for disease is an important parameter which may be more or less well-known , but can be varied within reasonable limits in the assessment. In the following example for demonstration (Fig. 4.2.6-8) the prevalence for sickness and health is equal (= 0.5). The rest of the values of the parameters of the two distributions are as follows: *for the healthy reference sample group* $M_B \approx M_T$ = 103 nmol/L, s_B = 19 nmol/L and *for the diseased reference sample group* $M_B \approx M_T$ = 220 nmol/L, s_B = 36 nmol/L, with a decision limit of 140 nmol/L. In Fig. 4.2.6 the two reference distributions are shown in a probit-plot together with the decision limit. Assumed values of analytical bias (+5, +10, +15, +20 nmol/L) are illustrated in Fig. 4.2.7. and assumed values of imprecision (5, 10, 15, 20 nmol/L) in Fig. 4.2.8. The same procedure may be performed for a given bias and variable imprecision or *vice versa* in order to investigate combined effects.

The basis for evaluation is the tail-areas representing the number of false positives, FP, and false negatives, FN. One way is to calculate the number of FN and FP, but it may be easier to compare results expressed as fractions - the sum of misclassifications (FP+FN) related to total number of classifications - and to plot these fractions *vs* bias or imprecision.

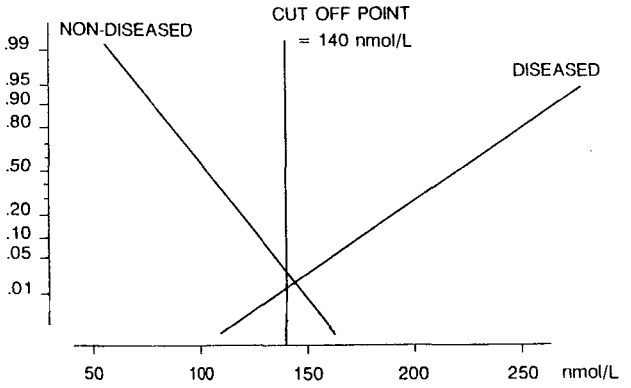


Fig. 4.2.6. Bimodal classification model with the non-diseased reference distribution (M_B = 103 nmol/L, s_B = 19 nmol/L) cumulated from high values and the diseased reference distribution (M_B = 220 nmol/L, s_B = 36 nmol/L) cumulated from low values. The decision limit (cut off point) (140 nmol/L) is illustrated as a vertical line.

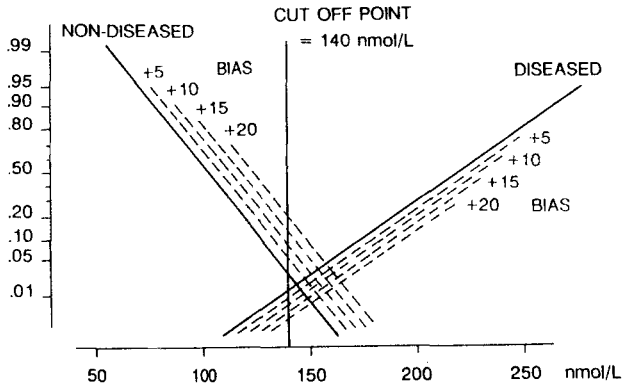


Fig. 4.2.7. Same as in Fig. 4.2.6 together with reference distributions for assumed bias-values of + 5, + 10, + 15, and + 20 nmol/L.

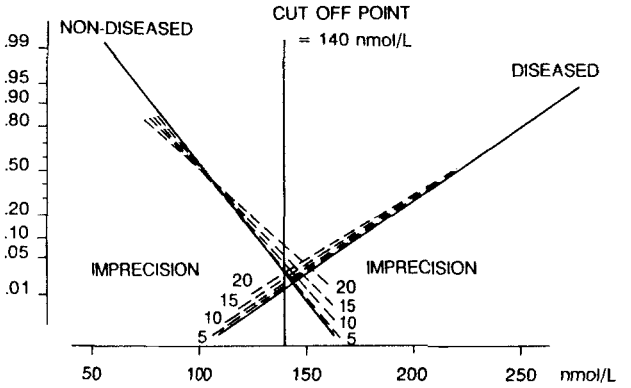


Fig. 4.2.8. Same as in Fig. 4.2.6 together with reference distributions for assumed imprecision of 5, 10, 15 and 20 nmol/L

Computer based cost analysis method

Computer programs may be used for calculation of

- the expected frequency of FN and FP outcome in connection with "bimodal classification" using a specified decision limit;
- the cost related to misclassification;
- the optimal discriminatory level in the sense that the costs are minimized;
- the diagnostic sensitivity, specificity and the prediction value of a positive and a negative test

The input data to such a program *cf.* (23) are

- mathematical frequency distributions, *e.g.* gaussian or log-gaussian, representing healthy and pathological reference sample groups;
- numerical weights, w_1 and w_2, representing the relative costs for making misclassifications;
- estimates of pre-analytical and analytical imprecision, expressed as coefficients of variation, CV_{preA} and CV_A;
- estimate of analytical bias, B_A;

The variances of the simulated reference distributions are calculated from

$$s_T^2 = s_B^2 + s_A^2 + s_{preA}^2$$

where s_B = total biological standard deviation

s_A = total analytical standard deviation

s_{preA} = pre-analytical standard deviation

The program calculates the number of FP and FN as the "tail-areas" cut off by the decision limit from the various distributions. The relative loss calculated as the weighted sum of false positives and false negatives

$$L = w_1 \cdot FP + w_2 \cdot FN$$

The optimal decision limit is determined as the value minimizing the relative loss. Diagnostic sensitivity and specificity, predictive values of positive and negative tests are calculated from conventional formulas.

Monte Carlo simulation techniques

Another computer method of great interest in this connection is the Monte Carlo simulation technique, based on random number generators and appropriate frequency distributions, and which can be used to generate synthetic data.

The method has the advantage that the outcome of classification can be studied on a

"case-by-case" basis, considering *e.g.* repeated measurements and reclassification of borderline cases (39). Furthermore, the influence of the size of the reference sample groups on the estimated decision limit and the outcome of the classification process can be studied.

This method is described in more detail in ref. (20; p.57) as applied to screening of hypocalcaemia; and in ref. (39) as applied to "bimodal screening" of pancreatic insufficiency with use of serum pancreatic iso-amylase measurements.

Biodynamic modelling technique

Monte Carlo techniques are also very useful in combination with computer simulation models describing the pathophysiological and/or biochemical dynamics of selected quantities of special interest in diagnosis or monitoring of patients. A good introduction to this approach is found in ref. (53) as applied to the design of an allopurinol loading test for characterization of liver function. In this paper the analytical requirements were assessed in connection with an optimized design of the blood sampling schedule in terms of number and location of measurements required for identification of diagnostic parameters derived from a combined biochemical and physiological compartmental model. Illustrative examples of the same technique are found in refs. (21, 22, 31, 41).

4.3. Assessment Procedure

Prerequisites

Specification of:

1. The clinical strategy and the perceived consequences of the results, expressed in quantitative terms as a measurable outcome.
2. The distribution(s) of reference values, and decision limits (DL), estimates of biological within- and between-subject variation, preanalytical factors, and the characteristics of the analytical measurement and quality control procedures.

Procedure:

1. Describe the ideal error-free situation and calculate the clinical outcome.
2. Assume various analytical conditions: systematic error (bias) and random error (imprecision); one at the time and in combinations.
3. Calculate the outcome for the various analytical conditions.

4. Decide on the acceptable decrease in outcome and determine the maximal allowable combinations of systematic and random error. Preanalytical variation ought to be included in the calculations.

Illustrative examples:

In Fig. 4.3.1 references to published models for assessing clinical goals and AQSpecs, and some applications of these models, are listed according to name of first author and year of publication. These references may give suggestions and guidance for your assessment process. The papers are grouped according to the scheme in the beginning of Chapter 4.2. To the left in Fig. 4.3.1 it is indicated whether the problem is investigated according to *'the biological approach'* or *'the clinical usefulness approach'*, and to the right whether it deals with imprecision (s_A or CV_A) or bias (B_A), difference in bias between two laboratories (B_{bL}) or at two different occasions in the same laboratory (B_{wL}). For definition of imprecision, bias and systematic error see ref. (7) (Chapter 11). From left to right papers are listed dealing with: single point measuring, two point monitoring, several point monitoring, and other situations.

When papers deal with several clinical situations and models the same reference is listed in all the appropriate areas. To provide some guidance in the selection of relevant literature for your own assessment some main clinical approaches/models will be commented on [*cf.* ref. (30) (Chapter 2) for the biological models].

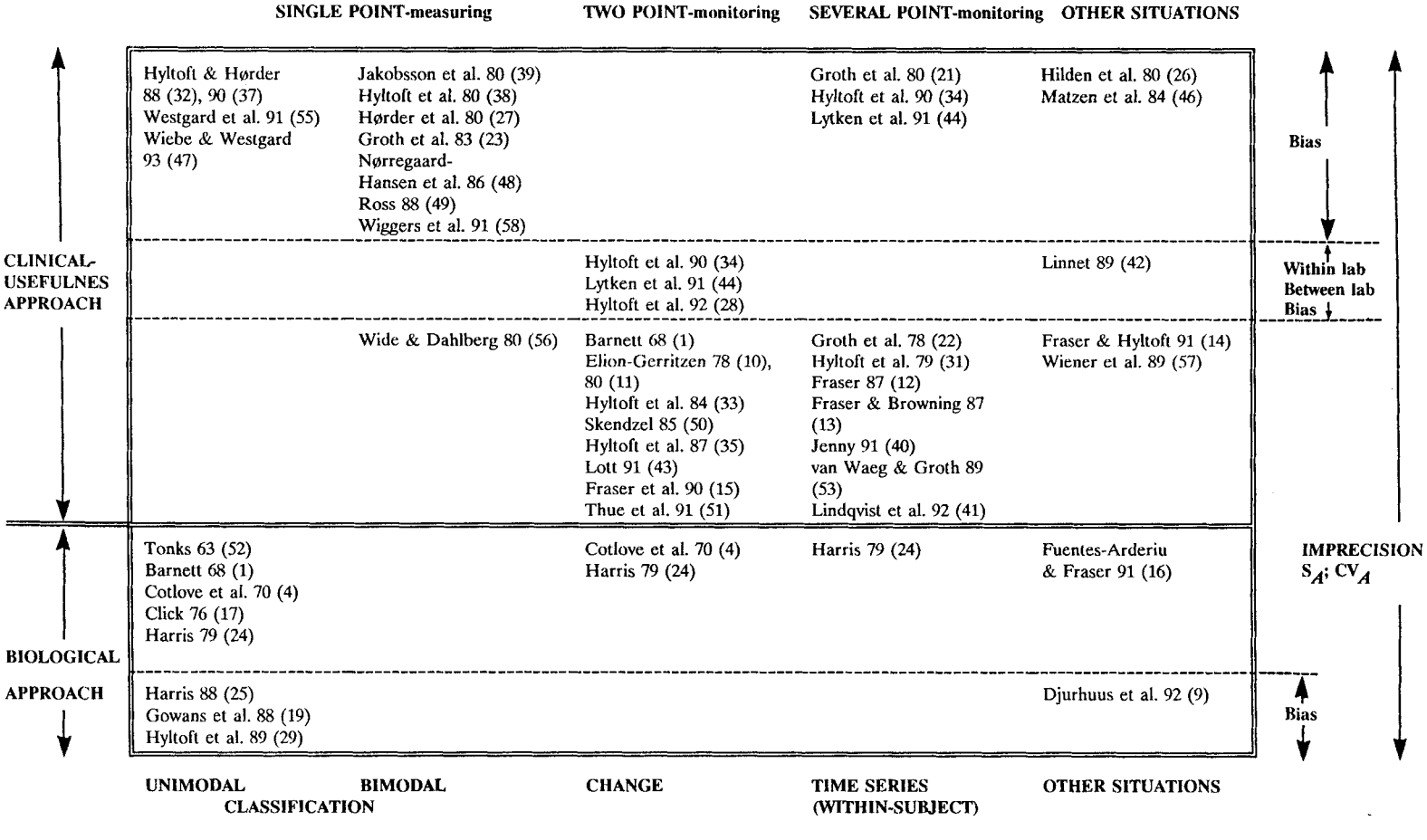
1. One reference distribution ('Unimodal classification')

This is used for single point measuring where only one biological reference distribution is assumed and the clinical interpretation is gradual, *e.g.* in S--Cholesterol screening where high values indicate increased risk of heart diseases (47) (in contrast to the use of cholesterol in diagnosis of familiar hypercholesterolaemia, which is a clearly bimodal classification problem).

Using the graphical example in Figs. 4.2.1 - 4 and 4.2.4. alone or in combination, it is always possible to study the effect of imprecision and bias on the clinical decisions. The specific situation depends on the problem under investigation and the reader is referred to four *ref.* (19, 36, 54, 55).

Fig. 4.3.1.

EVALUATION OF MODELS AND APPLICATIONS OF THESE FOR ESTIMATION OF QUALITY SPECIFICATIONS



2. *Two reference distributions ('Bimodal classification')*

From the demonstration example of a bimodal classification (Fig. 4.2.6) it can be seen that with B_A and s_A equal to zero the fraction of misclassified "diseased" is about 0.01 and the fraction of misclassified 'non-diseased' about 0.02 of the total number of classified; thus for a prevalence of 0.50 the total fraction of misclassified individuals is equal to 0.04.

From Fig. 4.2.7. and Fig. 4.2.8. it is possible to graphically estimate the effect of bias and imprecision, respectively. Furthermore, it is possible to construct graphs that combine the effects of bias and imprecision. It will be the tolerable fraction of misclassifications that determines the analytical quality specifications.

The following comments or conclusions can be made from the references:

a. Functions of misclassifications have been presented as combinations of imprecision and bias (23, 27, 39) and imprecision and decision limits where also nomograms describing the type are given (48, 58).
b. In the paper of Jacobsson *et al.* (39) a strategy of repeated measurements, has been investigated using Monte Carlo computer simulation technique. It was found that the influence of pre-analytical variation, duplicate measurements, reclassification of borderline cases and sizes of reference sample groups had some importance, but that the cost weighting ratio of false negatives: false positives was critical for the results.
c. Ross (49) used a theoretical model based on strict assumptions about the reference distributions, and the outcome was presented as efficiency.
d. In the work of Wide and Dahlberg (56) the analytical precision was optimized using a more sensitive function test for evaluation.

3. *Two point monitoring*

A unimodal model of repeated measuring has been described by Harris, from a statistical point of view, including the biological within-subject variation (24). An overlapping paper by Hyltoft Petersen *et al.* (33) describes the conversion of a two-point monitoring into a bimodal classification problem by calculating the ratios between measurement results from the same individual.

A series of investigations on 'the clinical usefulness approach' (10, 11, 15, 43, 50, 51) have been described and discussed previously (30) (Chapter 2). These papers deal with how physicians and practitioners on the average react on a change in test result. In

another series of papers a model is studied for evaluation of quality specifications in situations, where there is an agreement that a certain change of the analytical result ought to lead to clinical actions (15, 28, 34, 35, 44).

The predetermined change, Δ, is often empirical in nature, and the evaluation of quality specifications are based on the assumption that under stable (steady-state) conditions the probability of measuring a change $\geq \Delta$ should be less than a certain probability, P.

The general formula is

$$\Delta \geq z_P \sqrt{2} \, (s_A^2 + s_{Bw}^2)^{1/2} + B_{wL}$$

where the z_P factor is set to a value corresponding to a specified probability (P), s_A and s_{Bw} are the analytical and biological within-subject variation, respectively, and B_{wL} is a possible difference in bias, either within the same instrument at the two measurement occasions or when two instruments with different bias are used interchangeable. (If the two measurements were made in different laboratories B_{bL} is used.)

Based on this formula the interrelationships between s_A and B_{wL} can be expressed, when s_{Bw} for the quantity is known as well as the predetermined change (Δ) or the probability for exceeding Δ under steady state conditions. Rearrangement of the formula gives $B_{wL} \leq \Delta - z_p \cdot \sqrt{2} \, (s_A^2 + s_{Bw}^2)^{1/2}$, from which B_{wL} can be calculated for various assumed values of s_A. The shapes of curves for combined values of B_{wL} and s_A depend on Δ and z_P as illustrated in Fig. 4.3.2 for HbA_{1c} used for control of diabetic patients. HbA_{1c} is measured as an amount-of-substance fraction and expressed in per cent. The diagram shows that the maximal CV_A is nearly the same when you use Δ = 1% which gives a probability for a change of 0.80 and Δ = 2% which gives a probability for a change of 0.99.

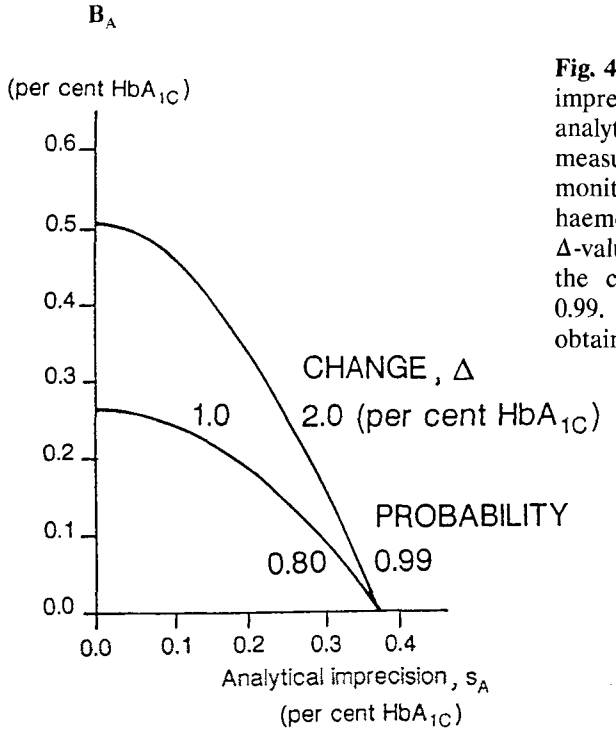


Fig. 4.3.2. Combined values of maximal imprecision (CV_A) and change in analytical bias (B_A) between the measurements of two samples during monitoring of a patient for gluconated haemoglobin. The two curves illustrate Δ-values of 1.0 and 2.0% HbA_{1c} with the corresponding P-values 0.80 and 0.99. The data used in this figure are obtained from (44).

If bias (B_{wL} or B_{bL}) is assumed to be zero another rearrangement of the formula can be made, which gives:

$s_A \leq [(\Delta^2/2z_P^2) - s_{Bw}^2]^{½}$. This expression has been used to study if multiple sampling can give a more accurate estimate of Δ, which can help to reduce the requirements on s_A. This approach is only possible for components with comparatively high s_{Bw}-values (*e.g.* S--Creatinine and S--Cholesterol) (6).

4. Several point monitoring

One example of the drug-monitoring is described in refs (12, 13). Three other papers (21, 22, 31) deal with complicated turn-over models investigated by computer simulation techniques, which would need more space than available to describe in detail.

One paper (44) deals with the same situation of keeping concentrations of a selected quantity below (or above) a certain value in the treatment of a patient. Even if the problem is different from the unimodal classification, the theoretical handling is similar. It is, however, easier here to introduce the extra information or assumption, by looking at patients with decision limits above (or below) the value, assuming that they should be interpreted analogously.

5. Other situations

These are very different in nature covering special aspects of quality specifications.

- Two papers (26, 46) deal with computer - aided diagnosis of jaundice, where the clinical chemical quantities are not determinative. This leads to the (expected) conclusion that analytical quality here is less important.
- Two papers (14, 42) deal with preanalytical variation related to sampling technique.
- In one paper computer supported decision analysis is used (57).
- One theoretical paper (16) deals with quality specifications for interference.
- Finally a paper (9) deals with the actual quality of reference intervals related to the current analytical quality.

REFERENCES

1. Barnett RN. Medical significance of laboratory results. Am J Clin Path 1968;50:671-676.
2. Bliss G. Statistica in Bioloy, New York: Mc Graw - Hill Book Co., 1967.
3. Colton T. Statistics in Medicine, Boston, Mass 02106, USA: Little, Brown and Co.
4. Cotlove E, Harris EK, Williams GZ. Biological and analytical components of variation in long-term studies of serum constituents in normal subjects. III. Physiological and medicial implications. Clin Chem 1970;16:1028-1032.
5. CEN/TC251/PT 002 L. Terminology and coding systems of medical procedures. b. Quantities in Laboratory Medicine. First working document approved at 14th CEN/TC251 meeting 1993-12-07.
6. de Verdier C-H, Groth T, Hyltoft Petersen P. Medical need for quality specifications - a NORDKEM project for selecting the appropriate quality in clinical laboratories. Scand J Clin Lab Invest 1993;53, Suppl 215:29-37.
7. de Verdier C-H, Groth T, Hyltoft Petersen P. Explanation of some other concepts and terms used in this report. Upsala J Med Sciences 1993;98(3):487-491. In Press.
8. de Verdier C-H. Approaches towards a combined data- and knowledge-base for analytical quality specifications in clinical chemical laboratories. Upsala J Med Sciences 1993. Chapter 7. In Press.
9. Djurhuus MS, Roholt A, Vadstrup S, Hyltoft Petersen P, Uldall A. Reference intervals based on hospitalized 'healthy' patients and medical students in relation to analytical bias for serum potassium. Scand J Clin Lab Invest 1992;52:305-312.

10. Elion-Gerritzen W.E. Requirements for analytical performance in clinical chemistry. An evaluation from the poit of view of practising physian, Thesis. Erasmus University. Rotterdam, 1978.
11. Elion-Gerritzen WE. Analytical precision in clinical chemistry and medical decisions. Am J Clin Path 1980;73:183-195.
12. Fraser CG. Desirable standards of performance for therapeutic drug monitoring. Clin Chem 1987;33:387-389.
13. Fraser CG, Browning MCK. Deciding the optimum interval between specimen collections: Theory and nomograms. Clin Chem 1987;33:1436-1438.
14. Fraser CG, Hyltoft Petersen P. The importance of imprecision. Ann Clin Biochem 1991;28:207-211.
15. Fraser CG, Hyltoft Petersen P, Lytken Larsen M. Setting Analytical Goals for Random Analytical Error in Specific Clinical Monitoring Situations. Clin Chem 1990; 36:1625-1628.
16. Fuentes-Arderiu X, Fraser CG. Analytical goals for interference. Ann Clin Biochem 1991;28:393-395.
17. Glick JH. Expression of random analytical error as a percentage of the range of clinical interest. Clin Chem 1976;22:475-483.
18. Gowans EMS, Fraser CG, Hyltoft Petersen P. A guide to the ose of probit transformation of Gaussian distributions. Biochemica Clinica 1989;13:327-336.
19. Gowans EMS, Hyltoft Petersen P, Blaabjerg O, Hoerder M. Analytical goals for the acceptance of common reference intervals for laboratories throughout a geographical area. Scand J Clin Lab Invest 1988;48:757-764.
20. Groth T. Methods and approaches to determine "optimal" quality for clinical chemical data. Scand J Clin Lab Invest 1980;40 Suppl 155:47-63.
21. Groth T, Hakman M, Hällgren R, Roxin L-E, Venge P. Diagnosis, size estimation and prediction of acute myocardial infarction from S-myoglobin observations: A system analysis to assess the influence of various sources of variability. Scand J Clin Lab Invest 1980;40 suppl 155:111-124.
22. Groth T, Hyltoft Petersen P, Hjelm P. Characterization of increased synthesis of acute phase proteins from plasma concentration measurements. Computer Programs in Biomedicine 1978;8:295-310.
23. Groth T, Ljunghall S, de Verdier C-H. Optimal screening for patients with hyperparathyroidism with use of serum calcium observations: A decision-theoretical analysis. Scand J Clin Lab Invest 1983;43:699-707.

24. Harris EK. Statistical principles underlying analytical goal-setting in clinical chemistry. Am J Clin Pathol 1979;72:374-382.
25. Harris EK. Proposed goals for analytical precision and accuracy in single-point testing: theoretical basis and comparison with data from the College of American pathologists proficiency surveys. Arch Pathol Lab Med 1988;112:416-420.
26. Hilden J, Matzen O, Malchow-Moeller A, Bryant S, and the COMIC group. Precision requirements in a study of computer-aided diagnosis of jaundice (the COMIC study). Scand J Clin Lab 1980;40 Suppl 155:125-128.
27. Hørder M, Hyltoft Petersen P, Groth T, Gerhardt W. Influence of analytical quality on the diagnostic power of a single S-CK B test in patients with suspected acute myocardial infarction. Scand J Clin Lab Invest 1980;40 Suppl 155:95-100.
28. Hyltoft Petersen P, Fraser CG, Westgard JO, Lytken Larsen M. Analytical goal-setting for monitoring patients when two analytical methods are used. Clin Chem 1992;38(11):2256-2260.
29. Hyltoft Petersen P, Gowans EMS, Blaabjerg O, Hoerder M. Analytical goals for the testimation of non-Gaussian reference intervals. Scand J Clin Lab Invest 1989;49:727-737.
30. Hyltoft Petersen P, Groth T, de Verdier C-H. Principles for Assessing Analytical Quality Specifications and their Use in Design of Control Systems. Upsala J Med Sciences 1993;98 (3):In Press
31. Hyltoft Petersen P, Groth T, Hjelm M. Characterization of increased synthesis of acute phase proteins from plasma concentration measurements: correction of sampling errors and of disturbing variations of plasma volume, exchange fluxes and catabolic rates in the clinical situation. J Clin Computing 1979;8:180-201.
32. Hyltoft Petersen P, Hoerder M. Ways of assessing quality goals for diagnostic tests in clinical situations. Arch Pathol Lab Med 1988;112:435-443.
33. Hyltoft Petersen P, Hoerder M, Blom M, Frimodt-Moeller P, Prag J. An example of choice of test quality in clinical decision making based on two consecutive measurements. Scand J Clin Lab Invest 1984;44 Suppl 171:47-58.
34. Hyltoft Petersen P, Lytken Larsen M, Fraser CG. The Quality Needed for Measuring Glycated Haemoglobin. An application. Upsala J Med Sci 1990;95 (3):185-190.

35. Hyltoft Petersen P, Lytken Larsen M, Hoerder M. Prerequisites for the maintenance of a certain state of health by biochemical monitoring. In: Harris EK, Yasaka T., eds. Maintaining a healthy state within the individual, Amsterdam: Elsevier, 1987:147-158.
36. Hyltoft Petersen P, Lytken Larsen M, Hoerder M. Goals based on clinical application. Scand J Clin Lab Invest 1990;50 suppl 202:71-74.
37. Hyltoft Petersen P, Lytken Larsen M, Hoerder M, Blaabjerg O. Influence of analytical quality and preanalytical variations of measurements of cholesterol in screening programmes. Scand J Clin Lab Invest 1990;50 Suppl 198:66-72.
38. Hyltoft Petersen P, Rosleff F, Rasmussen J, Holbolth N. Studies on the required analytical quality of TSH-measurements in screening for congenital hypothyroidism. Scand J Clin Lab Invest 1980;40 Suppl 155:85-89.
39. Jacobson G, Groth T, de Verdier C-H. Pancreatic iso-amylase in serum as a diagnostic test in different clinical situations: A simulation study. Scand J Clin Lab Invest 1980;40 suppl 155:77-84.
40. Jenny RW. Analytical goals for determination of theophylline concentrations in serum. Clin Chem 1991;37:154-158.
41. Lindqvist U, Groth T, Lööf L, Hellsing K. A hyaluronan-loading test applied to patients with liver and joint diseases. Clin Chim Acta 1992;210:119-132.
42. Linnet K. Choosing quality-control systems to detect maximum clinically allowable analytical errors. Clin Chem 1989;35:284-288.
43. Lott JA, Surufka N, Massion CG. Proficiency testing of serum enzymes based on medical-needs criteria. Arch Pathol Lab Med 1991;115:11-14.
44. Lytken Larsen M, Fraser CG, Hyltoft Petersen P. A comparison of analytical goals for haemoglobin A1c assays derived using different strategies. Ann Clin Biochem 1991;28:272-278.
45. Magar EM. Data Analysis in Biochemistry and Biophysics, New York and London: Academic Press, Inc, 1972:
46. Matzen P, Hilden J, Thomsen C, Malchow-Moeller A, Juhl E. Does test quality influence the outcome of algorithmic classification of jaundiced patients? Scand J Clin Lab Invest 1984;44 Suppl 171:35-46.
47. Wiebe DA, Westgard JA. Cholesterol - a model system to relate medical needs with analytical performance. Clin Chem 1993;39(7):1504-1513.

48. Noerregaard Hansen K, Hyltoft Petersen P, Hangaard J, et al. Early observations of S-myoglobin in the diagnosis of acute myocardial infarction: The influence of discrimination limit, analytical quality, patient's sex and prevalence of disease. Scand J Clin Lab Invest 1986;46:561-569.
49. Ross JW. A theoretical basis for clinically relevant proficiency testing evaluation limits. Sensitivity analysis of the effect of inherent test variability on acceptable method error. Arch Pathol Lab Med 1988;112:421-434.
50. Skendzel LP, Barnett RN, Platt R. Medical useful criteria for analytical performance of laboratory tests. Am J Clin Pathol 1985;83:200-205.
51. Thue G, Sandberg S, Fugelli P. Clinical assessment of haemoglobin values by general practitioners related to analytical and biological variation. Scand J Clin Lab Invest 1991;51:453-459.
52. Tonks DB. A study of the accuracy and precision of clinical chemistry determinations in 170 Canadian laboratories. Clin Chem 1963;9:217-233.
53. van Waeg G, Groth T. Allopurinol kinetics in humans as a means to assess liver function: design of a loading test. Am J Physiol 1989;257:R237-R245.
54. Westgard JO, Burnett RW. Precision Requirements for Cost-Effective Operation of Analytical Processes. Clin Chem 1990;36:1629-1632.
55. Westgard JO, Hyltoft Petersen P, Wiebe A. Laboratory Process Specifications for Assuring Quality in the U.S. National Cholesterol Education Program (NCEP). Clin Chem 1991;37:656-661.
56. Wide L, Dahlberg PA. Quality requirements of basal S-TSH assays in predicting a S-TSH response to TRH. Scand J Clin Lab Invest 1980;40 Suppl 155:101-110.
57. Wiener F, Groth T, Nilsson U. A knowledge-based system for automatic interpretation of an analytical profile of complement factors. J Clin Lab Anal 1989;3:287-295.
58. Wiggers P, Dalhöj J, Hyltoft Petersen P, Blaabjerg O, Hörder M. Screening for haemochromatosis: influence of analytical imprecision, diagnostic limit and prevalence on test validity. Scand J Clin Lab Invest 1991;51:143-148.

Correspondence:
Per Hyltoft Petersen
Department of Clinical Chemistry
Odense University Hospital
DK-5000 Odense C, Denmark