THE PURSUIT OF QUALITY

IN CLINICAL CHEMISTRY

Gedrukt door:  Drukkerij Haveka B.V., Alblasserdam

THE PURSUIT OF QUALITY IN CLINICAL CHEMISTRY

(HET STREVEN NAAR KWALITEIT IN DE KLINISCHE CHEMIE)

PROEFSCHRIFT

Ter verkrijging van de graad van doctor
aan de Erasmus Universiteit Rotterdam
op gezag van de Rector Magnificus
Prof. Dr. P.W.C. Akkermans M.A.
en volgens besluit van het College van Dekanen.
De openbare verdediging zal plaatsvinden op
woensdag 1 februari 1995 om 15.45 uur.

door
Jeroen Dirk Evert van Suijlen
geboren te Zeist.

Promotie-commissie

Promotor	:	Prof. Dr. H.G. van Eijk (Erasmus Universiteit Rotterdam)
Co-promotor	:	Dr. J. Lindemans (Erasmus Universiteit Rotterdam)
Overige leden	:	Prof. dr. J.F. Koster (Erasmus Universiteit Rotterdam)
	:	Prof. J.H.P. Wilson (Erasmus Universiteit Rotterdam)
	:	Prof. dr. P.J. Brombacher (Rijksuniversiteit Limburg)

Men dwaalt zowel door te veel als door te weinig te denken

Niccolini

Het onderzoek dat leidde tot de totstandkoming van dit proefschrift kon alleen worden uitgevoerd dankzij de steun en inzet van de medewerkers van het Instituut **Chemische Pathologie** van de Erasmus Universiteit te Rotterdam en alle medewerkers van de afdeling **Klinische Chemie** van het Academisch Ziekenhuis te Rotterdam.

CONTENTS page

List of abbreviations

AHA	American Heart Association
CBC	complete blood count
Cbl	cobalamin
CCKL	Coördinatie Commissie Kwaliteitsbeheersing Laboratoriumonderzoek Gezondheidszorg
CEN	Organisation Commune Européenne de Normalisation
CLIA	Clinical Laboratory Act
dest	twice distilled water
DM	definitive method
ELISA	enzyme-linked immunosorbent assay
EN	European Norm
FAES	flame atomic emission spectrometry
FDA	Food and Drug Administration
Fol	folate (folic acid)
fte	full-time equivalent
GLP	Good Laboratory Practice
h	hour(s)
Hb	hemoglobin
Hcy	total homocysteine
HPLC	High performance liquid chromatography
Ht	hematocrit
ICU	intensive care unit
IFCC	International Federation of Clinical Chemistry
IHDF	International Health Development Foundation
ISE	ion-selective electrode
ISO	International Organization for Standardization
lab	laboratory
LIS	laboratory information system
MCV	mean cell volume (fl)
min	minute(s)
Mma	methylmalonic acid
Mw	molecular weight

NBSB	National Biological Standards Board
NCCLS	National Committee for Clinical Laboratory Standards
NIST	National Institute for Standards and Technology
no.	number
OC	opinion of clinicians
ODC	oxygen dissociation curve
OR	operating room
PAA	polyacrylamide
POC	point-of-care
POCT	point-of-care testing
ppm	parts per million
QA	quality assurance
QC	quality control
QLP	quality laboratory practices
RIVM	Rijksinstituut voor Volksgezondheid en Milieuhygiëne
RM	reference method
ROC	receiver operating curve
SD	standard deviation
SKZL	Stichting Kwaliteitscontrole Ziekenhuis Laboratoria
SoA	state-of-art
SRM	standard reference material
SSV	system specific value
stat	short turnaround time ("cito")
TAT	turnaround time
TATmax	maximum allowable turnaround time
TQM	total quality management
VC	coefficient of variation
WBC	white blood cell(s)
wsbv	within-subject biological variation

CHAPTER 1

GENERAL INTRODUCTION

Introduction

A clinical chemistry laboratory should have as it's main goal the production of high quality test results. In order to fully appreciate the impact of quality in the clinical chemistry laboratory, the exact meaning of quality in laboratory testing must be understood. In general terms, quality has been best defined as "fitness for use" [1], "conformance to requirements" [2] and "the capacity to achieve goals" [3]. With these definitions in mind, a quality clinical chemistry laboratory test may then be defined as "the most appropriate laboratory test, correctly performed at the lowest cost, and reported within an optimal time frame".
For the user of the clinical chemical information, only a few performance characteristics are of interest: the possible errors in the information (precision and accuracy), the clinical sensitivity and specificity of laboratory tests, the time between the submission of the samples and the results to become available (turnaround time) and the costs associated with the analysis.
These performance characteristics (table 1.1) are direct indicators of a test's quality and are called quality aspects.

table 1.1 *Quality aspects of clinical chemistry laboratory testing*

precision and accuracy
clinical sensitivity and specificity
turnaround time
costs

Specifications or requirements have to be set for each of these quality aspects in order to have quantitative and objective criteria for evaluating laboratory performance. In this chapter the main focus is on precision and accuracy. Clinical sensitivity and specificity are described only briefly, a detailed discussion is beyond the scope of this thesis. Both turnaround time and costs are viewed in the context of centralized versus decentralized testing in the chapters 5 and 6.

The nature of analytical error

Analytical error can be thought of as a combination of two types of error: random error and systematic error. Random error is always present to a measurable degree and can be defined as a scatter (dispersion) of data around a mean value. Random errors are inevitable, but can be reduced, for instance, by repeated measurements. The amount of random error is referred to as imprecision and is defined by the standard deviation (SD) and the coefficient of variation (CV).
Systematic errors are errors within the test system or methodology. These errors occur in only one direction away from the "true value" and cannot be reduced by repeated measurements, but only by removing their cause. The discrepancy between the result of a measurement and the "true value" is termed inaccuracy or bias.

Precision and accuracy

Many strategies have been used to set desirable specifications, particulary for imprecision but also for inaccuracy [4-10].
In summary, there are six different strategies to set quality specifications for imprecision (table 1.2). All of these approaches have disadvantages.
Fractions of reference intervals [11-13]. The first and most favoured approach is that proposed by Tonks [11] namely that the allowable error (2 CV) is:

(¼ reference interval / mean reference interval) x 100.

Different fractions have been proposed as well [12], pointing at the major disadvantage of this strategy: the fraction chosen is more or less empirical. Furthermore, the allowable error depends upon the imprecision and inaccuracy of the method, the nature of the population studied, and the method of data reduction.
Opinions of clinicians [14-18] are subjective in nature, moreover these encompass biological, pre-analytical and analytical variation.

table 1.2 *Strategies to set quality specifications for imprecision*

Fractions of reference intervals
Opinions of clinicians
State-of-art performance
Views of expert individuals and groups
Data on biological variation
Analysis of the effect of analytical performance on interpretation in specific clinical situations

The state-of-art [19][20] is derived from external quality assessment procedures in which the materials may not perform and be treated as patient samples [21][22]. Furthermore, the state of the art is always improving and thus goals continuously become higher although they still may not yet be good enough.

The views of expert individuals and groups [4-10,23,24] are fascinating but generally subjective and sometimes contradictory in nature.

Data on biological variation [25-28] are viewed as best. The theory is that the analytical CV should be:

less than or equal to ½ the within-subject biological variation.

The major advantages of this strategy are that (1) the model is simple to understand and apply and (2) there are many data available on biological variation [29]. Disadvantages are that (1) the goals are based on data on biological variation in the healthy person (and not the diseased patient), (2) there exist little data on long-term biological variation or on biological variation at unusual levels and (3) the fraction ½ is empirical, although chosen so that imprecision adds only 12 % or less to the total observed variability.

Analysis of the effect of analytical performance. Some suggest that different goals are required for analyses performed for different clinical purposes such as screening, monitoring and aiding in diagnosis [30][31][32]. However, the clinical laboratory must, in my opinion, ensure the same goals for imprecision

and inaccuracy of test results regardless of where the tests are performed, the clinical purpose and by whom. The challenge such responsibility entails is great, but not insurmountable.

In conclusion, it is necessary to relate the allowable imprecision to within-subject biological variation or, if that is too stringent, to achievable analytical performance characteristics (that is the state-of-art).

Very few attempts have been made to define specific numerical goals for inaccuracy. Because laboratory tests are now performed not only in the central laboratory but also at the point-of-care (POC), methods should ideally have no bias so that results are comparable in time and place. Gowans and coworkers [33] show that if it is desired to have acceptance of the same reference interval throughout a geographical area, as small as a hospital or as big as Europe, the inaccuracy should not exceed one quarter of the group biological variation.

Taking into account all the advantages and disadvantages of the available techniques for setting quality specifications, and what is currently technically achievable, a working group of the European group for the evaluation of reagents and analytical systems in laboratory medicine has proposed the following quality specifications for analytical systems for clinical chemistry [34]:

Total imprecision should be less than one-half of the average within-subject biological variation or less than the state of art achieved by the best 0.20 fractile of laboratories, whichever is the less stringent. The second approach may be used when data on biological variation do not exist.

Inaccuracy should be less than one-quarter of the group (within- plus between-subject) biological variation or less than one-sixteenth of the reference interval, when data on group biological variation do not exist or less than twice the ideal imprecision, if the above specifications are too demanding.

Quality specifications are already proposed in European countries. For example in the Dutch "Combi" quality assessment scheme [35] goals are based on biological variation or the state-of-art, whichever is considered most realistic.

Clinical sensitivity and specificity

Clinical sensitivity and specificity provide for quantitation of clinical accuracy. The sensitivity of an assay is the fraction of those with a specific disease that the assay correctly predicts. The specificity is the fraction of those without the disease that the assay correctly predicts. These properties can be represented by the receiver operating characteristic (ROC) curve, a plot of sensitivity versus one minus specificity [36]. The procedure for deriving ROC curves and assessment of clinical accuracy were recently published in a guideline by the NCCLS [37]. An ideal diagnostic test would be 100% sensitive, giving positive results in all diseased subjects, and would be 100% specific, giving negative results in all healthy subjects. Factors which increase the specificity of a test tend to decrease the specificity and vice versa. Whether it is desirable to maximize specificity or sensitivity depends on the clinical situation the test is used for.

Quality management

There is a growing recognition of the fact that quality of clinical chemical information must be judged not only by the yardstick of the internal quality parameters precision and accuracy. Total laboratory quality is being evaluated by how well the patient is served, and thus, recognized as a medical service to be judged by medical criteria which are far broader than a proficiency test report or the point on a control chart.

The traditional model for quality management in clinical chemistry includes three major components: (1) Quality laboratory practices (*QLP*), (2) Quality control (*QC*) and (3) Quality assurance (*QA*) [38]. *QLP* provides for policies, procedures and protocols needed for laboratory performance. *QC* refers to statistical quality or process control. *QC* charts are in wide spread use, each measurement procedure in a clinical chemistry laboratory is usually monitored by a periodical statistical control procedure. Common control procedures are Levey-Jennings charts [39] and Westgard's multi-rule *QC* [40]. Less common

procedures include the application of moving averages (e.g. Bull's moving average [41]) and the use of average of (normal) patient samples for *QC* purposes. *QA* is primarily concerned with broader monitoring of laboratory performance, such as patient and specimen identification and turnaround time. Two additional components, quality improvement (*QI*) and quality planning (*QP*) are needed for quality to be properly managed [42] in a so called total quality management (*TQM*) framework.

According to Donabedian, the quality of care is a concept that has many components, which may be grouped under the following seven headings: efficacy, effectiveness, efficiency, optimality, acceptability, legitimacy and equity [43]. With these, two important components are added to quality management: costs and customer service. Taking into account the call for increased service [44][45] quality in clinical chemistry can be expressed in the form of a triangle (figure 1.1), the angles of which are formed by the quality of laboratory practices (*QLP*), the quality of laboratory methods (*QM*) and the quality of laboratory service (*QS*). *QLP* can and should be guaranteed by laboratory certification. In *QM*, apart from the individual quality aspects (see table 1.1), the importance of comparability of methods should be emphasized. This can be pursued by standardisation.

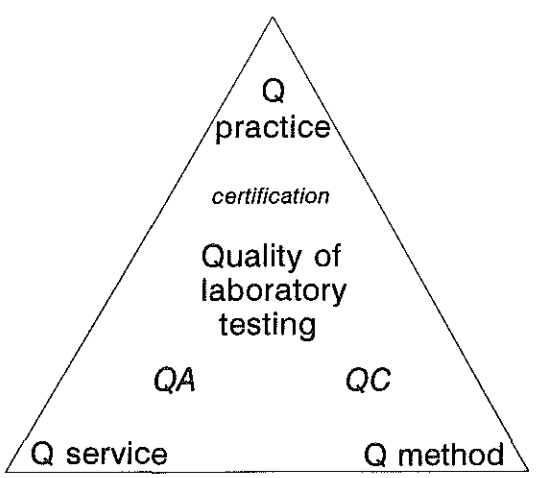


figure 1.1 *Clinical chemistry quality trilogy*

Laboratory certification

Certification has been defined as "the procedure by which a third party gives written assurance that a product, process or service conforms to specific requirements" (International Organization for Standardisation (ISO)) and "the formal recognition that a testing laboratory is competent to carry out specific tests or specific types of tests" (Organisation Commune Européenne de Normalisation (CEN/CENELEC)).

General criteria for laboratory certification have been developed by several organizations [46]. The FDA (Food and Drug Administration) in 1976 issued Good Laboratory Practice (GLP) regulations. GLP concerns the organization and the circumstances under which laboratory research is planned, conducted, monitored, registered and reported. For clinical chemistry laboratories, GLP regulations are only worthwhile attaining, if participation in international research trials is considered. Generally applicable criteria are published as the ISO 9000 series and the CEN/CENELEC EN (European Norm) 45000 series of standards. These certification criteria are developed to guarantee a laboratory's competence and cover such elements as management and organization; use of qualified and continuously trained personnel; availability, calibration and maintenance of equipment; validation of test methods and procedures; availability of standard operating procedures (SOP's); and participation in proficiency testing. All together these elements compose a quality system. A "Code of practice", a guide for the implementation of a quality system has been published by the Dutch CCKL (Coördinatie Commissie bevordering Kwaliteitsbeheersing Laboratoriumonderzoek Gezondheidszorg).

In the Netherlands, clinical chemistry laboratories applying for certification can obtain a quality certificate from STERLAB (based on EN 45000) or from another third party, CCKL*test*. This new independent organization uses the ISO 9000 series of standards as the basis for certification.

Finally, clinical chemistry laboratory certification could become a part of a hospital wide certification (better called accreditation) program. Institutional accreditation is under investigation in the International Health Development Foundation's (IHDF) PACE-project.

Standardisation

An improvement of the agreement between measurement results obtained under different conditions can be achieved by all using the same analytical methods or instruments. This implies that a consensus must be attained on which method or instrument is the most adequate for determination of the analyte concerned. One can imagine that this concept of standardisation has some drawbacks. A more flexible way of standardisation is to describe criteria for the performance of a specific method e.g. the criteria described by the Dutch enzyme commission for the performance of enzyme measurements [47].

Optimal comparability of assay results can only be obtained however, if the analysis is placed in a framework that is based on accuracy. Accuracy can be defined as conformity to the "true value" in the case of measurement. To understand accuracy in chemistry one must understand the chemical measurement process. Chemical measurements are relative in nature. Measurements are generally made by comparison with a reference material for which the concentration of the analyte to be measured has been previously assigned. Multiple methods almost always exist for the measurement of any particulate analyte. These methods have procedural differences and often even have different physical principles on which the measurement is based. Chemical measurements are also frequently subject to significant matrix interferences (non specific measurement). Finally, chemical measurements are always subject to change due to the introduction of new measurement methods. All of these measurement characteristics point to accuracy as the necessary transfer mechanism from method to method. Different methods and different procedures, with their varying biases for measuring any particular analyte, must be understood, and differences must be accommodated if one is to make meaningful, medically relevant interpretation of the resulting data.

How then does one achieve accuracy in a measurement system? The simple answer is by implementation of an overall quality assurance scheme through an accuracy based measurement system. A hierarchy of methods and materials is recognised among the analytical methods and materials needed for reference purposes of this system. In this system, definitive methods provide the closest

approach to "true values". Where definitive methods exist, they are to be used for establishing the accuracy of reference methods for the analytes concerned. Reference methods in turn can be used to evaluate the accuracy of routine methods used in patient care. The most commonly used definitions for definitive and reference methods (for discussion see [48]) are :

Definitive method (DM):

(1) An analytical method that has been subjected to thorough investigation and evaluation for sources of inaccuracy, including nonspecificity. The magnitude of the of the DM's final imprecision and bias, expressed in the uncertainty statement, is compatible with the DM's stated end purpose. The mean DM value is taken as the "true value" (National Committee for Clinical Laboratory Standards (NCCLS) [49]).

(2) A method which after exhaustive investigation, is found to have no known source of inaccuracy or ambiguity (International Federation of Clinical Chemistry (IFCC) [50]).

Reference method (RM):

(1) A thoroughly investigated method in which exact and clear descriptions of the necessary conditions and procedures are given for the accurate determination of one or more property values; the documented accuracy and precision of the method are commensurate with the method's use for assessing the accuracy of other methods for measuring the same property values or for assigning reference method values to reference materials. (NCCLS [49]).

(2) A method which after exhaustive investigation has been shown to have negligible inaccuracy in comparison with its imprecision (IFCC [50]).

Unfortunately not for all analytes definitive or reference methods are available at this moment as can be seen from table 1.3.

table 1.3 *(Candidate) definitive and (candidate) reference methods*

Analyte	Literature
Alkaline Phoshatase	[51]
Alanine aminotransferase	[52-53]
Anti-epileptic drugs	[54]
Apolipoprotein B-100	[55]
Aspartate aminotransferase	[56]
Bilirubin	[57-59]
Calcium	[60-61]
Chloride	[62]
Cholesterol	[63-69]
Clomipramine	[70]
Creatine kinase	[71]
Creatinine	[72-77]
Digoxine	[78]
γ-glutamyltransferase	[79]
Glucose	[80-85]
Glycerol	[86]
Hemoglobin	[87-88]
Inorganic phosphate	[89]
Iron	[90]
Lithium	[91]
Magnesium	[92]
Nickel	[93]
Oxalate	[94-95]
pH	[96]
Phenylalanine	[97]
Potassium	[98-99]
Protein, total	[100]
Sodium	[101-104]
Steroids	[105-110]
Thyroxine	[111-112]
Urea	[113-115]
Uric acid	[116-118]

Laboratory service

For a laboratory test to be useful to its customers, the clinicians, results should be available in a timely manner. Specimen turnaround time (TAT) happens to be the far most important component of laboratory service, this according to the opinions of clinicians. Appropriate goals for specimen TAT need to be developed and should be based on medical necessity and clinical expectation.
From the clinical chemistry point of view attention should be focused not merely on the reduction of the number of tests requested, but also on reducing the amount of inappropriate testing. Inappropriate (including excessive) requesting can lead to harmful further diagnostic investigations, needlessly prolonged hospital stay, poor understanding of results and incorrect actions. The efficiency of the requesting pattern is therefore also a service quality parameter for clinical chemical information.

Scope of the thesis

The studies described in this thesis were undertaken to:

(1) illustrate the importance and the proper use of the traditional quality aspects precision and accuracy and
(2) to focus on quality of laboratory service as a relatively new component of quality and its consequences for laboratory performance, with special emphasis on point-of-care testing.

A specific example illustrates the need for accuracy and comparability of results obtained by different methods for the same analyte. In chapter 2 is demonstrated that the results obtained by different serum ferritin immunoassays are not comparable to each other, even after introduction of a reference preparation for the calibration of serum ferritin immunoassays. The importance of an accuracy based reference system (reference preparation and method) is emphasized and the methodological concept for a serum ferritin reference method is described.

The way in which the results of measurements are reported can affect the accuracy of test results. For the determination of sodium a definitive as well as a reference method are developed (see table 1.3). However, both methods report the assay results as molarity i.e. sodium concentration per litre serum or plasma. Therefore assay results are sensitive to the serum volume occupied by proteins and lipids. This sensitivity may lead to misinterpretation of electrolyte status if serum samples with abnormal protein and or lipid status are concerned. In chapter 3 the development of a proposed reference method for the sodium determination on a plasma water basis is described. Misinterpretation due to abnormal protein (or lipid) status is excluded now because results are presented as sodium molality (concentration per litre plasma water) or sodium molarity after correction for abnormal water contents. The inaccuracy of a routine method is determined by analyzing a series of patient specimens by the method in question and the alternative reference method. The resulting method comparison data (slope and intercept) are a measure of proportional (slope) and constant (intercept) bias at any particular concentration. In chapter 4 the accuracy of 21 routine methods equipped on a multitest clinical chemistry analyzer is evaluated. In addition to this evaluation a procedure for multi-point calibration of the routine methods based on accurate reference method values is presented, to correct possible inaccuracy.
With regard to the quality of laboratory service, the specimen turnaround time and the quality of test requesting patterns are presented in more detail. In chapter 5 a system for continuous monitoring of specimen turnaround time is described. This system is implemented in the clinical chemistry laboratory and used to justify laboratory organization and procedures. If it is not possible to produce information within the desired turnaround time in the central laboratory, decentral testing (point-of-care testing) should be considered. In chapters 6 the author's opinion about how to manage point-of care testing is evaluated. In chapter 7 all quality aspects presented previously pass in review in the evaluation of the first representative of a pH, blood gas and electrolyte analyzers especially developed for point-of-care testing.
In chapter 8 the second important component of laboratory service, the efficiency of test requesting patterns, is demonstrated, using the differential diagnosis of cobalamin deficiency as an example.

References

1. Juran JM, Gryna FM. Quality planning and analysis. McGraw-Hill, New York, 1980.
2. Crosby PB. Quality is free. McGraw-Hill, New York, 1979.
3. Steffen G. Quality medical care. JAMA 1988; 260: 56-61.
4. Fraser CG. Analytical goals in clinical biochemistry. Prog Clin Pathol 1981; 8: 101-122.
5. Fraser CG. Desirable performance standards for clinical chemistry tests. Adv Clin Chem 1983; 23: 299-399.
6. Fraser CG. Analytical goals for glucose analyses. Ann Clin Biochem 1986; 23: 379-389.
7. College of American Pathologists Conference XIII. Evaluation of proficiency testing results for quantitative methods in relation to clinical usefulness. Arch Pathol Lab Med 1988; 112: 313-474.
8. Fraser CG, Hyltoft Petersen P. Desirable standards for laboratory tests if they are to fulfill medical needs. Clin Chem 1993; 39: 1447-1455.
9. Hyltoft Petersen P, Fraser CG. Setting quality standards in clinical chemistry: Can competing models based on analytical, biological and clinical outcomes be harmonized? Clin Chem 1994; 40: 1865-1868.
10. Clinical Chemistry Forum. Accuracy and precision goals in clinical chemistry testing: can they be defined by medical relevance? Clin Chem 1993; 39: 1446.
11. Tonks DB. A study of the accuracy and precision of clinical chemistry determinations in 170 Canadian laboratories. Clin Chem 1963; 9: 217-233.
12. Glick JH. Expression of random analytical error as a percentage of the range of clinical interest. Clin Chem 1976; 22: 475-483.
13. Schoen I, Graham CA, Custer E, Bandi Z, Georgewill D. Clinically useful limits (CUL) criteria as a basis for quality control including minimal and optimal goals for quality control. Am J Clin Pathol 1989; 91: 172-181.
14. Barnett RN. Medical significance of laboratory results. Am J Clin Pathol 1968; 50: 671-676.
15. Skendzel LP. How physicians use laboratory tests. JAMA 1978; 239: 1077-1080.
16. Barret AE, Cameron SJ, Fraser CG. A clinical view of analytical goals in clinical biochemistry. J Clin Pathol 1979; 32: 893-896.
17. Elion-Gerritzen WE. Analytic precision in clinical chemistry and medical decisions. Am J Clin Pathol 1980; 73: 183-195.
18. Skendzel LP, Barnett RN, Platt R. Medically useful criteria for analytical performance of laboratory tests. Am J Clin Pathol 1985; 83: 200-205.
19. Vassault A, Grafmeyer D, Naudin C, Dumont G, Bailly M, Henny J, Gerthardt MF, Georges P. SFBC protocol for the validation of methods (document B, stage 3). Ann Biol Clin 1986; 113, 829-830.
20. Barnett RN. Error limits ans quality control. Arch Pathol lab med 1989; 113: 829-830.
21. Fraser CG, Peake MJ. Problems associated with clinical chemistry quality control materials. CRC Crit Rev Clin Lab Sci 1980; 12: 59-86.
22. Rowan RM, Laker MF, Alberti KGMM. The implications of assaying external quality control sera under special conditions. Ann Clin Biochem 1984; 21: 64-68.
23. Hyltoft Petersen P, Horder M. Influence of analytical quality on test results. Scand J Clin Lab Invest 1992; 52 (suppl 208): 65-88.
24. Lott Ja, Manning NR, Kyler MK. Proficiency testing in a medical-needs context. Clin Chem 1989; 35: 347-354.

25. Cotlove E, Harris EK, Williams GZ. Biological and analytic components of variation in long-term studies of serum constituents in normal subjects III. Physiological and medical implications. Clin Chem 1970; 16: 1028-1032.
26. Elevitch R ed. College of American Pathologists Conference Report. Conference on analytical goals in clinical chemistry, at Aspen, Colorado, IL, 1976.
27. Harris EK. Statistical principles underlying analytical goal-setting in clinical chemistry. Am J Clin Pathol 1979; 72: 374-382.
28. Proceedings of the Sub-Committee on Analytical Goals in Clinical Chemistry, World Association of Societies of Pathology. Analytical goals in clinical chemistry: their relationship to medical care. Am J Clin Pathol 1979; 79: 624-630.
29. Fraser CG. The application of theoretical goals based on biological variation data in proficiency testing. Arch Pathol Lab Med 1988; 112: 404-415.
30. Harris EK. Proposed goals for analytical precision and accuracy in single-point testing: theoretical basis and comparison with data from CAP proficiency surveys. Arch Pathol Lab Med 1988; 118: 416-420.
31. Ross JW. A theoretical basis for clinically relevant proficiency testing limits. Arch Pathol Lab Med 1988; 118: 421-434.
32. Fraser CG, Hyltoft Petersen P, Larsen ML. Setting analytical goals for random analytical error in specific clinical monitoring situations. Clin Chem 1990; 36: 1625-1628.
33. Gowans EMS, Peterson PH, Blaabjerg O, Horder 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.
34. Fraser CG, Hyltoft Petersen P, Ricos C, Haeckel R. Proposed quality specifications for the imprecision and inaccuracy of analytical systems in clinical chemistry. Eur J Clin Chem Clin Biochem 1992; 30: 311-317.
35. Steigstra H, Jansen RTP, Baadenhuijsen H. Cobi scheme: new combined internal/external quality assessement scheme in the Netherlands. Clin Chem 1991; 37: 1196-1204.
36. Metz CE. Basic principles of ROC analysis. Semin Nucl Med 1978; 8: 283-298.
37. National Committe for Clinical Laboratory Standards. Assessment of the clinical accuracy of laboratory tests using receiver operating characteristic (ROC) plots; tentative guideline. NCCLS document GP10-T 1993; vol. 13, no. 28.
38. Westgard JO, Burnett RW, Bowers GN. Quality management science in clinical chemistry: a dynamic framework for continuous improvement of quality. Clin Chem 1990; 36: 1712-1716.
39. Levey S, Jennings ER. The use of control charts in the clinical laboratory. Am J Clin Pathol 1950; 20: 1059-1066.
40. Westgard JO, Barry PL, Hunt MR. A multi-rule Shewhart chart for quality control in clinical chemistry. Clin Chem 1981; 27: 493-501.
41. Bull BS, Elashoff RM, Heilbron DC, Couperus JA. A study of various estimators for the derivation of quality control procedures from patient erythrocyte indices. Am J Clin Pathol 1974; 61: 473-481.
42. Westgard JO, Barry PL.Total quality control: evolution of quality control systems in health care laboratories. Lab Med 1989; 20: 377-384.
43. Donabedian A. The seven pillars of quality. Arch Pathol Lab Med 1990; 114: 1115-1118.

44. Schuler RS. Customerizing the clinical laboratory. Repositioning for enhanced service and competitive advantage. Clin Lab Manag Rev 1989; 5: 303-310.
45. O'Connor SJ. Service quality: understanding and implementing the concept in the clinical laboratory. Match service quality to customer expectations. Clin Lab Manag Rev 1989; 6: 329-335.
46. Van Suijlen JDE, Leijnse B. Vademecum: Termen, definities, richtlijnen en algemene informatie bij het CCKL symposium "Labcertificatie in de gezondheidszorg?", maart 1989; Utrecht.
47. Enzym commissie van de NVKC. Tijdschrift NVKC 1987; 12: 231-236 and 1989; 14: 180.
48. Leijnse B. Definitive methods- reference methods (absolute methods?): their important impact on clinical chemistry. Ann Clin Biochem 1982; 19: 289-294.
49. National Committee for Clinical Laboratory Standards. Sourcebook of reference methods, materials and related information for the clinical laboratory; proposed guideline. NCCLS document NRSCL 12-P 1994; vol. 14, no. 1.
50. Approved recommendation (1978) on quality control in clinical chemistry. J Clin Chem Clin Biochem 1980; 18: 69-77.
51. Tietz NW, Rinker AD, Shaw LM. IFCC methods for the measurement of catalytic concentration of enzymes. Part 5: IFCC method for alkaline phosphatase. J Clin Chem Clin Biochem 1983; 21: 731-748.
52. Bergmeyer HU, Horder M, Rej R. Approved recommendation on IFCC methods for the measurement of catalytic concentration of enzymes. Part 3 IFCC method for alanine aminotransferase. J Clin Chem Clin Biochem 1986; 24: 481-495.
53. Okorodudu AO, Pelletier PR, Valcour AA, Bowers GN, McComb RB. Evaluation of the IFCC reference method for alanine aminotransferase: spurious blank ALT activity due to contamination of D-alanine with L-alanine, and recommendations for a correction. Clin Chem 1989; 35: 153-156.
54. Juergens U, May T, Hillenkoetter K, Rambeck B. Systematic comparison of three basic methods of sample pretreatment for high-performance liquid chromatographic analysis of antiepileptic drugs; using gas chromatography as a reference method. Therapeutic Drug Monitoring 1984; 6: 334-343.
55. Albers JJ, Lodge MS, Curtiss LK. Evaluation of a monoclonal antibody-based enzyme-linked immunosorbent assay as a candidate reference method for the measurement of apolipoprotein B-100. J Lipid Res 1989; 30: 1445-1458.
56. Bergmeyer HU, Horder M, Rej R. Approved recommendation on IFCC methods for the measurement of catalytic concentration of enzymes. Part 2 IFCC method for aspartate aminotransferase. J Clin Chem Clin Biochem 1986; 24: 487-510.
57. Doumas BT, Kwok-Cheung PP, Perry BW, Jendrzejczak B, McComb RB, Schaffer R, Hause L. Candidate reference method for determination of total bilirubin in serum: development and validation. Clin Chem 1985; 31: 1779-1789.
58. Hajzer S. Comparison of direct spectrophotometric determinations of bilirubin with candidate reference method in sera of newborns. J Clin Chem Clin Biochem 1989; 27: 445-449.
59. Schlebusch H, Axer K, Schneider C, Liappis N, Röhle G. Comparison of five routine methods with the candidate reference method for the determination of bilirubin in neonatal serum. J Clin Chem Clin Biochem 1990; 28: 203-210.

60. Cali JP, Mandel J, Moore L, Young DS. A reference method for the determination of calcium in serum. NBS 1974; NBS special publication 260-36. US Government Printing Office, Washington DC.
61. Moore LJ, Machlan LA. High accuracy determination of calcium in blood serum by isotope dilution mass spectrometry. Anal Chem 1972; 44: 2291-2296.
62. Velapoldi RA, Paule RC, Schaffer R, Mandel J, Murphy TJ, Gramlich JW. A reference method for determination of chloride in serum. NBS 1979; NBS Special publication 260-67. US Government Printing Office, Washington DC.
63. Cohen A, Hertz HS, Mandel J, Paule RC, Schaffer R, Sniegoski LT, Sun T, Welch MJ, White E. Total serum cholesterol by isotope dilution mass spectrometry: A candidate definitive method. Clin Chem 1980; 26: 854-860.
64. Pelletier O, Wright LA, Breckenridge WC. Isotope dilution mass spectrometry of serum cholesterol with [3,4-^{13}C] cholesterol: proposed definitive method. Clin Chem 1987; 33: 1403-1411.
65. Ellerbe P, Meiselman S, Sniegoski LT, Welch MJ, White E. Determination of serum cholesterol by a modification of the isotope dilution mass spectrometric definitive method. Anal Chem 1989; 61: 1718-1723.
66. Abell LL, Levy BB, Brodie B, Kendall FE. A simplified method for the estimation of total cholesterol in serum and demonstration of its specifity. J Biol Chem 1952; 195: 357-366.
67. Derks HJ, van Heiningen A, Koedam HC. Gas-chromatographic determination of cholesterol in serum: candidate reference method. Clin Chem 1985; 31: 691-694.
68. Ellerbe P, Myers GL, Cooper GR, Hertz HS, Sniegoski LT, Welch MJ, White E. A comparison of results for cholesterol in human serum obtained by the reference method and by the definitive method of the National Reference System for cholesterol. Clin Chem 1990; 36: 370-375.
69. Bernett JT, Akins JR, Cooper GR, Poulouse AK, Meyers GL, Sampson EJ. Factors influencing the accuracy of the National Reference System total cholesterol reference method. Clin Chem 1991; 37: 2053-2061.
70. Gaskell SJ. Gas chromatography high-resolution mass spectrometry as a reference method for clomipramine determination. Postgrad Med J 1980; 56 (suppl 1): 90-93.
71. Horder M, Elser RC, Gerhardt W, Mathieu M, Sampson EJ. Approved recommendation on IFCC methods for the measurement of catalytic concentration of enzymes. Part 7 IFCC method for creatine kinase. Eur J Clin Chem Clin Biochem 1991; 29: 435-456.
72. Siekmann L. Determination of creatinine in human serum by isotope dilution-mass spectrometry. Definitive methods in clinical chemistry IV. J Clin Chem Clin Biochem 1985; 23: 137-144.
73. Welch MJ, Cohen A, Herts HS, Kwokei JN, Schaffer R, Van der Lijn P, White E. Determination of serum creatinine by isotope dilution mass spectrometry as a candidate definitive method. Anal Chem 1986; 58: 1681-1685.
74. Hartmann L, Doree G, Ambert JP. Gas-liquid chromatography reference method for the assay of urinary creatinine. J Chromatography 1983; 273: 235-243.
75. Rosano TG, Ambrose RT, Wu AH, Swift TA, Yadegari P. Candidate reference method for determining creatinine in serum: method development and interlaboratory validation. Clin Chem 1990; 36: 1951-1955.
76. Zwang LJ, Blijenberg BG. Assessment of a selected method for creatinine with special emphasis on bilirubin interference. Eur J Clin Chem Clin Biochem 1991; 29: 795-800.

77. Linnet K, Bruunshuus I. HPLC with enzymatic detection as a candidate reference method for serum creatinine. Clin Chem 1991; 37: 1669-1675.
78. Reh E, Jork H. Determination of digoxin in human serum. Comparison of some immunological assays with an HPLC candidate reference method. Fresenius Z Anal Chem 1987; 327: 105-110.
79. Shaw LM, Stromme JH, London JL, Theodorsen L. IFCC method for the measurement of catalytic concentration of enzymes. Part 4 IFCC method for γ-glutamyltransferase. J Clin Chem Clin Biochem 1983; 21: 633-646.
80. Schaffer R, Mandel J, Sun J, Cohen A, Hertz HS, Neese JW. Standard reference materials: Evaluation by an ID MS method of the AACC reference method for serum glucose. NBS 1982. NBS special publication 260-80. US Government Printing Office, Washington, DC.
81. White E, Welch MJ, Sun T, Sniegoski LT, Schaffer R, Hertz HS, Cohen A. The accurate determination of serum glucose by isotope dilution mass spectrometry-two methods. Biomed Mass Spectr 1982; 9: 395-405.
82. Pelletier O, Arratoon C. Precision of glucose measurements in control sera by isotope dilution mass spectrometry: proposed definitive method compared with a reference method. Clin Chem 1987; 33: 1397-1402.
83. Schumann G, Buttner J. Determination of glucose in reference materials by isotope dilution mass spectrometry. Fresenius Z Anal Chem 1988; 330: 424-425.
84. Neese JW, Duncan P, Bayse D, Robinson M, Cooper T, Stewart C. Development and evaluation of a hexokinase / glucose-6-phosphate dehydrogenase procedure for use as a national glucose reference method. HEW publication no (CDC) 77-8330. HEW USPHS 1976, Centers for Disease Control.
85. Magni F, Paroni R, Bonini PA, Kienle MG. Determination of serum glucose by isotope dilution-mass spectrometry: candidate definitive method. Clin Chem 1992; 38: 381-385.
86. Siekmann L, Schönfelder A, Siekmann A. Isotope dilution mass spectrometry of total glycerol in human serum - a reference method in clinical chemistry. Fresinius' Z Anal Chem 1986; 324: 280-281.
87. Van Kampen EJ, Zijlstra WG. Determination of hemoglobin and its derivates. Adv Clin Chem 1965; 8: 141-187.
88. Recommendations for reference method for haemoglobin in human blood (ICSH standard 1986) and specifications for international haemiglobincyanide reference preparation. International Standard. Clin Lab Haematol 1987; 9: 73-79.
89. Schumann G, Büttner J. Highly accurate determinations of orthophosphate using high performance ion chromatography, proposed as a reference method. J Clin Chem Clin Bioche 1989; 27: 766-777.
90. Valcour AA, Krzymowski G, Onoroski M, Bowers GN, McComb RB. Proposed reference method for iron in serum used to evaluate two automated iron methods. Clin Chem 1990; 36: 1789-1792.
91. Velapoldi RA, Paule RC, Schaffer R, Mandel J, Machlan LA, Garner EL, Rains TC. A reference method for the determination of lithium in serum. NBS 1980; NBS special publication 260-69 . US Government Printing Office, Washington, DC.
92. Külpmann WR, Ruschke D, Büttner J, Paschen K. A candidate reference method for the determination of magnesium in serum. J Clin Chem Clin Bioche 1989; 27: 33-39.
93. Brown SS, Nomoto S, Stoeppler M, Sundermann FW. IUPAC reference method for analysis of nickel in serum and urine by electrothermal atomic absorption spectrometry.

Clin Biochem 1981; 14: 295-299.
94. Koolstra W, Wolthers BG, Hayer M, Elzinga H. Development of a reference method for determining urinary oxalate by means of isotope dilution-mass spectrometry (ID-MS) and its usefulness in testing existing assays for urinary oxalate. Clin Chim Acta 1987; 170: 227-235.
95. Koolstra W, Wolthers BG, Hayer M, Rutgers H. An improved high performance liquid chromatographic method for determining urinary oxalate making use of an ID-MS reference method. Clin Chim Acta 1987; 170: 237-243.
96. Maas AHJ, Weisberg HF, Burnett RWQ, Müller-Plathe O, Wimberly PD, Zijlstra WG, Durst RA, Siggaard-Andersen O. International Federation of Clinical Chemistry. Reference method (1986) for pH measurement in blood. J Clin Chem Clin Biochem 1987; 25: 281-289 and Clin Chim Acta 1987; 165: 97-109.
97. Lynes GW, Hjelm M. Phenylalanine in whole blood and plasma-a candidate reference method. J Chromatogr 1991; 562: 213-224.
98. Velapoldi RA, Paule RC, Schaffer R, Mandel J, Machlan IA, Gramlich JW. A reference method for determination of potassium in serum. NBS 1979; NBS special publication 260-63. US Government Printing Office, Washington, DC.
99. Gramlich JW, Machlan LA, Brletic KA, Kelly WR. Thermal-ionization isotope-dilution mass spectrometry as a definitive method for determination of potassium in serum. Clin Chem 1982; 28: 1309-1313.
100. Doumas BT, Bayse DD, Carter RJ, Peters T, Schaffer R. A candidate reference method for determination of tatal protein in serum. I. development and validation. Clin Chem 1981; 27: 1642-1650.
101. Velapoldi RA, Paule RC, Schaffer R, Mandel J, Moody JR. A reference method for the determination of sodium in serum. NBS 1978; NBS special publication 260-60. US Government Printing Office, Washington, DC.
102. Schaffer R, Velapoldi RA, Paule RC, Mandel J, Bowers GN, Copeland BE, Rodgerson DO, White JC. A multilaboratory-evaluated reference method for the determination of serum sodium. Clin Chem 1981; 27: 1824-1828.
103. Lemmens AA, Reijenga JC, Everaerts FM, Janssen RT, Hulsman JA, Meijers CAM. Isotachophoresis as a candidate reference method in analytical chemistry. Determination of sodium in serum. J Chromatography 1985; 320: 193-197.
104. Van Suijlen JDE, Berrevoets CA, Leijnse B. A candidate reference method for coupled sodium-water determination in human serum. J Clin Chem Clin Biochem 1990; 28: 817-824.
105. Siekmann L, Röhle G. Reference methods for the determination of steroids and uric acid by use of isotope dilution-mass spectrometry. Fresenius Z Anal Chem 1986; 324: 208-209.
106. Siekmann L. Determination of steroid hormones by the use of isotope dilution-mass spectrometry: A definitive method in clinical chemistry. J Steroid Biochemistry 1979; 11: 117-123.
107. Siekmann L, Breuer H. Determination of cortisol in human plasma by isotope dilution-mass spectrometry. Definitive methods in clinical chemistry I. J Clin Chem Clin Biochem 1982; 20: 883-892.
108. Jonckheere JA, De Leenheer AP. Candidate definitive method for the determination of cortisol in human serum. Biomedical Mass Spectrometry 1983; 10: 197-202.

109. Siekmann L. Determination of oestradiol-17 beta in human serum by isotope dilution-mass spectrometry. Definitive methods in clinical chemistry II. J Clin Chem Clin Biochem 1984; 22: 551-557.
110. Patterson DG, Patterson MB, Culbreth PH, Fast DM, Holler JS, Sampson EJ, Bayse DD. Determination of steroid hormones in human serum by isotope dilution-mass spectrometry: a candidate definitive method for cortisol. Clin Chem 1984; 30: 619-626.
111. Moeller B, Balk O, Bjoerkhem I. Isotope dilution-mass spectrometry of thyroxin proposed as a reference method. Clin Chem 1983; 29: 2106-2110.
112. Siekmann L. Measurement of thyroxine in human serum by isotope dilution-mass spectrometry . Definitive methods in clinical chemistry V. Biomed Environ Mass Spectrometry 1987; 14: 683-688.
113. Welch MJ, Cohen A, Hertz HS, Ruegg FC, Schaffer R, Sniegoski LT, White E. Determination of serum urea by istope dilution-mass spectrometry as a candidate definitive method. Anal Chem 1984; 56: 713-719.
114. Sampson EJ, Baird MA, Burtis CA, Smith EM, White DL, Bayse DD. A coupled-enzyme equilibrium method for measuring urea in serum: optimization and evaluation of the AACC study group on urea candidate reference method. Clin Chem 1980; 26: 816-826.
115. Pei P, Steiner FP, Vonderschmitt DJ. Isotachophoretic determination of urea-ammonium in plasma: a candidate reference method. J Clin Chem Clin Biochem 1990; 28: 447-451.
116. Siekmann L. Definitive methods in clinical chemistry III. Determination of uric acid in human serum by isotope dilution-mass spectrometry. J Clin Chem Clin Biochem 1985; 23: 129-135.
117. Kock R, Delvoux B, Tillmanns U, Greiling H. A candidate reference method for the determination of uric acid in serum based on high performance liquid chromatography, compared with an isotope dilution-gas chromatography-mass spectrometer method. J Clin Chem Clin Biochem 1989; 27: 157-162.
118. Ellerbe P, Cohen A, Welch MJ, White E. Determination of serum uric acid by isotope dilution-mass spectrometry as a new candidate definitive method. Anal Chem 1990; 62: 2173-2177.

CHAPTER 2

ACCURACY OF SERUM FERRITIN DETERMINATIONS IN TISSUE PREPARATIONS AND HUMAN SERUM

This chapter is based on:

J Clin Chem Clin Biochem 1990; 28: 43-48.

Summary

Introduction of the National Biological Standards Board (NBSB) 80/602 reference preparation for the calibration of ferritin immunoassays has reduced the inter-assay variability, and represents the first important step towards standardisation of ferritin immunoassays. However, our investigations show that comparison between assay results is still impossible. This is caused by differences in assay methodology and performance, differences in the specificities of the antibodies and antigens used and possibly interfering substances.

We conclude that all kits detect mainly the more basic isoferritins in serum, but all isoferritins (acidic, intermediate and basic) are systematically underestimated. Since we also showed immunologic differences between reference or kit standards and serum ferritin, we conclude that the kits evaluated have poor accuracy.

To diminish inter-assay variability and to increase the accuracy of serum ferritin determinations, a method is needed that quantifies all basic, intermediate and acidic isoferritins in serum under normal and pathological circumstances, in order to measure the true ferritin concentration in serum.

The introduction of a reference method in combination with the NBSB 80/602 human liver reference preparation, is the second important step towards accurate standardisation of ferritin immunoassays.

Introduction

The study of accuracy in clinical chemistry is a difficult problem especially if all information on composition of reagents and analytical procedures is manufacturer dependent. Serum ferritin determination is a good example for such a black box assay. It has been shown that compatibility between assay results from different kits is lacking [1][2][3][4][5]. The variability in determinations with the various kits has been attributed to the immunological heterogeneity of (serum) isoferritins [6], the different specificities of the antisera [1], the methodology used [7] and the lack of an international standard [2][8].

It is generally accepted that the introduction of the NBSB first international standard for human liver no. 80/602 [9] improves the accuracy and inter-assay variability of serum ferritin determinations [10]. It is beyond doubt that the dispersion of the values obtained using different commercial kits is reduced after calibration with the NBSB reference standard. But are the assays more accurate?

In our opinion the introduction of the NBSB reference preparation is only the first step towards solving the accuracy problem. Because this reagent is used on a large scale to calibrate ferritin assays, the decreased inter-assay variability is only coupled with increased accuracy if the standard ferritin and serum ferritin are immunologically identical, and all isoferritins present in serum are detected.

We investigated inter-assay variability after introduction of the NBSB standard by comparing seven immunoassay kits for serum ferritin determination available in the Netherlands. The kits were used to evaluate serum ferritin concentrations in 40 sera with serum ferritin levels between 20 and 800 μg/l. To study the possible influence of structural and immunological differences between tissue and serum ferritins on accuracy, purified human heart (acidic isoferritins), human placental (intermediary isoferritins), human liver (basic isoferritins) and human serum ferritins were tested. Other possible interferences influencing accuracy (iron contents and isolation procedure of the standards [11]) are also taken into account. The differences from the nominal ferritin concentrations based on protein contents are shown as % cross reactivity in a so-called kit index. This kit index shows which isoferritins are measured accurately and indicates a further need for standardisation.

Materials and methods

Ferritin standards and antibodies

The first international reference human liver standard no. 80/602 was obtained from the NBSB (Hertfordshire, England). Ferritin from human liver type IV (cat no. F6754) and ferritin from human heart type VII (cat no. F4131) were obtained from Sigma chemical company (St. Louis, USA). Ferritin from human liver X551 was from DAKO (Copenhagen, Sweden).

Anti human liver ferritin antibodies were obtained from DAKO immunoglobulins (Denmark) and anti placental ferritin antibodies were a gift from Behringwerke AG (Marburg, Germany).

Serum ferritin kits

All kits used are non-competitive sandwich assays, and are named A - G in this chapter. For more details see [12].

Because there is no reference method available we used kit A, which is in routine use in our laboratory and is performed on an automated system, as a reference.

Ferritin preparation

Ferritins were purified from human tissue in essentially the same way as described by *May* and *Fish* [13], using standard procedures. Serum ferritin was purified from a serum pool from healthy persons. Purification was essentially as described for tissue ferritin, except the ultracentrifugation step which was replaced by affinity chromatography on CNBr-activated Sepharose 4B in two separate columns using purified antibody preparations against human liver ferritin and against human placental ferritin [14].

The purity of the preparations was examined with gradient-pore polyacrylamide gel electrophoresis using precast PAA 4/30 gels (Pharmacia, Uppsala, Sweden). Electrophoresis conditions used are recommended by the manufacturer. The gels are stained for protein with CBB R250 and stained for iron in potassium ferrocyanide. The same bands appeared after staining for iron and protein. Protein was determined by the method of *Lowry* according to *Worwood* [15].

Statistical analysis

Inter-assay comparisons were made by regression analysis according to *Passing and Bablok* [16]. Significant differences in slope and or intercept are considered at the 95% confidence level. Differences between means of series of assay results were investigated with the *Student's* paired t-test (significant difference if $p < 0.05$).

Results

Correlation data between kit A and the kit in question (for determination of ferritin in serum) are shown in table 2.1.

Table 2.2 shows the mean values of all sera measured together with the inter-assay variability presented as coefficient of variation, %CV.

Table 2.3 shows the influence of calibration with NBSB 80/602 on inter-assay variability. Our results indicate a significant decrease of inter-assay variability after introduction of the NBSB 80/602 reference standard. These results agree with other investigations [9,10].

table 2.1 *Method comparison by Bablok and Passing regression calculation between results obtained with kit A and the kit in question*

Kit	y = ax + b			
	a	b	r	Significances*)
B	1.072	-11.03	0.998	a,b
C	0.967	-10.87	0.999	a,b
D	1.044	-13.50	1.000	a,b
E	0.798	4.59	0.996	a,b
F	1.061	-6.54	0.999	a,b
G	1.036	3.05	0.999	a,b

*) a: significant difference from slope = 1.0,
b: significant difference from intercept = 0.0, in regression equations kit A versus kit in question (95% confidence level, calculation according to *Passing* and *Bablok*).

table 2.2 *Inter-assay variability after introduction of the NBSB 80/602 international human liver reference preparation in different concentration ranges*

Concentration range	Concentration (μg/l)	Mean value (μg/l)	CV (%)
Low reference	15 - 50	32	19
Medium reference	80 - 120	93	13
High reference	200 - 250	241	9
Pathological	> 500	612	10

table 2.3 *Inter-assay variability before and after introduction of the NBSB 80/602 international human liver reference preparation*

Source	Calibration with NBSB 80/602 standard	Between-kit variability (%)
van Oost [1]	no	52
Wood [2]	no	44
Grail [4]	no	56
Dunn [5]	no	16
Iacobello [10]	no	32
Authors' laboratory	yes	13

figure 2.1 shows the respective regression lines for the relation between the nominal concentration of reference standard 80/602 and the results of the analysis with each of the seven kits. Results from kits C, D, E and F differ significantly from the nominal values.

figure 2.2 shows the respective regression lines for the nominal concentration of purified human serum ferritin and the results of the analysis with each of the seven kits. Results from kits B, C, D, E and G differ significantly from the nominal values based on the protein contents.
figure 2.3 shows slight differences in relative migration values between some purified ferritin preparations.
Correlation data between nominal values for all ferritin preparations and the values measured with the different kits are expressed as % cross reactivity in table 2.4.
DAKO liver and SIGMA liver preparations are underestimated in all seven kits, the % reactivity for DAKO liver lying between 55 and 74 %, that for SIGMA liver between 49 and 98 % (table 2.4). This is remarkable because all kits use liver or spleen ferritin as the antigen source for their determinations and so reactivities between 85 (spleen) and 100 % (liver) should be expected. The situation is even more complicated for the NBSB 80/602 reference preparation, which showed reactivities between 76 and 112 % (figure 2.1, table 2.4). Kit A is the only kit with 100 % reactivity with the NBSB 80/602 reference preparation. The assay results of 4 kits differ significantly from the nominal NBSB 80/602 values, although all 4 kits are calibrated against it.

Discussion

Inter-assay variability The introduction of NBSB reference standard 80/602 decreased the inter-assay variability (table 2.3) of serum ferritin determinations, but comparison between assay results from different kits is not possible (table 2.1). Although there is a good correlation ($r > 0.98$) between the various kits, there are still considerable differences in the measured absolute values.
These present findings indicate that a reference standard is not the only relevant factor in the variability of serum ferritin determination. The observed discrepancies among tissue ferritin preparations may be explained by different procedures used to isolate and purify the tissue ferritins, leading to slight differences in electrophoretic behaviour (figure 2.3).

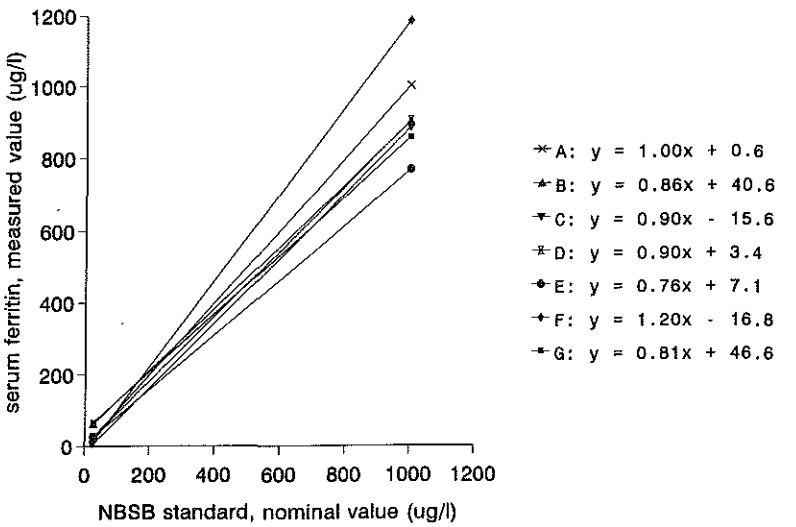


figure 2.1 *Comparison between the NBSB 80/602 standard nominal concentration and the measured concentration using the kit standards*

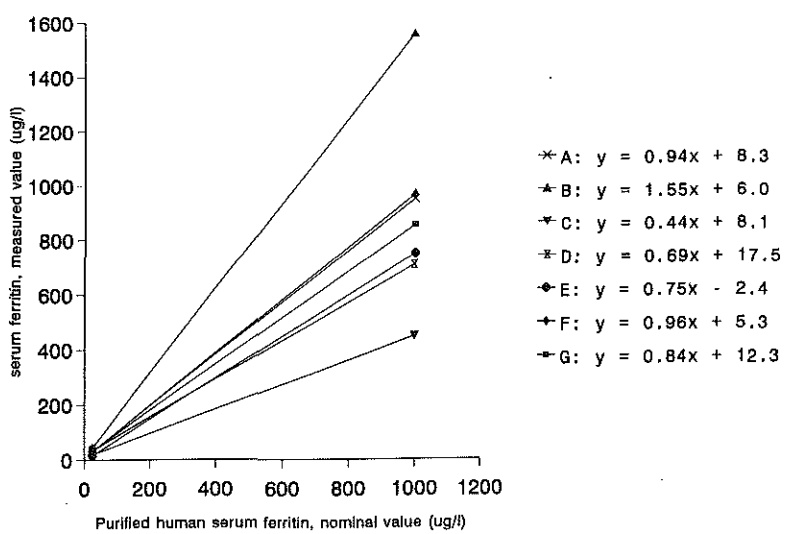
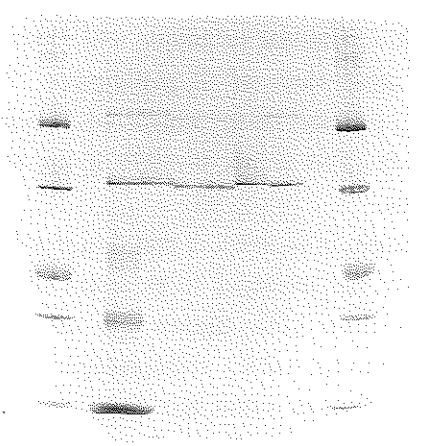


figure 2.2 *Comparison between purified human serum ferritin nominal concentration and the measured concentration using the kit standards*

table 2.4 *Kit index: relation between the nominal ferritin concentrations of purified tissue or serum and the concentrations measured with the different kits presented as % cross-reactivity*

Ferritin source	A	B	C	D	E	F	G
WHO 80/602	100	86	90	91	76	112	81
Dako liver	74	60	55	66	71	62	61
Sigma liver	86	98	68	74	85	49	89
Sigma heart	18	26	13	16	16	12	26
Placental	37	44	26	31	27	20	33
Placental (apo)	31	47	19	19	22	21	29
Serum	94	155	44	69	75	96	84



1 2 3 4 5 6 7 8

figure 2.3 *Gradient pore 4-30% polyacrylamide gel electrophoresis. Relative molecular mass markers: thyroglobulin (669 000), ferritin (440 000), catalase (232 000), lactate dehydrogenase (140 000) and albumin (67000) in the first and the last lanes. 2nd lane: NBSB 80/602, 3th lane: Dako liver ferritin, 4th lane: Sigma liver ferritin, 5th lane: Sigma heart ferritin , 6th lane: placental ferritin, 7th lane: purified serum ferritin*

In addition to the different assay methodologies and specificities of the antibodies used, iron contents is also a contributor to variability. Van Oost et al. [1] found a twofold increase in antibody binding in 2 of 4 kits tested after removal of iron from liver ferritin. Our results however, show decreased ferritin detection after removal of iron in 5 out of 7 kits (table 2.4; placental apo-ferritin).

The immunological heterogeneity of serum ferritin is a very important contributory factor to variability. Its structural and immunological characteristics appear to change with iron status and presence of pathology [17]. We showed reactivities for purified serum ferritin between 44 and 155% (table 2.4).

The observed discrepancies between kit results for sera and tissue or serum purifications could also be explained by quantitative differences between kit standards. Despite calibration against NBSB 80/602 four kits differ significantly from the nominal ferritin values. This indicates that comparison between assay results from different kits is only possible if all kits are standardised versus each other. Differences in methodology, source and purification of standards, specificity of antibodies and performances have to be eliminated.

Laboratories using such a particular technique will report results that are normally distributed around the mean value of all participants. This approach can lead to results that are, in fact, accurate being considered inaccurate because they are not close to the group mean. This is not possible if the assays used are essentially free of systematic errors.

Accuracy A basic principle in analysis is that dose/response relationships between an accepted standard and the substance to be tested show no significant deviation from linearity or parallelism [18]. figures 2.1 and 2.2 show dose (nominal value)/response (measured kit value) relationships for NBSB 80/602 reference preparation and purified human serum ferritin. Statistically significant deviations from parallelism are present in kits B, C, D and F. One possible explanation for these deviations from parallelism is that purified human serum ferritin is immunologically different from the NBSB 80/602 reference preparation. This explanation seems very likely because the inter-assay variability for purified serum ferritin determination is much larger than can be explained by the differences found for NBSB 80/602 and other tissue preparations. Because this serum ferritin was prepared from serum of healthy individuals (serum ferritin

between 15 and 300 μg/l), we conclude that under normal circumstances serum ferritin is immunological different from the tissue ferritins used as antigen sources in the different kits. This conclusion disagrees with a collaborative study organised by the International Committee for Standardisation in Haematology (ICSH), which suggests there are no immunological differences between the two ferritins in question.

The ferritin in human tissues contains two types of subunits: H and L [19]. H subunits have a relative molecular mass of 21,000 and are found in the more acidic isoferritins (pI 4.5-5.0) in heart [20], red blood cells [21] and other tissues. L subunits have a relative molecular mass of 19,000 and predominate in the more basic isoferritins (pI 5.3-5.8) of liver, spleen and placenta. The H-rich isoferritins are immunologically different from the L-rich isoferritins. The origin of serum ferritin is not certain. Immunologically serum ferritin resembles liver ferritin and it is detected with antibodies to ferritins rich in L subunits.

Table 2.4 shows that all kits detect mainly the more basic isoferritins in serum, and that basic, acidic and intermediate [22] isoferritins are systematically underestimated, although acidic isoferritins are detectable in 20 % of normal sera and in 65% of sera from patients with malignancy [23]. In many other circumstances, for example in diseases involving liver damage (with serum ferritin spectrum similar to the whole spectrum of liver ferritin) 2-40% (table 2.4) of the serum ferritin is systematically underestimated owing to its acidic characteristics. In iron deficiency underestimation can also be expected because a shift towards the more acidic isoferritins in serum is seen. Only in severe hemochromatosis no acidic isoferritins can be detected in serum, so that theoretically accurate assay values can be expected. However, our observations, in combination with differences in immunological properties between tissue ferritins used as antigens and serum ferritins to be analyzed, indicate that all immunological ferritin kits tested have poor accuracy.

Reference method Proper standardisation of ferritin assays is a major problem because an accuracy base is needed. Use of the mean values from different assays as a basis for accuracy assumes that the results from all methods will be randomly distributed around the true value. Regarding our results, this seems very unlikely.

The reference method concept [24] can be used to provide a basis for accuracy. The importance of such methods should increase with the accelerated development of new methods based on different methodologies. *Worwood* et. al. [25] developed a simple ELISA for use as a reference assay of serum ferritin using the 80/602 reference preparation as primary reference material and standardised reagents including two lyophilized antibody preparations to human liver ferritin. Use of this assay will certainly decrease inter-assay variability but not the accuracy of serum ferritin determinations. The latter can only be achieved if a proper mixture of anti-H and anti-L chain monoclonal antibodies is used to assure that not only the more basic isoferritins but all acidic, intermediate and basic isoferritins present in serum under normal and pathological conditions are detected. An example of this concept is presented in literature [26]. However, the high imprecision precludes the use of the assay described as a reference method. Moreover, the relative values for H and L subunits as indicated by immunoreactivity showed not to increase in proportion to the H and L subunit content in all isoferritins [27].

A totally different concept for a reference method is the detection of ferritin subunits after denaturation and separation by HPLC [28]. If the international reference preparation 80/602 is used for calibration and monoclonal antibodies against purified human serum ferritin subunit preparations are used for detection, assay results are no longer influenced by (serum)ferritin immunological heterogeneity and accurate and precise assay results can be expected.

References

1. Van Oost BA, Willekens FLA, Van Neerbos BR, Van de Belt B. Implications of using different tissue ferritins as antigens for ferritin in serum: four radioimmunoassay kits compared. Clin Chem 1982; 28: 2429-2433.
2. Wood WG. A comparison of eleven commercial kits for the determination of serum ferritin levels. J Clin Chem Clin Biochem 1981; 19: 947-952.
3. Lipschitz DA, Skikne BA, Thomson P. Precision and accuracy of serum ferritin measurements. Am J Clin Nutr 1981; 34: 951-956.
4. Grail A, Hancock BW, Harrison PM. Serum ferritin in normal individuals and in patients with malignancy. J Clin Pathol 1982; 35: 1204-1212.

5. Dunn CDR, Boden DJ. Three commercial immunoradiometric kit assays for serum ferritin evaluated. Clin Chem 1981; 27: 1280-1283.
6. McKeering LV, Halliday JW, Caffin TA, Mach U, Powell LM. Immunological detection of isoferritins in normal human tissue and serum. Clin Chim Acta 1976; 167: 189-197.
7. Ghielmi S, Pizzoccolo G, Iacobello C, Albertini A, Arosio P. Methodological effects on the quantitation of serumferritin by radio and enzym immunoassays. Clin Chim Acta 1982; 120: 285-294.
8. Mareschal JC, Dublet B, Wustefeld C, Charler C, Crichton RR. Immunological characteristics of human isoferritins; consequences for human serumferritin determination. Clin Chim Acta 1981; 111: 99-103.
9. International Committee for Standardisation in Hematology, Expert panel on iron. Proposed international standard of human ferritin for the serumferritin assay. Br J Haematol 1985; 61; 61-63.
10. Iacobello C, Ghielmi S, Belloli S, Arosio P, Albertini A. Use of a reference standard to improve the accuracy and precision of seven kits for the determination of ferritin in serum. Clin Chem 1984; 30: 298-301.
11. Revenant MC, Goudable J, Beaudonnet A, Mailliavin A, Pichot J, Monnet L. Determination of serumferritin by a one-step immunoenzymassay, and comparison of four liver-ferritin standards. Clin Chem 1985; 31: 640-642.
12. Van Suijlen JDE, Van Noord PCH, Leijnse B. Accuracy of serum ferritin determinations in tissue preparations and human serum. J Clin Chem Clin Biochem 1990; 28: 43-48.
13. May M, Fish WW. The isolation and properties of porcine ferritin and apoferritin. Arch Biochem Biophys 1977; 182: 396-403.
14. Gonyea LM. Purification and iodination of antibody for use in an immunoradiometric assay for serum ferritin. Clin Chem 1977; 23: 234-236.
15. Worwood M. Serum ferritin. In Methods in Haematology 1980; Vol 1: 59-89, J Cook ed. Churchill Livingstone, New York.
16. Passing H, Bablok W. A new biometrical procedure for testing the equality of measurements from two different analytical methods. J Clin Chem Clin Biochem 1983; 21: 709-720.
17. Arosio P, Iacobello C, Montesoro E, Albertini A. Serumferritin evaluation with radioimmunoassays specific for HeLa and liver ferritin types. Immunol Lett 1981; 3: 309-313.
18. Dunn CDR. Napier JAF. Technical comments on the bioassay of erythropoietin. Exp Hematol 1978; 6: 577-584.
19. Arosio P, Adelman TG, Drysdale JW. On ferritin heterogeneity. Further evidence for heteropolymers. J Biol Chem 1978; 253: 4451-4458.
20. Wagstaff M, Worwood M, Jacobs A. Iron and isoferritins in iron overload. Clin Sci 1982; 62: 529-540.
21. Cazzola M, Dezza l, Bergamaschi B, Barosi G, Bellotti G, Caldera D, Ciriello MM, Quaglini S, Arosio P, Ascari E. Biologic and clinical significance of red cell ferritin. Blood 1983; 62, 1078-1087.
22. Konijn AM, Tal R, Levy R, Matzner Y. Isolation and fractionation of ferritin from human term placenta. A source for human isoferritins. Anal Biochem 1985; 144: 423-428.

23. Cazzola M, Arosio P, Bellotto V, Bergamaschi G, Dezza L, Iacobello C, Ruggeri G, Zappone E, Albertini A, Ascari E. Immunological reactivity of serum ferritin in patients with malignancy. Tumori 1985; 71: 547-554.
24. Tietz NW. A model for comprehensive measurement system in clinical chemistry. Clin Chem 1979; 25: 833-839.
25. Worwood M, Thorpe SJ, Heath A, Flowers CH, Cook JD. Stable lyophilized reagents for the serum ferritin assay. Clin Lab Haematol 1991; 13: 297-305.
26. Cozzi A, Levi S, Bazzigaluppi E, Ruggeri G, Arosio P. Development of an immunoassy for all human isoferritins, and its application to serum ferritin evaluation. Clin Chim Acta 1989; 184: 197-206.
27. Roberts DD, Drysdale JW. Immunoreactivities of human isoferritins. Clin Chim Acta 1985; 150: 41-51.
28. Gowan LK, Collawn JF, Fish WW. Seperation by high-performance liquid chromatography of two types of subunit from horse spleen ferritin. Journal of Chromatography 1983; 264: 463-468.

CHAPTER 3

A CANDIDATE REFERENCE METHOD FOR COUPLED SODIUM-WATER DETERMINATION IN HUMAN SERUM

This chapter is based on:

J Clin Chem Clin Biochem 1990; 28: 817-824.

Summary

A candidate reference method is described for coupled sodium-water determination based on ion-exchange sodium separation from the serum matrix followed by gravimetry as Na_2SO_4 and serum water determination by means of microwave evaporation. For sera with normal sodium and water contents, the mean relative standard deviation is 0.6 % (0.8 mmol/l). Mean inaccuracy for the coupled sodium-water determination is -0.3 % (0.4 mmol/l). The candidate reference method can be considered a reference method because the reference method value did not differ significantly from the definitive value, there is no known source for interferences or bias, and misinterpretation due to abnormal protein or lipid levels is excluded because serum sodium is determined on a plasma water basis.

Results of proficiency testing show a good correlation between (in)direct field methods and the candidate reference method, if sera with approximately normal water contents are used.

Introduction

In clinical chemistry, the sodium concentration in serum is determined by flame atomic emission spectrometry (FAES) or (in)direct ion-selective electrode (ISE) analyzers. In direct ISE methods the activity of free sodium ions in serum water is determined, whereas in traditional FAES methodology and indirect ISE measurement the total sodium concentration in serum after dilution of the sample is measured [1].

In all but direct ISE measurement, assay values are sensitive to the serum volume occupied by proteins and lipids. This sensitivity may lead to misinterpretation of electrolyte status (pseudohyponatremia) [2][3] if serum samples with abnormal protein and/or lipid concentrations are determined. Hence it is stated that the interpretation of sample data should be based on measurement of sodium activity (direct ISE) which is insensitive to the serum volume displaced by dispersed or dissolved solids [1][4][5]. Experience with direct ISE analyzers has shown that the results obtained on one particular sample may differ significantly from one instrument to another [6][7]. These differences, caused by broad variation in instrument design and calibration solutions, make comparison of data very difficult so that direct ISE is not suitable as an accuracy base (reference method) for sodium measurement. Since confusion also exists concerning the quantity to be reported by direct ISEs (activity or concentration), the European Working Group on ISEs (EWGISE) of the IFCC Expert Panel on pH, Blood Gases and Electrolytes, has proposed that all instruments should report results of sodium measurement as concentrations (mmol/l plasma) that agree with those obtained by FAES measurement in normal plasma samples [8].

To date, instrument manufacturers produce direct ISE analyzers which mathematically convert sodium activity in sodium concentration with the same numerical value as would have been obtained by FAES measurement, if normal serum is concerned [4].

Flame photometers (even if internal standard instruments are used) are subject to inherent bias from e.g. flame temperature, viscosity, surface tension and matrix effects. In order to exclude possible misinterpretation, data should also be expressed on a plasma-water basis [9].

We describe the development of a candidate reference method for coupled sodium-water determination based on ion-exchange sodium separation from the serum matrix followed by gravimetry as Na_2SO_4 and serum water determination by means of microwave evaporation. The procedure is based on a modification of the National Institute for Standards and Technology (NIST) definitive method for the determination of sodium in serum [10] and an improved version of a new water determination developed in our laboratory [11].
Finally, we used this method to study the accuracy of consensus values in proficiency testing.

Materials and Methods

Water
Twice distilled water with a specific resistance >10kOhm.m at 25°C was used.

Chemicals
Sodium, potassium, lithium standard solutions: standard solutions were prepared from Standard Reference Material (SRM) Sodium Chloride (SRM 919), Potassium Chloride (SRM 918) and Lithium Carbonate (SRM 924), NIST, Washington, D.C., USA. The SRMs were dried by heating at 110°C for four hours and stored in a desiccator for 1 hour before use.
Hydrochloric acid 8.2 mol/l, "suprapur", max. 0.5 ppm sodium, sulphuric acid 9.8 mol/l, "suprapur", max. 0.1 ppm sodium, nitric acid 10.3 mol/l, "suprapur", max. 0.1 ppm sodium, and boric acid "suprapur" were obtained from Merck, Darmstadt, Germany.
All other reagents were of analytical grade.

Accuracy assessment
SRM 909, human serum, was obtained from NIST, Washington, D.C., USA.
Lyophilised horse serum for clinical chemical use (RIVM high component concentrations), was obtained from RIVM, Bilthoven, the Netherlands.

Control sera from proficiency testing
Obtained from SKZL, Nijmegen, the Netherlands.
no. 90 A/B; 14-12-87: Nycomed Autonorm 2244/2280,
no. 91 A/B; 01-02-88: Biotrol,
no. 92 A/B; 05-04-88: Boehringer Mannheim Precinorm/path,
no. 93 A/B; 06-06-88: Seronorm 179/Pathonorm H-22,
no. 94 A/B; 15-09-88: Autonorm 2362/2361,
no. 95 A/B; 10-10-88: Technicon Testpoint 1/2,
no. 96 A/B; 05-12-88: Fisher Serachem,
no. 97 A/B; 30-01-89: Seronorm (human)/Lyotrol N-X,
no. 98 A/B; 03-04-89: Precinorm/Precipath and
no. 99 A/B; 29-05-89: RIVM (human).

Ion-exchange resin and chromatographic columns
AG 50W-X8, 100-200 mesh, a strongly acidic cation exchange resin and econo-columns, 1.0 x 30.0 cm, were obtained from Bio-Rad Laboratories, Richmond/CA, USA.

Microwave oven
Philips AVM Cooktronic 706/PH, 1.3 Kw, 2450 MHz, Philips, the Netherlands.

Containers
All volumetric glassware was of borosilicate material, confirming to class A specifications. Teflon beakers used were from TPX Azlon, Labplex, U.K. (50 and 250 ml beakers were used).

Balance
Sartorius 2462 analytical balance (d=0.1 mg), Göttingen, Germany.

Cleaning of the containers
All containers were soaked for 60 min in 0.75 mol/l nitric acid and afterwards rinsed 5 times with dest. The containers were dried at room temperature and used immediately after drying.

Sampling

To minimize sampling errors, standard and serum samples were taken by weighing. Aliquots were converted from weight in volume, using the density of each standard or serum sample, measured with a pycnometer (5.00 ml) at 25°C.

Weighing procedure

All standards and samples were weighed until constant weight: maximal allowed weight difference between two subsequent weighings 0.0005 g. Usually two subsequent weighings were sufficient to achieve constant weight.

Sodium determination

Ion-exchange separation

- Approximately 8 ml of sample is weighed into a teflon beaker (50 ml) and diluted with dest to about 25 ml.
- The sample is loaded onto the column using a peristaltic pump and teflon tubing. For complete transfer, the teflon beaker is washed with 25 ml dest, the washing solution is completely transferred onto the column. This washing procedure is repeated twice, so 100 ml dest is used for complete transfer and to wash non-ionogenic matter from each column.
- Elution is started with 0.4 mol/l HCl, flow 60 ml/h. Fraction collection is started after detection of the first acidic effluent.
- Fractions 0-75 ml (fraction A) and 75-225 ml (fraction B) are collected in 250 ml teflon beakers.
- The resin is regenerated with 100 ml 5 mol/l HCl and washed with 75 ml dest.
- Fractions A and B are evaporated on a hot plate to about 10 ml (8-12 hours at 100°C).
- Fraction A is examined by FAES to determine the presence of sodium. If present, the determination is rejected.

Gravimetric determination

- A platinum crucible (with cover) is heated for 15 min at 700°C and weighed after 15 min cooling in a desiccator. This procedure is repeated until constant weight is achieved.

- The concentrated fraction B is transferred quantitatively to the previously weighed crucible.
- 1 ml sulphuric acid (1 mol/l) is added (via the crucible wall) and the sample is evaporated on a hot plate. After cooling for 15 min in a desiccator, this procedure is repeated twice, the last time the crucible has to be partly covered with the platinum lid).
- The crucible (covered with lid) is slowly heated from 400-700 °C, and held at 700°C for 15 min.
- After cooling for 15 min in a desiccator, 1 ml ammonium carbonate (15.8%) is added (via crucible wall) and slowly evaporated to dryness on a hot plate.
- The sample is again heated for 15 min at 700°C, cooled 15 min in a desiccator and weighed.
- The last 2 steps are repeated until constant weight is achieved.

Calculation:

$$\text{Na (mmol/l)} = \frac{\text{yield } Na_2SO_4 \text{ (mg) x 2 x density (g/l)}}{\text{Mw } Na_2SO_4 \text{ (g/mol) x sample weight (g)}}$$

Remarks

All samples are analyzed in duplicate. The maximal allowed difference between duplicate measurements is 0.9 mmol/l. If this limit is exceeded, an additional measurement must be performed.

In sera with extremely elevated lithium levels (> 1.5 mmol/l) separation of lithium and sodium may not be complete (see results). For this reason the precipitate is dissolved in 10.00 ml dest, analyzed for lithium with the NIST reference method [12], and corrected if necessary. So far, this has never been necessary in our determinations.

Before use, the ion-exchange column has to be regenerated until constant column height (28 ± 0.5 cm) is achieved. This step is performed twice.

Normally two working days are needed for the determination of one reference method value, four columns can be handled efficiently by one technician.

Water determination

- All samples are analyzed in duplicate, the maximal allowable difference between duplicate measurements being 0.0010 kg/l. If this limit is exceeded, an additional measurement must be performed.
- A glass petri dish (with cover) is irradiated for 3 min, (2 min at 175 W, 1 min at 575 W), cooled 30 min in a desiccator (containing calcium sulphate as desiccant) and weighed.
- This procedure is repeated until constant weight is achieved (M1).
- Approximately 1 ml of sample is weighed into the glass petri dish and immediately hereafter it is weighed again (M2).
- 0.5 ml ethanol is added and after gently swirling of the petri dish, the dish is placed (uncovered) into the microwave.
- The petri dish is again irradiated for 3 min, cooled 30 min in a desiccator and weighed.
- This procedure is repeated until constant weight is achieved (M3).

Calculation:

$$[H_2O]\ (kg/l) = \frac{M2\ (g) - M3\ (g)}{M2\ (g) - M1\ (g)} \times \text{density (kg/l)}.$$

Accuracy water determination

- 4 vials of control serum are stored overnight in a desiccator containing $CaSO_4$ as a desiccant.
- The vials are weighed (W1), 5.00 ml of dest is added and the vials are weighed again (W2).
- After determination of density and water content (four determinations per vial) the vials are washed with distilled water, dried, stored overnight in a desiccator and weighed again (W3).

Calculation water content:

$$[H_2O]\ (kg/l) = \frac{W2\ (g) - W1\ (g)}{W2\ (g) - W3\ (g)} \times \text{density (kg/l)}.$$

Interference study

The possible interference of lithium and potassium was studied by determination of the sodium content in standard solutions containing different amounts of lithium (0.5-2.5 mmol/l) and potassium (5-50 mmol/l) but the same (weighed in) sodium concentration. In combination with these measurements, the influence of sodium-bicarbonate complex formation and sodium-protein binding on the assay results was investigated. Three solutions were prepared:

Solution A: 0.9368 g sodium chloride in dest,

Solution B: 0.9380 g sodium chloride in 0.05 mol/l $LiOH/H_3BO_3$,

Solution C: 1.3506 g sodium bicarbonate in 0.05 mol/l $LiOH/H_3BO_3$.

The sodium concentration in A was determined according to the protocol after incubation (2 hours at 20 °C) with 75 g/l human serum albumin (HSA). Reference method values are corrected for sodium present in the albumin fraction. The sodium concentration in B and C was measured by FAES [8] after separation on the ion-exchange column using 1 mol/l LiCl in LiOH boric acid buffer (pH=8.2). All samples are measured in duplicate on two consecutive days.

Expression of the sodium concentration in serum

The sodium concentration measured by the candidate reference method is divided by the measured water content and expressed as molality or substance concentration in the water phase (*total molal concentration; mmol/kg serum water)*.If this value is multiplied by the normal serum water value (kg/l) [13], the reported value is thus the *total molar concentration (mmol/l serum)*, corrected for abnormal water content.

Results

Precision

The imprecision of the sodium determination was calculated after twelve measurements on each of two human serum pools. The mean relative standard deviation of the candidate reference method value was 0.52 % at the 120-140 mmol/l level (table 3.1). For calculation of the precision of the water

determination, twenty measurements on two human serum pools were used. The mean relative standard deviation was 0.09 % (table 3.2). These results indicate a mean imprecision of the coupled sodium water determination of approximately 0.6 % (0.8 mmol/l) for sera having normal sodium and water contents.

table 3.1 *Precision of the sodium determination*

Control serum	Mean candidate reference method value mmol/l	SD mmol/l	CV %
Human serum pool A	141.0	0.66	0.47
Human serum pool B	121.1	0.68	0.56

table 3.2 *Precision of the water determination*

Control serum	Mean candidate reference method value kg/l	SD kg/l	CV %
Human serum pool C	0.9401	0.0007	0.08
Human serum pool D	0.9288	0.0009	0.10

Accuracy

The accuracy of the sodium determination was investigated with twelve measurements on a standard solution and four measurements on NIST SRM 909. The mean candidate reference method value for the standard solutions is 147.4 mmol/l (range 146.8-148.7 mmol/l), which equals the weighed-in value. The mean candidate reference method value for SRM 909 differed by -0.3 % from the definitive value, all four values were within the confidence limits of the definitive value given by NIST (table 3.3).

table 3.3 *Accuracy of the sodium determination*

Control sample	Definitive sodium value (95% confidence limit) mmol/l	Candidate reference method value mmol/l	Difference %
Standard	147.4	147.4	0.00
SRM 909 A	134.2 (133.4 - 135.1)	133.5	-0.52
SRM 909 B	134.2 (133.4 - 135.1)	134.1	-0.08
SRM 909 C	134.2 (133.4 - 135.1)	133.8	-0.29
SRM 909 D	134.2 (133.4 - 135.1)	134.0	-0.15

The accuracy of the water determination was investigated according to the procedure mentioned in the methods section. The mean measured water content differed -0.01 % (range -0.03-0.04 %) from the calculated value (table 3.4). Mean inaccuracy for the coupled sodium water determination therefore is approximately -0.3 % (0.4 mmol/l) with SRM 909.

table 3.4 *Accuracy of the water determination*

Control serum	H_2O calculated kg/l	H_2O measured kg/l	Difference kg/l
Vial 1	0.9374	0.9371	-0.03
Vial 2	0.9366	0.9368	0.02
Vial 3	0.9357	0.9357	0.00
Vial 4	0.9452	0.9456	0.04

Interference study

Sodium measurement is not influenced by potassium concentrations up to 50 mmol/l and lithium concentrations up to 1.5 mmol/l. If the lithium concentration in a serum sample exceeds 1.5 mmol/l, separation between lithium and sodium is

not complete and reference method values are biased by lithium sulphate precipitation.
Binding of sodium to protein and bicarbonate at physiological pH has been reported in literature[14][15] and this would be expected to increase with increasing pH. Because there is no significant decrease in measured sodium value after incubation with HSA under reaction conditions (solution A) and at pH = 8.2 (solution B) (table 3.5), the candidate reference method can be considered to determine both bound and unbound sodium in human serum. At pH=8.2 approximately 5% of the sodium present is bound to bicarbonate. Our results (table 3.5, solution C) indicate, as expected, that all bicarbonate bound sodium is also determined with the candidate reference method, because no significant difference between measured and calculated sodium concentration can be detected.

table 3.5 *Influence of sodium protein binding and bicarbonate complex formation on candidate reference method values*

Solution	Na calculated mmol/l	Na measured mmol/l	Binding %
A	160.3	160.9	0.37
		159.8	-0.32
B	160.5	160.2	-0.19
		161.0	0.31
C	160.8	161.0	0.12
		160.4	-0.25

Proficiency testing
Using the data base of the SKZL we studied the accuracy of the group mean of sixteen external quality control sera in three selected methods. These consensus values for FAES, direct ISE and indirect ISE were compared with the candidate reference method values (corrected for abnormal water content). Consensus values agreed well with the candidate reference method values (table 3.6).

table 3.6 *Comparison of candidate reference method sodium molarity with method dependent consensus values*

Control serum	H_2O kg/l	Candidate reference method value A mmol/l	FAES value B mmol/l	(B-A)/A x 100 %	ISE direct value C mmol/l	(C-A)/A x 100 %	ISE indirect value D mmol/l	(D-A)/A x 100 %
90 A	0.9333	133.7	134.1	0.3	133.8	0.1	134.4	0.5
90 B	0.9320	148.6	148.4	-0.1	150.3	1.1	149.3	0.5
91 A	0.9362	144.0	145.2	0.8	144.0	0.0	145.1	0.8
91 B	0.9221	155.3	154.8	-0.3	155.1	-0.1	154.9	-0.3
93 A	0.9447	134.8	135.4	0.4	145.7	0.7	135.2	0.3
93 B	0.9268	161.7	160.8	-0.6	162.5	0.5	160.6	-0.7
94 A	0.9384	126.5	128.4	1.5	125.0	-1.2	128.1	1.3
94 B	0.9283	145.5	145.1	-0.3	146.4	0.6	145.1	-0.3
95 A	0.9585	115.6	116.3	0.6	116.9	1.1	117.0	1.2
95 B	0.9350	147.5	146.3	-0.8	148.3	0.5	147.1	-0.3
96 A	0.9591	118.3	121.6	2.8	118.0	-0.3	120.9	2.2
96 B	0.9322	147.2	147.5	0.2	146.1	-0.7	147.5	0.2
98 A	0.9117	128.8	123.9	-3.8	129.7	0.7	124.6	-3.3
98 B	0.9173	141.6	136.8	-3.4	143.6	1.4	138.0	-2.5
99 A	0.9445	116.5	114.7	-1.5	118.1	1.4	115.2	-1.1
99 B	0.9475	144.1	143.8	-0.2	144.8	0.5	144.9	0.6

The mean bias of FAES was -0.3 % (range -3.8 -2.8 %) and of indirect ISE -0.1 % (range -3.3 -2.2 %). At low water content (samples 98A and B) negative bias and at high water content (sample 96A) a positive bias was predominant. Consensus values of direct ISE differed on average by -0.4 % (range -1.2 - 1.4 %) from the candidate reference method value.

Discussion

Good agreement between direct (ISE) and indirect (ISE and FAES) determinations is found in samples in which protein and lipid, and therefore, serum water concentrations are normal. The expected differences are found with samples which contain abnormally low or high levels of protein and lipids (and therefore an abnormal water concentration). In these samples, indirect methods give results which are susceptible to clinical misinterpretation.

Approaches which seek to correct flame photometric and indirect ISE sodium determinations by measurement of total protein and/or lipids in serum [9] or by measuring serum water [11] gain ready acceptance in clinical chemistry laboratories [16]. These methods vary in performance and generally exhibit a bias versus each other, thus indicate the need for all sodium measurement (both direct and indirect) to be performed on a serum-water concentration basis. Our method provides the necessary accuracy base, i.e. a candidate reference method for coupled sodium-water determination. This candidate reference method can be considered a reference method for the following reasons.

- The candidate reference method sodium value did not differ significantly from the definitive value given by NIST and mean imprecision and inaccuracy for the coupled sodium water determination are below the limits set by the Expert Commission on Electrolytes (1 SD imprecision limit of 1.5 mmol/l and a maximum difference from the definitive value of 2.0 mmol/l for serum sodium at the 140 mmol/l level) [10].
- There is no known source for interferences or systematic errors (bias), mainly because the sodium ion is separated from the matrix,
- It produces results which approximate the "true value" (total sodium value) within narrow limits because all sodium in serum (free ions as well as bound to protein or bicarbonate) is determined.
- Misinterpretation due to abnormal protein or lipid levels is excluded because reported values are expressed as sodium molality or as sodium molarity after correction for abnormal water content.

The candidate reference method is an accuracy base for sodium measurements and can be used in the calibration procedure of field methods and to study the accuracy of consensus values in proficiency testing.
It is argued that it would be physiologically desirable to measure sodium activity instead of molality [1,4,5]. Although there is an analytical and statistical difference between activity and molality, clinically these differences are not significant [4]. The major advantage of sodium molality measurement is that it provides for an accuracy base (the described candidate reference method) which is

absent in (direct ISE) sodium activity measurement. A disadvantage of presentation of sample data as sodium molality is that high capacity instruments are needed to determine the water content if indirect field methods are used. Because these instruments are not available and reference ranges would be significantly different from those currently used, the most practical solution is to correct the candidate reference method sodium concentration for a possible abnormal water content.

The sodium concentration measured by the candidate reference method is divided by the measured water content and multiplied by the normal serum water value (kg/l). The reported value is thus the total sodium molarity that would be measured if serum water concentration was normal. Because this "water correction" is used, current reference ranges can still be used and misinterpretation due to abnormal protein or lipid levels is still impossible.

The candidate reference method sodium molarity can be used in a calibration procedure of (in)direct field methods. This multi-point calibration procedure based on method comparison data is described in chapter 4.

Calibrating (in)direct field methods in the way described ensures that field method values are similar to the candidate reference method values, if samples with normal protein and lipid levels are used. Because inter-method comparability is increased and more accurate assay results are produced, this method comparison procedure should be used in all field methods in addition to the calibration procedure mentioned by the manufacturers.

The candidate reference method is also used to establish the accuracy of method dependent consensus values in a Dutch external quality control program. The expected differences are found in samples which contain abnormally low (96A) or high (98 A and B) protein and or lipid levels.

According to the new Dutch guidelines based on within-subject biological variation and the state-of-art [17], sodium determinations may differ maximal 3 % from the reference method value (at a sodium level of 140 mmol/l). Although mean differences are less than 0.5 % (table 3.6) FAES (98 A and B) and indirect ISE (98 A) have to be considered invalid when water content is abnormally low.

The candidate reference method might be used to determine the urgently needed reference method values of sodium in serum on a plasma water basis. If the regression equation of comparison against the candidate reference method is used for calibration of field methods, all samples which contain abnormally high levels of protein and or lipids will still, if indirect measured, give results which are susceptible to clinical misinterpretation. Only after correction of results from indirect measurement for abnormal water content, both indirect and direct measurement can be used for correct patient management. Therefore development of high capacity instruments for water determination in human serum is desirable. Meanwhile it must be kept in mind that for most patient samples indirect and direct measurement correlate well, and only in a few cases the difference is clinically significant.

References

1. Maas AHJ, Siggaard-Andersen O, Weisberg HF, Zijlstra WG.Ion-selective electrodes for sodium and potassium: A new problem of what is measured and what should be reported. Clin Chem 1985; 31: 482-485.
2. Ladenson JH, Apple FS, Aguanno JJ, Koch DD. Sodium measurements in multiple myeloma: two techniques compared. Clin Chem 1982; 28: 2383-2386.
3. Albrink MJ, Hald PM, Man EB, Peters JP. The displacement of serum water by the lipids of hyperlipemic serum. A new method for the rapid determination of serum water. J Clin Invest 1955; 34: 1483-1488.
4. Worth HGJ. A comparison of the measurement of sodium and potassium by flame photometry and ion-selective electrode. Ann Clin Biochem 1985; 22: 343-350.
5. Külpmann WR. Determination of sodium with ion-selective electrodes: A new method or a new quantity? J Clin Chem Clin Biochem 1990; 28: 813-815.
6. Oesch U, Ammann D, Simon W. Ion-selective membrane electrodes for clinical use. Clin Chem 1986; 32: 1448-1459.
7. Burnett D, Ayers GJ, Rumjen SC, Woods, TF. Sodium measurements in the presence of paraproteins by four direct ISE methods and flame photometry compared. Ann Clin Biochem 1988; 25: 102-109.
8. Boink ABTJ, Buckley BM, Christiansen TF, Covington AK, Maas AHJ, Mueller-Plathe O, Sachs Ch, Siggaard-Andersen O. Ion-selective electrode measurements of sodium and potassium in undiluted serum, plasma or whole blood: recommendations for expression of results in clinical practice. In: Proceedings of the 4th meeting of the European Working Group on Ion Selective ellectrodes 1986; 237-238, Radiometer, Copenhagen.
9. Waugh WH. Utility of expressing serum sodium per unit of water in assessing hyponatremia. Metabolism 1969; 18: 706-712.

10. Velapoldi RA, Paule RC, Schaffer R, Mandel J, Moody JR. A reference method for the determination of sodium in serum. NBS 1978; NBS Special Publication 260-60. US Government Printing Office, Washington DC.
11. van Kreel BK, Wulkan RW, Leijnse BL. Serum water determination by means of microwave evaporation. J Clin Chem Clin Biochem 1988; 26: 579-581.
12. Velapoldi RA, Paule RC, Schaffer R, Mandel J, Machlan LA, Garner EL, Rains TC. A reference method for the determination of lithium in serum. NBS 1980; NBS Special Publication 260-69. US Government Printing Office, Washington DC.
13. Broughton PMG, Smith SCH, Buckley BM. Calibration of direct ion-selective electrodes for plasma Na to allow for the influence of protein concentration. Clin Chem 1985; 103: 806.
14. Carr CW. Studies on the binding of small ions in protein solutions with the use of membrane electrodes. VI. The binding of sodium and potassium ions in solutions of various proteins. Arch Biochem Biophys 1956; 62: 476-484.
15. Coleman RL, Young CC. Evidence for formation of bicarbonate complexes with Na and K under physiological conditions. Clin Chem 1981; 27: 1938-1939.
16. Faye S, Payne RB. Rapid measurement of serum water to assess pseudohyponatremia. Clin Chem 1986; 32: 983-986.
17. Baadenhuijsen H, Jansen RTP, Meinders AE, Thijssen JHH. Normen externe kwaliteitsbewaking: uitwerking voor combi-enquete algemene chemie. Tijdschr NVKC 1992; 17: 255-261.

CHAPTER 4

CALIBRATION OF MULTITEST ANALYZERS: EVALUATION OF THE ACCURACY OF ROUTINE METHODS

This chapter is based on:

Eur J Clin Chem Clin Biochem 1991; 29: 205-208.

Summary

Calibration of two Technicon Chem 1 multitest clinical chemistry analyzers according to the instructions of the manufacturer resulted in significant differences in assay results of the two analyzers in twelve of twenty-two methods. To remove these differences we used an additional procedure based on the NCCLS EP-9T protocol for method comparison and bias estimation. Four reference methods were used as comparison methods. Our results indicate that for three of these methods results on both analyzers are biased in slope and intercept, and corrections of these parameters are necessary to produce results comparable to the reference method values.

For the remaining eighteen assays, methods in routine use in our laboratory were used as comparison methods. In fifteen of eighteen assays corrections of slopes and (or) intercept were necessary to prevent changes in the cumulative average patient data produced by our laboratory. Adjustment of slopes and intercepts according to the multi-point calibration procedure presented here, resulted in significant differences between system specific values measured in our laboratory and those assigned by the manufacturer in ten of thirteen methods.

The use of the multi-point calibration procedure presented in this chapter (based on calibration according to the instructions of the manufacturer and NCCLS EP-9T) greatly improves the intra-laboratory comparability. If the procedure is based on reference method data, inter-laboratory comparability can be improved as well, and unbiased results are produced.

We conclude that the multi-point calibration procedure therefore should be used on a regular basis for all measurement systems in clinical chemistry to (re)assure accurate assay results.

Introduction

Many of the analytical methods routinely used in clinical chemistry have to be calibrated with commercially produced serum calibrators. Some clinical chemists change the target values assigned by the manufacturer to values which have been obtained by the method they use(d) routinely in the belief the use of these values will compensate for methodological inaccuracies [1]. However, the use of changed and unchanged calibration values may not compensate for non-specificity of the method concerned because differences may exist between the matrices of calibrator and patient samples. If the matrix of the calibration material is not identical with that of fresh human serum, the manufacturer may give this material a system-specific value (SSV) to correct for matrix interferences. If the SSV is used for calibration, the results obtained for patient samples are presumed to be accurate. However, calibrations with only one calibrator can still cause matrix shifts in routine methods with limited specificity, even if the SSV concept is used [2]. Therefore a multi-point calibration procedure has to be preferred.

This is demonstrated by the calibration procedures of two multitest clinical chemistry analyzers in routine use in our laboratory. Before the implementation of these Technicon Chem 1 analyzers in the laboratory, a screening protocol was performed after calibration according to the instructions of the manufacturer. Surprisingly, significant differences were found between results from the two Chem 1 analyzers for several assays. We decided therefore to study these differences in more detail by following the NCCLS EP-9T method comparison protocol [3]. As comparison methods we used the (proposed) reference methods for total protein [4], total bilirubin [5], total cholesterol [6] and sodium [7]. For all except the four methods mentioned no (proposed) reference method was available in our laboratory and the methods in routine use were used as comparison methods. Our goal in this situation was not to change the cumulative means of patient data the laboratory produces.

The method comparison data thus obtained (slope and intercept) are used for additional multi-point calibration of both Chem 1's.

Materials and Methods

Materials

The system calibrator used was Setpoint Chem 1 (Technicon), lot no. 8V1614, a bovine lyophilised serum product. Together with the calibrator a list of SSV's as determined by the manufacturer, is enclosed. For the initial screening and the calibration procedure fresh human sera were used.

Routine methods and instruments

The instruments in routine use in our laboratory and the methods used as comparative methods described in our previous article [2]. Briefly, the comparison instruments are SMAC and RA 1000 from Technicon, Tarrytown, USA, ACP 5040 and EPOS from Eppendorf, Hamburg, Germany.
The two Chem 1 analyzers are from Technicon, Tarrytown, USA and are named Chem 1A and Chem 1B in this chapter.

Calculation of results

The Chem 1 analyzer automatically calculates the reported result according to the following expression:

$$\text{Reported result} = \frac{(\text{assay result - intercept})}{\text{slope}}$$

Because Technicon recommends a slope of 1.0 and an intercept of 0.0 (except for creatinine), initially all reported results equal the measured assay results.

Screening protocol

Approximately 20 (range 19-24) fresh human sera per analyte, were selected after routine analysis. The range of values extended from below the reference range to substantially above it. We analyzed these samples simultaneously with Chem 1 and comparison routine methods. All analyzers were previously calibrated according to the instructions of the manufacturers, using calibration sera. A regression equation was calculated and the resulting slope and intercept are entered in the software of the corresponding Chem 1 if there was a significant difference (see statistical analysis) between the means of assay results.

NCCLS EP-9T protocol

The complete protocol is followed with exception of the linear regression procedure (see statistical analysis). In brief, the protocol is as follows:
All analyzers are calibrated according to the instructions of the manufacturers.
On 5 consecutive days, 8 patient samples and 3 control samples are analyzed in duplicate with the Chem 1 analyzers and the comparison method within 2 hours.
The first aliquots of the patient samples are analyzed in a randomized sequence, the second (duplicate) measurement in reverse order, the range of measurements must extend from values below the expected reference range to substantially above it.

Statistical analysis

Inter-assay comparisons were made by regression analysis according to *Passing* and *Bablok* [8]. Significant differences in slope and or intercept are considered at the 95% confidence level. Significant differences between means of series of assay results were investigated with the *Student's* paired t-test (significant difference if $p < 0.05$).

Multi-point calibration procedure

A Calibration according to the instructions of the manufacturer.
B Method comparison according to NCCLS EP-9T.
C Regression analysis according to *Passing* and *Bablok* [8].
D If a significant difference in slope and or intercept of the comparison data is demonstrated, the slope and or intercept in the system software are recalculated and entered.
E Repeat starting from B.

Setpoint Chem 1 system-specific calibration values

After the multi-point calibration procedure the SSV's of Setpoint Chem 1 were measured. On 10 consecutive days, Setpoint was determined twice (morning and afternoon) in duplicate. The mean values from duplicate measurements were taken for calculation. These values are compared with the SSV's assigned by Technicon.

Results

The regression equations resulting from the screening procedure are presented in table 4.1. In addition to these equations, the assay results from Chem 1A are compared with those from Chem 1B. Five analytes (chloride, alanine aminotransferase, aspartate aminotransferase, lactate dehydrogenase and creatine kinase) show significant differences in slopes and intercepts between Chem 1A and Chem 1B. The methods for alkaline phosphatase, α-amylase, urea and uric acid show significant differences in the slopes while the assays for triglycerides, total bilirubin and creatinine differ significantly in the intercepts where the comparison Chem 1A versus Chem 1B is concerned.

The final slope and intercept values resulting from the entire calibration procedure (and therefore entered in the software of the corresponding Chem 1) are shown in table 4.2. Only the slopes and intercepts for the calcium, inorganic phosphate, cholesterol and uric acid determinations are within the 95% confidence limits set for slope = 1.0 and intercept = 0.0. All other methods are significantly different in slope and (or) intercept from the comparison method.

The system specific values measured in our laboratory after this multi-point calibration are compared with the system specific values established by the manufacturer. The mean results are presented in table 4.3. Owing to the multi-point calibration procedure, there was of course no significant difference between analyzer values for patient samples. The mean measured values for all but three methods (chloride, inorganic phosphate and total protein) show a significant difference (*Student* t-test, $p < 0.001$) from the values assigned by Technicon.

Discussion

The use of a large number of serum specimens from patients for calibration of (new) analyzers in clinical chemistry most closely approaches the goal of calibration, i.e. direct comparability of the results of a (new) method with those obtained with the comparison method. This leads to the question which comparative value is correct, i.e. which comparison methods yield good estimates of the "true value". Reference methods should be used for this purpose because

table 4.1 *Method comparison by Passing and Bablok regression calculation after the screening procedure*

method	Chem 1A versus comparison method	Chem 1B versus comparison method	Chem 1A versus Chem 1B	Signifi-cances*)
Sodium	y=0.927x+10.6	y=0.951x+8.3	y=1.024x-2.3	-,-
Potassium	y=1.140x-1.13	y=1.134x-1.18	y=0.994x-0.05	-,-
Calcium	y=0.943x+0.069	y=0.948x+0.114	y=1.005x+0.044	-,-
Magnesium	y=0.774x+0.010	y=0.771x+0.009	y=0.997x-0.001	-,-
Chloride	y=1.176x-16.7	y=1.072x-6.8	y=0.897x+9.8	a,b
Inorganic phosphate	y=1.006x+0.002	y=0.989x+0.030	y=0.983x+0.028	-,-
Alanine aminotransferase	y=1.279x+4.4	y=1.329x-0.1	y=1.040x-4.7	a,b
Aspartate aminotransferase	y=1.441x+2.2	y=1.471x-2.5	y=1.020x-4.7	a,b
Lactate dehydrogenase	y=1.782x+6.8	y=1.946x-15.2	y=1.093x-22.8	a,b
γ-Glutamyltransferase	y=1.068x-4.7	y=1.062x-5.4	y=1.007x-0.6	-,-
Alkaline Phosphatase	y=1.305x-4.5	y=1.330x-4.9	y=1.019x-0.3	a,-
Creatine kinase	y=1.747x+1.5	y=1.773x+4.1	y=1.015x+2.6	a,b
α-Amylase	y=0.628x+2.6	y=0.581x+1.7	y=0.926x-0.7	a,-
Albumin	y=1.052x+1.9	y=1.026x+2.2	y=0.975x+0.4	-,-
Total protein	y=0.951x+2.6	y=0.953x+1.8	y=1.002x-0.7	-,-
Cholesterol	y=1.011x+0.78	y=1.007x+0.060	y=0.997x-0.17	-,-
Triglycerides	y=1.137x-0.087	y=1.123x-0.35	y=0.988x+0.051	-,b
Total bilirubin	y=1.254x-5.1	y=1.250x-7.75	y=0.996x-2.6	-,b
Glucose	y=1.034x+0.00	y=1.021x+0.00	y=0.987x-0.05	-,-
Creatinine	y=1.023x+27.8	y=1.026x+13.8	y=1.003x-14.1	-,b
Urea	y=0.935x+0.52	y=0.981x+0.94	y=0.952x+0.44	a,-
Uric acid	y=0.961x+0.032	y=1.032x-0.013	y=1.079x-0.049	a,-

*) a: significant difference from slope = 1.0,
b: significant difference from intercept = 0.0, in regression equations Chem 1A versus Chem 1B (95% confidence level, calculation according to *Passing and Bablok*).

table 4.2 *Correction of slope and intercept Chem 1 according to the multi-point calibration procedure*

method	Chem one	Slope	Intercept	Significances*)
Sodium	1A	0.891	14.9	c,d
	1B	0.884	15.9	c,d
Potassium	1A	1.016	0.79	-,d
	1B	1.009	- 0.82	-,d
Calcium	1A	0.980	0.030	-,-
	1B	1.033	- 0.070	-,-
Magnesium	1A	0.774	0.070	c,d
	1B	0.769	0.076	c,d
Chloride	1A	0.888	11.5	c,d
	1B	0.856	10.7	c,d
Inorganic phosphate	1A	1.012	- 0.020	-,-
	1B	0.976	- 0.013	-,-
Alanine aminotransferase	1A	1.342	2.7	c,d
	1B	1.396	- 0.3	c,-
Aspartate aminotransferase	1A	1.428	1.2	c,d
	1B	1.471	- 2.0	c,d
Lactate dehydrogenase	1A	1.806	1.6	c,-
	1B	1.841	- 2.4	c,-
γ-Glutamyltransferase	1A	1.629	- 3.9	c,d
	1B	1.660	- 3.9	c,d
Alkaline Phosphatase	1A	1.203	5.6	c,d
	1B	1.227	4.9	c,d
Creatine kinase	1A	1.654	0.9	c,-
	1B	1.699	1.0	c,-
α-Amylase	1A	0.593	0.7	c,-
	1B	0.581	6.2	c,d
Albumin	1A	1.000	2.5	-,d
	1B	0.950	2.9	c,d
Total protein	1A	0.941	2.1	c,-
	1B	0.939	1.9	c,-
Cholesterol	1A	1.000	0.00	-,-
	1B	1.000	0.00	-,-
Triglycerides	1A	1.153	- 0.058	c,d
	1B	1.127	- 0.019	c,-
Total bilirubin	1A	1.194	- 0.5	c,d
	1B	1.222	- 0.6	c,d
Glucose	1A	1.034	- 0.10	c,-
	1B	1.021	- 0.10	c,-
Creatinine	1A	1.036	15.7	c,d
	1B	1.021	15.9	-,d
Urea	1A	0.936	0.17	c,d
	1B	0.908	0.41	c,d
Uric acid	1A	0.999	- 0.003	-,-
	1B	0.969	0.002	-,-

*) c: significant difference from slope = 1.0,

d: significant difference from intercept = 0.0, in regression equations Chem 1A versus Chem 1B (95% confidence level, calculation according to *Passing and Bablok*).

these methods "have shown to have negligible inaccuracy in comparison with their imprecision" (definition of the IFCC). If these reference methods are absent, routine methods have to be used. In that case, calibration resulting in an adjustment of new data in accordance with cumulative patient means is the approach of choice.

The methods programmed in the Chem 1 are either factory calibrated (zero order rate methods) or require a one-point (endpoint and first order rate methods) or a two-point (sodium and potassium assays) calibration. For one-point procedures a system calibrator with SSV's for the analytes concerned is advised. These calibration procedures showed several disadvantages as will be discussed below.

Firstly, it was surprising and disappointing that results from two Chem 1 analyzers differed significantly from each other in twelve of twenty-two methods (table 4.1) after calibration according to the instructions of the manufacturer. To remove these differences the values for slope and intercept resulting from the screening procedure had to be entered in the system software. These values are corrected for day-to-day variability if the sample data from the method comparison according to NCCLS EP-9T are used.

Secondly, the accuracy of the assay results may not be guaranteed after a one-point calibration procedure. This certainly is the case for the total protein and total bilirubin assays in human serum. For these methods corrections on slope and (or) intercept are necessary to produce results comparable to the reference method values (table 4.2). Even if a two-point calibration procedure is used (total sodium) results on both Chem 1's can be biased in slope and intercept. In contrast, the determination of total cholesterol is unbiased and no further correction was needed (table 4.2).

Thirdly, additional corrections may always be necessary to assure good correlation with other routine methods and to prevent changes in the cumulative patient means after introduction of a new method or analyzer. In our laboratory only three methods (calcium, inorganic phosphate and uric acid) were introduced without significant corrections for slope or intercept.

Also for the potassium, magnesium, chloride, albumin, triglycerides, creatinine and urea assays corrections are necessary to adjust assay data in accordance to cumulative patient means. The accuracy of these Chem 1 methods is questionable,

table 4.3 *System specific calibration values; A: estimated by Technicon, B: mean measured values after multi-point calibration procedure*

Test		A	B	SD	(B-A)/A x 100 %	p value (t-test)
Calcium	mmol/l	2.54	2.46	0.02	-3	<0.001
Magnesium	mmol/l	1.00	1.17	0.01	17	<0.001
Chloride	mmol/l	115	115	2.2	0	>0.500
Inorganic phosphate	mmol/l	1.80	1.79	0.01	-1	0.057
Albumin	g/l	30	31	0.3	-9	<0.001
Total protein	g/l	64	64	0.8	0	>0.500
Cholesterol	mmol/l	4.42	4.24	0.07	-4	<0.001
Triglycerides	mmol/l	2.09	1.78	0.02	-15	<0.001
Total bilirubin	μmol/l	97	87	2.6	-10	<0.001
Glucose	mmol/l	14.7	13.9	0.09	-5	<0.001
Creatinine	μmol/l	778	702	13.0	-10	<0.001
Urea	mmol/l	21.8	23.1	0.14	6	<0.001
Uric acid	μmol/l	547	570	5.0	4	<0.001

because all routine methods used as comparison methods perform well in the quality control programs of the SKZL. The corrections of slopes and intercepts of all zero order methods (enzymes), however, are predominantly caused by differences in reaction temperature between Chem 1 and comparison methods.

Fourthly, an assay result may be biased even if the measured calibration value equals the assigned calibration value. The mean measured SSV's for both chloride and total protein determinations do not differ from the assigned SSV's (table 4.3) although both analytes are overestimated at low serum levels and underestimated at high serum levels as indicated by the regression equations in table 4.2.

It is evident from our data that the use of the multi-point calibration procedure presented in this chapter greatly improves the intra-laboratory comparability. To improve the inter-laboratory comparability and to produce unbiased results, reference methods must be used as comparison methods.

It is important to verify the calibration on a regular basis e.g. when lots of calibrators change, or when QC data reveal deteriorating performance. To (re)assure accurate assay results within 1 day, the data resulting from the screening protocol can be used for readjustment of slope and intercept.

Finally, we performed the study described above with two Chem 1 systems and we cannot judge directly on other analyzers. However, Chem 1 methods and calibration procedures are based on officially designated or well defined procedures that are also representative for other systems. Therefore, we do not believe that other systems would behave differently from these two Chem 1 analyzers.

We conclude that the multi-point calibration procedure should be used on a regular basis for all measurement systems in clinical chemistry to (re)assure accurate assay results.

References

1. Burnett D, Woods TF, Jennings RD. Acceptance of values assigned to serum-based calibration materials. Ann Clin Biochem 1886; 23: 157-167.
2. Blijenberg BG, Boerma GJM, Leijnse B. An evaluation of the improved Technicon Chem 1 system. J Clin Chem Clin Biochem 1989; 27: 457-469.
3. National Committee for Clinical Laboratory Standards. Method comparison and bias estimation using patient samples; tentative guideline. NCCLS 1993; vol. 13, no. 4.
4. Doumas BT, Bayse BB, Carter RJ, Peters T, Schaffer R. A candidate reference method for determination of total protein in serum; I development and validation. Clin Chem 1981; 27: 1642-1650.
5. Doumas BT, Kwok-Cheung PP, Perry BW, Jendrzejczak B, McComb B, Schaffer R, Hause L. Candidate reference method for determination of total bilirubin in serum: development and validation. Clin Chem 1985: 31, 1179-1189.
6. Duncan IW, Mather A, Cooper GR. The procedure for the proposed cholesterol reference method. Clinical Chemistry Division, Centers for Disease Control 1982, Atlanta, GA 30333, 75pp. Available on request.
7. Van Suijlen JDE, Berrevoets CA, Leijnse B. A candidate reference method for coupled sodium-water determination in human serum. J Clin Chem Clin Biochem 1990; 28: 817-824.
8. Passing H, Bablok W. A new biometrical procedure for testing the equality of measurements from two different analytical methods. J Clin Chem Clin Biochem 1983; 21: 709-720.

CHAPTER 5

TURNAROUND TIME AS A QUALITY ASPECT OF CLINICAL CHEMISTRY LABORATORY TESTING

Submitted for publication

Summary

In this chapter a fully automated turnaround time (TAT) monitoring system for all stat test requests is presented. Results for all grouped tests are presented on a daily basis to the laboratory staff. On a weekly basis all TAT related data are plotted and made available as well. All TAT data are saved in day, week and month files, which can be filtered and plotted automatically on demand. These procedures take less than 30 minutes per day and are used to establish pragmatic criteria for acceptability of TAT, to evaluate the impact of TAT monitoring and organizational structures on TAT and to examine the need for point-of-care testing.

Introduction

The product of the clinical chemistry laboratory, the test result, should be functional, precise, accurate and timely. Internal process monitoring and external proficiency testing programs support the laboratory in its efforts to obtain optimal analytical quality. Periodic monitoring of precision and accuracy are an essential part of quality assurance procedures since many years.
Timeliness of result reporting is not routinely included among standard procedures of laboratory quality monitoring, despite its importance and the availability of laboratory computer systems which make fully automated turnaround time (TAT) monitoring possible.
For a test to be useful, it's result must be available in a timely manner. Nevertheless, it is only recently that the evaluation of test TAT is introduced as a quality aspect of laboratory performance [1]. Ever since, TAT of laboratory performance has been the subject of systematic study. These studies quantitate TAT, factors influencing TAT, laboratorian and clinician definition of and goals for TAT and the conformance with these goals [2-6].
In this chapter, the application of a system of computer programs for fully automated continuous TAT monitoring of all stat test requests is presented. The system is used to establish pragmatic criteria for acceptability of TAT, to investigate the impact of TAT monitoring and laboratory organizational structures on TAT and to examine the need for point-of-care testing (POCT).

Materials and methods

Definition of turnaround time
In this chapter, TAT is related to three separate times. The time between phlebotomy and result reporting was defined as the *testing time*, whereas *laboratory time* was defined as the time between laboratory specimen accessioning and result reporting. The *pre-laboratory* time is the time between phlebotomy and laboratory specimen accessioning.

Data collection and analysis

The study was conducted in the department of Clinical Chemistry of the University Hospital Rotterdam (CCUHR; the Netherlands). This department exists of three laboratories at different work areas in the hospital. The laboratories use a laboratory information system (LIS), which is operational as a part of an integrated hospital information system, (LABZIS; BAZIS, Leiden, the Netherlands) for ordering, accessioning and transmittal of test results. In addition LIS groups the various tests on work lists for individual instruments and work areas (table 5.1).

table 5.1 *High-volume test groups, working areas and procedures in the Department of Clinical Chemistry.*

Test group	Test panel	Analyzer	Work area	Working procedures	Stat tests/ month
A	pH, pCO_2, pO_2	Radiometer ABL 3	CCCL adult hospital	*phlebotomy*: clinical staff	1557
B	Glucose	Merck EPOS		*transport*: pneumatic tube or clinical staff	2588
C	General chemistry	Technicon Chem 1			2137
D	General chemistry	DuPont Dimension		*monitoring*: laboratory time	1334
E	CBC, Hb, Ht, differentiation	Sysmex NE 8000			3937
F	pH, pCO_2, pO_2 and electrolytes	Radiometer ABL 505	Satellite lab thoracic surgery department	*phlebotomy*: clinical staff *transport*: pneumatic tube or clinical staff *monitoring*: testing time	1832
G	CBC, Hb, Ht differentiation	Sysmex K 1000			894
H	pH, pCO_2, pO_2 and electrolytes	IL BGE	CCCL children's hospital	*phlebotomy* and *transport*: laboratory personnel *monitoring*: testing time	2101
I	CBC, Hb, Ht differentiation	Technicon H1			1718
J	General chemistry	DuPont Dimension			938

CBC = Complete Blood Count, CCCL = Central Clinical Chemistry Laboratory

A LIS-satellite program was developed that both creates a TAT record as soon as a test result is reported to LIS en adds this record to a TAT file. The TAT record contains patient and sample identification, clinical unit, date, test group and profile, time of phlebotomy, time of request entering in LIS after arrival of the sample at the laboratory (i.e. laboratory specimen accessioning time) and time of reporting to LIS. For all stat testing, these data are gathered in the department of clinical chemistry, starting approximately 1 year ago. Accessioning and reporting time are captured in real time as all accessioning and reporting transactions are being made in LIS. Time of phlebotomy was obtained from the test request form (or from LIS) and was entered in LIS by laboratory technicians or clerks.
The raw data from the TAT record were downloaded to a PC on a daily basis (or more frequently if necessary) and further extracted and edited into ASCII-format using Turbo Pascal 7.0 (Borland International Inc) All raw data are saved in day files. Additionally, cumulative week and month files are prepared by test group, and all files can be further filtered by using test groups (day files), clinical unit, period of day and maximum TAT as selection criteria. Finally, the homogeneous data are downloaded to Quattro Pro 4.0 (Borland International Inc) and graphics of the data are created automatically by a program developed in Quattro Pro. The entire procedure requires only 30 minutes handling including evaluation of the results.

Data display and evaluation
On a daily basis, scattergrams of all high-volume grouped stat tests TAT's are made available to the laboratory staff. These scattergrams display all individual TAT's versus the time of sample collection over the preceding 24 hours (figure 5.1). Testing time is considered when blood specimens are collected and brought to the laboratory by laboratory technicians or phlebotomists. If blood specimens are collected by non-laboratory employees, laboratory time is considered (table 5.1). The reason for this discrepancy is twofold. Firstly, it proved to be impossible to gather reliable phlebotomy times when specimens were collected by non-laboratorians. Secondly, the goal was to monitor that part of the testing process that was carried out under direct responsibility of the laboratory. An exception to this rule is stat testing in the satellite lab in the thoracic surgery

department, where testing time was taken into account although samples are collected by non-laboratorians. Reliability is guaranteed by the fact that sampling time is used as sample identification in the operating room.

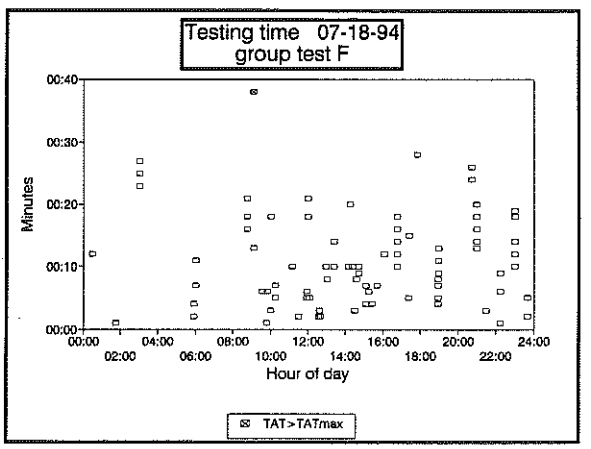


figure 5.1 example of a daily TAT plot as presented to laboratory staff

To each scattergram, a list of samples exceeding the maximum allowable TAT (TATmax) is added and the cause of the time delay is investigated according to the criteria presented in table 5.2. On a weekly basis, all TAT related data are plotted by test group and are made available to the laboratory staff (figure 5.2).

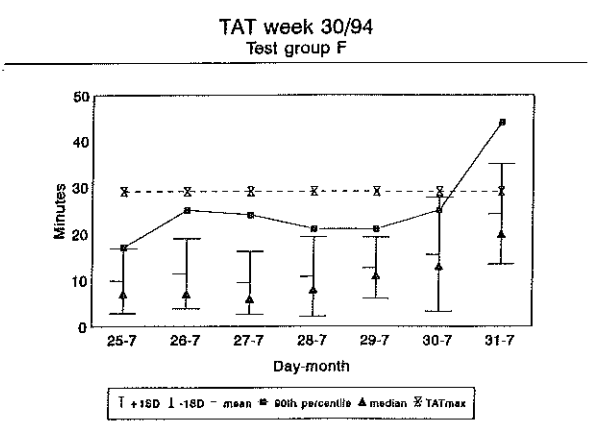


figure 5.2 example of a weekly TAT plot as presented to laboratory staff

table 5.2 *Criteria for evaluation of testing procedures exceeding TATmax*

TATmax = 90th percentile TAT over the preceding 3 months	
Daily TAT test results	*Action*
10% or less results > TATmax and no result(s) > 3x TATmax	None
10% or less results > TATmax but 1 or more result(s) > 3x TATmax	Look for cause TAT > 3x TATmax for all individual samples
more than 10% results > TATmax and no result(s) > 3x TATmax	Look for cause general procedural problems
more than 10% results > TATmax and 1 or more result(s) > 3x TATmax	Look for cause TAT > 3x TATmax for all individual samples and for general procedural problems

Statistical analysis

The sample numbers, mean, median and 90th percentile are calculated for all data sets prepared by using the computer programs presented in the data collection and analysis part.

TAT does not follow a normal distribution and therefore, significant differences between mean TAT results were investigated with the *Wilcoxon* signed rank test (abnormal distribution, significant difference if $p < 0.05$).

The maximum allowable TAT (TATmax) is defined as the 90th percentile of test group TAT over the last three months. If the 90th percentile TAT is decreased significantly in this period, TATmax is altered. If there happens to be no significant decrease in 90th percentile TAT, the previous TATmax is maintained.

Differences between numbers of tests exceeding TATmax were investigated with the *Student's* unpaired t-test (significant difference if $p < 0.05$).

Results and discussion

The TAT of stat laboratory tests is the most visible aspect of laboratory performance to clinicians. Perhaps therefore more than 80% of laboratories receive complaints about TATs [7]. There are two important reasons why a large number of test results are not available within a time period required by clinicians. Firstly, clinical chemists and clinicians have neglected to jointly develop TAT definitions and goals for acceptable TAT. Clinical chemists often use 60 minutes laboratory time as a time limit for acceptable stat TAT [8], and organize their laboratories to fulfil this criterion. However, a recent study [6] indicated that clinician median TAT goals for testing time are 10 min for pO_2, 20 min for hemoglobin and 30 minutes for glucose and potassium measurements. These tests are ordered frequently as part of a stat test panel and are therefore representative for other stat tests performed simultaneously.

Secondly, there may be discrepancies between expected and actual TAT because there are unknown sources for result delay which remain unknown because TAT is not monitored periodically. Recently, the principles of Continuous Quality Improvement (CQI) have been used to evaluate the impact of general testing procedures on TAT [9][10]. After procedural changes to remove the cause of unnecessary delay, TAT was reduced significantly. Of course removal of unnecessary delay caused by general procedures is the first step towards optimizing TAT. The second thing to be done, is the implementation of a system for continuous TAT monitoring for all individual stat testing to be able to track down every accidental TAT delay immediately.

Recently, two studies on periodic monitoring of TAT and the impact on laboratory performance have been published [11][12]. Rollo et al [11] have chosen for the production of weekly scattergrams showing hourly TAT patterns. Although a valuable additional quality assurance tool, this program contains the impossibility to identify all causes for delay in test reporting. To be able to do so, all TAT results have to be displayed on a daily basis and results have to be acted upon if preset TAT limits are exceeded. The freckle plot approach [12] fulfils these criteria, but is too laborious to allow monitoring of all stat testing in a department of clinical chemistry.

The use of TAT as a quality aspect in clinical chemistry starts with continuous gathering of all stat test TATs in the laboratory. This is the first indication of the impact of organizational structure on actual TAT by an objective method and allows to quantitate unnecessary TAT delay. Mean, median and 90th percentile TAT of all stat laboratory testing in the CCUHR in the preceding year are shown in table 5.3.

table 5.3 *Statistics of all stat grouped tests in the one year evaluation period*

	Testing time (min)			*Laboratory time* (min)		
Test group	Mean	Median	90th percentile	Mean	Median	90th percentile
A	42.8	28.3	95.1	7.0	3.5	17.5
B	102.3	81.1	179.7	19.9	14.6	41.6
C	166.7	164.8	226.7	74.7	73.9	102.2
D	68.0	48.2	122.8	25.4	20.9	49.1
E	56.3	39.2	120.6	9.0	5.7	18.8
F	15.5	10.0	29.6	4.0	2.5	9.7
G	19.3	12.5	36.9	6.4	3.4	14.5
H	27.6	22.0	47.4	7.0	4.7	13.3
I	41.4	31.6	84.3	22.7	13.6	58.9
J	51.9	46.3	79.5	31.3	27.2	52.2

The second thing to be done is establishment of more pragmatic criteria for acceptability of TAT and to act upon unacceptable results. Acceptable and pragmatic in this is a reflection of staffing, instrumentation and testing procedures on TAT limits for each test group. This is realized by defining TATmax as the 90th percentile of grouped test TAT results over a representative time period (e.g. three months) and by using the criteria for evaluation of testing procedures exceeding TATmax as stated in table 5.2. Because the TATmax is recalculated if the 90th percentile TAT is decreased significantly in the last three months, it indicates the actual *state-of-art* of laboratory TAT performance.

The department of CCUHR states to its customers that laboratory time for all stat testing should be less than 60 minutes. As can be seen from 90th percentile TATs in table 5.3, this limit is exceeded for test group 4. These tests however, being processed as stat tests are in fact requested with asap (as soon as possible) priority and for this samples a TAT limit of 3 hours is used.
Due to procedural changes, triggered by repeating long lists of samples exceeding TATmax, TATs for test group F and I tests are decreased significantly in the evaluation period ($p < 0.001$). The 90th percentile laboratory TAT for test group I (CBC and diff) was 69 minutes in the first three months of the evaluation period and therefore beyond preset TAT limits. Research pointed out that unnecessary time delay was caused by working up manual WBC diff because the procedure was to wait for the automatic (H1) diff which failed in over 80% of samples. Immediate working up of manual diff (if requested) effected in a significant decrease (53 minutes 90th percentile TAT over the last three months) in group I TATs. Other test group TATs didn't change significantly, but for all test groups the number of tests exceeding TATmax was reduced significantly because of consistent follow-up of TAT delay.
Organizational structures have a tremendous influence on TAT. In figure 5.3 median TATs are presented for four different working areas in the department of CCUHR where pH, blood gas and electrolyte measurements are performed.

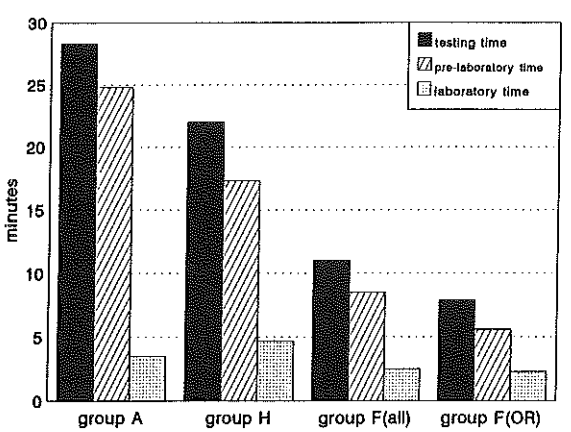


figure 5.3 Impact of laboratory organizational structure on TAT

Important among these are samples from the operating rooms (OR) analyzed in the satellite lab at the thoracic surgery department (24 % of all stat tests at the satellite lab).
Median laboratory time is less than 5 minutes for all areas, with the shortest TAT for stat samples from the OR in the satellite lab in the thoracic surgery department. The significant differences in TAT are caused by differences in pre-laboratory time, which account for up to 88% of total testing time when sampling and transport is beyond the direct responsibility of the clinical chemistry department (group A figure 5.3). The shortest overall testing time is achieved in the satellite lab, with a median TAT of 11.0 min. This is reduced to 7.9 min if only samples from the OR are regarded. The possible explanation for this difference is that samples from the OR are send to the lab by pneumatic tube, and all other samples are brought to the lab by clinical staff.
Further, *state-of-art* TAT has to be announced to all laboratory related clinicians, and has to be compared to clinician goals for TAT. Announcement of TAT performance and comparison with clinician expectations is important regardless whether possible discrepancies represent unreasonable clinician expectations, inadequate laboratory performance, or a combination of both. If a clinician expects the laboratory not to meet his or her TAT requirements or is not aware of laboratory TAT performance, regardless of the true test urgency, that person is more inclined to order tests stat, disrupting the specimen flow through the laboratory and lengthening average routine and stat TATs. This improper use of stat procedures is one of the major causes for exceeding TATmax (table 5.4), certainly if this leads to a higher accessioning overload frequency.

table 5.4 *Major causes for exceeding TATmax*

Phlebotomy and or transport delay
Accessioning overload
Shift overlap, staff breaks, evening and night shift
Sample related problems
Improper use of stat procedures
Instrument (QC) or LIS failure

Educating clinicians on the proper use of stat procedures may reduce the number of stat tests requested, and increase laboratory TAT performance.
If a mutual agreement between clinical chemist and clinicians can be achieved on TATmax under certain well defined conditions, e.g. pH, blood gas and electrolyte measurements in heart surgery, the clinical chemist has to investigate whether this goal can be met in his or her laboratory. TAT data filter options as available in the TAT software can be of great help in this matter because within a few minutes laboratory and or testing time for any POC (ward, intensive care, emergency room, OR) can be generated and evaluated. If this results in the conclusion that certain TAT goals cannot be met with the present laboratory testing procedures, procedural and or organizational changes have to be considered, including POCT. In the thoracic surgery department of CCUHR, most stat testing is requested for pH, blood gases, electrolytes, CBC and Hb analysis. As can be seen from the results in table 5.3, median testing time for both test groups concerned (F and G) are significantly shorter than for the same test groups in the central lab and therefore satellite lab testing is justified if, as can be expected [6], clinician TAT requirements are also well below central lab median TAT.
In the satellite lab in the thoracic surgery department, special interest is paid to samples from the OR during open heart surgery. In open heart surgery, the recommendations of the American Heart Association (AHA) are adapted, which means that test results for pH, gases and electrolytes should be available within 5 minutes. Although the implementation of new analyzers and procedures in october 93 resulted in a significant decrease of median and 90th percentile TAT (figure 5.4) it is obvious that the AHA recommendations cannot be met in the satellite lab (figure 5.5). Based on these objective TAT data there was a mutual agreement between the departments of clinical chemistry and thoracic surgery to implement a POCT analyzer in the operating theatre. This resulted in a 90th percentile testing time of 4.0 minutes because pre-laboratory time was reduced to less than 1 minute (results not shown).
In this chapter is recommended that clinical chemists should start using TAT as a quality aspect of clinical chemistry laboratory procedures. This means that they should start monitoring and publicizing their TAT for all stat testing in their

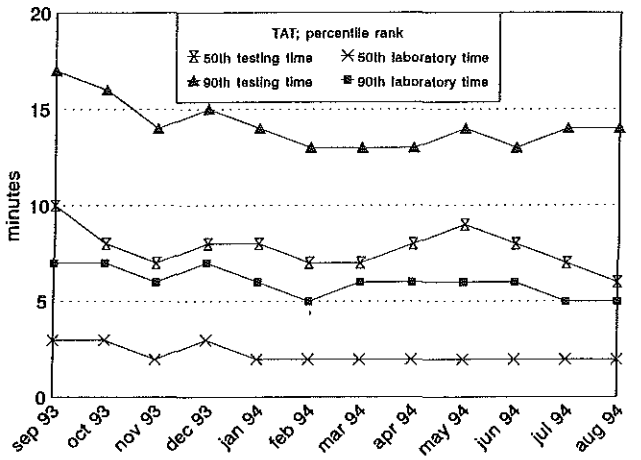


figure 5.4 *TAT statistics for samples from the OR, analyzed in the thoracic surgery department*

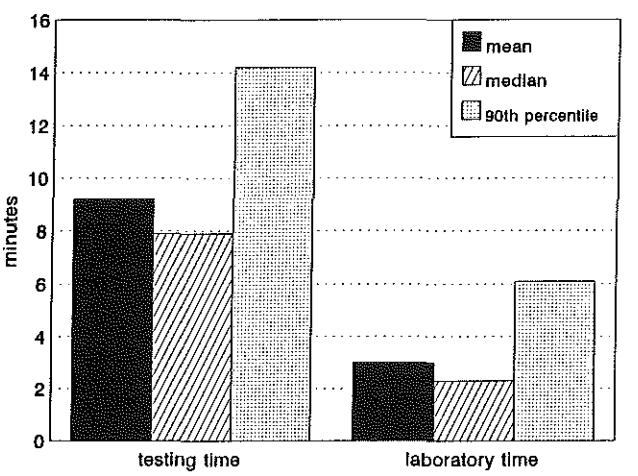


figure 5.5 *TAT components for samples from the OR, analyzed in the period september 93 - august 94 in the thoracic surgery department*

institution and that in a dialogue with the clinicians mutual definitions and goals for TAT should be realized. With continuous monitoring and identification of factors influencing TAT or causing unacceptable TAT delay, TAT and therefore total quality improvement can be expected and the percentage displeased clinicians will certainly be reduced.

References

1. Hilborne LH, Oye RK, McArdle JE, Repinski JA, Rodgerson DO. Evaluation of stat and routine turnaround times as a component of laboratory quality. Am J Clin Pathol 1989; 91: 331-335.
2. Valenstein PN, Emancipator KA. Sensitivity, specificity and reproducibility of four measures of laboratory turnaround time. Am J Clin Pathol 1989; 91: 452-457.
3. Hilborne LH, Oye RK, McArdle JE, Repinski JA, Rodgerson DO. Use of specimen turnaround time as a component of laboratory quality. A comparison of clinician expectations with laboratory performance. Am J Clin Pathol 1989; 92: 613-618.
4. Howanitz PJ, Steindel SJ. Intralaboratory performance and laboratorians expectations for stat turnaround times. A College of American Pathologists Q-probes study of four cerebrospinal fluid determinations. Arch Pathol Lab Med 1991; 115: 977-983.
5. Howanitz PJ, Steindel SJ, Cembrowski GS, Long TA. Emergency department stat test turnaround times. A College of American Pathologists Q-probes study for potassium and hemoglobin. Arch Pathol Lab Med 1992; 116: 122-128.
6. Howanitz PJ, Cembrowski GS, Steindel SJ, Long TA. Physician goals and laboratory test turnaround times. A College of American Patholgists Q-probes study of 2763 clinicians and 722 institutions. Arch Pathol Lab Med 1993; 117: 22-28.
7. Hallam K. Tornaround time: Speeding up, but is it fast enough? Med Lab Observer 1988; 20: 28-34.
8. Tiersten D. The stat problem. Clin Lab Med 1983; 3: 499-507.
9. Bluth EI, Lambert DJ, Lohmann TP, Franklin DN, Bourgeois M, Kardinal CG, Dalovisio JR, Williams MM, Becker AS. Improvement in stat laboratory turnaround time. A model continuous quality improvement project. Arch Intern Med 1992; 152: 837-840.
10. Saxena S, Wong ET. Does the emergency department need a dedicated stat laboratory? Continuous quality improvement as a management tool for the clinical laboratory. Am J Clin Pathol 1993; 100; 606-610.
11. Rollo JL, Fauser BA. Computers in total quality management. Statistical process control to expedite stats. Arch Pathol Lab Med 1993; 117: 900-905.
12. Pellar TG, Ward PJ, Tuckerman JF, Henderson RA. The freckle plot (daily turnaround time chart): A technique for timely and effective quality improvement of test turnaround times. Clin Chem 1993; 39: 1054-1059.

CHAPTER 6

MANAGING POINT-OF-CARE TESTING: A NEW CHALLENGE FOR THE CLINICAL CHEMIST

This chapter is based on:

Option Bio 1993; 105 (supplement) 1-14.

Summary

Quality aspects of analytical results are applicable to all testing regardless of where the tests are performed and by whom.
Point-of-care testing, if used for specific test in areas where rapid turnaround times are critical for clinical and/or financial reasons, can be a cost-effective extension of central laboratory services.
Point-of-care technology offers new possibilities for the clinical chemistry profession, provided the clinical chemists and their staff are prepared to change, to bring the service of the laboratory to the patient and to take responsibility for maintenance of the laboratory level of quality at the point-of-care.

Introduction

The shift towards patient-focused care marks a significant development in the healthcare environment of the 90's. One of the major effects of this shift on the laboratory is the implementation of point-of-care testing (POCT), defined as testing in all hospital areas where it is medically indicated, with (trans)portable or hand-held instruments. In the last few years, many organizations of clinical chemistry have held conferences on POCT and several documents have been published [1-10].

There are several aspects which play an important role in the creation of POCT in the current healthcare delivery system. Firstly, intensive care has expanded to serve increasing numbers of patients requiring more complicated treatment. Secondly, clinicians expect laboratories to deliver reliable results quickly, enabling them to make immediate patient management decisions. For example, the American Heart Association and the Emergency Care Research Institute recommended that blood gas and electrolyte measurements in open heart surgery should be available within five minutes [11]. In addition, economic factors continue to encourage improvement of clinical efficiency e.g. by shortening hospital stays and costs to the bare minimum. And finally, although the mere availability of POCT devices is not ample justification for their use, it is certainly a very important reason for their implementation.

The two most important criteria for the implementation of POCT ought to be if it can increase the quality of patient care and decrease hospital costs.

The scope of this chapter is to focus on the comparison of quality aspects of analytical results in the central laboratory and at the point-of-care. The quality of analytical results is usually expressed in terms of accuracy and precision, but other issues, such as pre-analytical factors, quality assurance procedures, registration of data, blood conservation, TAT, and cost aspects must also be considered. Further, the responsibility for POCT is defined and the clinical requirements for POCT in the hospital setting, together with selected aspects of measuring technology and system concepts developed to meet these requirements are illustrated.

Pre-analytical quality aspects

A test result is only as good as the specimen that is analyzed and it is vital therefore that each laboratory sets appropriate goals for important pre-analytical factors. One of the attractions of POCT may be the removal of some of these pre-analytical factors.
At our laboratory we recently evaluated a multichannel microchemistry instrument for pH, blood gas and electrolyte measurements, which can be used for POCT. Evaluation was first conducted in our central lab, then at the point of care. In the laboratory evaluation paired samples were analyzed within one minute on both analyzers. In the second part of the evaluation specimens were first analyzed in the pediatric intensive care unit where the analyzer was placed, and subsequently at the central laboratory (see chapter 7).
It was quite obvious that the method comparison correlation coefficients were closer to 1.0 in our laboratory evaluation than in our evaluation at the point of care. The same results were obtained previously [12] using other analyzers. Cell metabolism, the presence of air bubbles and influence of anti-coagulant are a possible explanation for the discrepancy in data between comparison studies at the central laboratory and the POC. Lost patient specimens, inaccurate specimen labelling and errors in data transmission and distribution also represent important sources of error in the pre-analytical phase. POCT should allow for the avoidance of such errors. It also reduces the time interval between sample collection and analysis and, consequently, the influence of environmental conditions. POCT does not, however, remove all pre-analytical factors affecting results. It has the disadvantage of involving a greater number of different sites and personnel.

Analytical quality aspects: precision and accuracy

Many strategies have been used to set desirable standards for the analytical goals for the performance of tests, particularly imprecision but also inaccuracy.
Some would suggest that analyses performed for screening, monitoring and aiding in diagnosis all require separate standards. But it is our opinion that the clinical

chemistry laboratory must ensure the same goals for imprecision and inaccuracy of test results regardless of where the tests are performed, the clinical purpose and by whom. The biological variation between individuals should provide the basis for the establishment of these specifications [13]. In daily practice the accuracy of POCT analyzers can be obtained by method comparison, using the central laboratory method as a reference. The importance of comparability with the central laboratory must be realized, because patients may be treated on a mixture of centralized- and decentralized results, depending on time and origin of the request.

Quality assurance procedures

Point-of-care microchemistry instrument quality assurance procedures should be identical to those in central laboratories to ensure the highest quality blood test results and to meet current or future regulations and or standards. It is unrealistic to believe that conventional quality assurance procedures will be applied routinely in POCT and, if they are, that the operator would necessarily act upon the results. The quality control of performance must therefore be engineered into the system to make it as fool-proof as possible.
The quality of results at the POC varies according to the complexity of the instruments used. Therefore, instruments requiring the least number of steps to operate, produce the most accurate results (figure 6.1).

Registration of data

A very important aspect of point-of-care testing is the registration of measurement data from patient samples as well as from quality control material. Central laboratories nowadays have extended facilities for data handling. Most larger analyzers have single- or bi-directional connections with a computer system which makes it possible to store the data efficiently in a laboratory data bank and present the results directly after analysis on a ward terminal or printer. The small,

sometimes hand-held, point-of-care equipment should contain data storage facilities so as to collect measurement data for several patients together with patient identification information and registration of date and time of day. Preferentially, this data can be easily transmitted to the laboratory computer system and added to the electronic laboratory data file of the patient where it can be reviewed in a later stage. Some point-of-care instruments allow for transmission through infrared signals to a receptive computer system, other systems store the data on a diskette so that later these can be downloaded in another computer system such as a personal computer. Other aspects of registration not to be forgotten are the data of the evaluation study of the point-of-care instrument and the education and training status of the testing personnel.

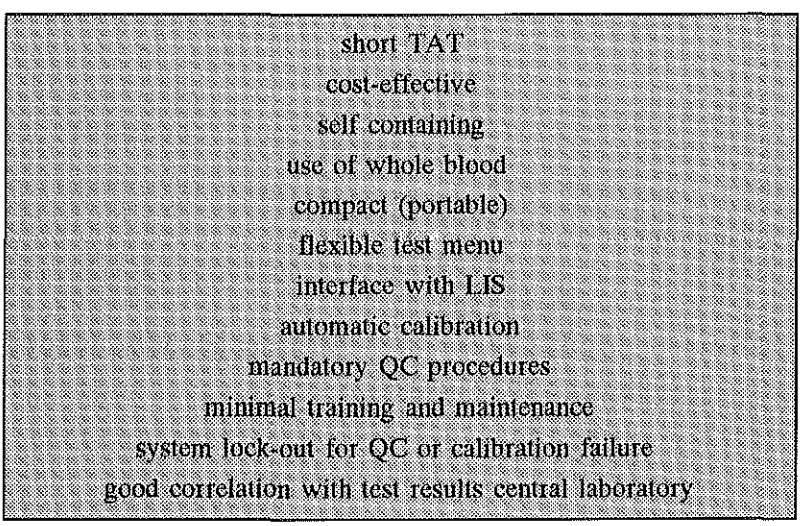


figure 6.1 *Ideal features of instruments for use at the point-of-care*

Blood conservation

Frequent blood sampling for diagnostic purposes is a common procedure with critically-ill patients. Smoller and Kruskall [14] have reviewed phlebotomy patterns for IC patients and reported that 47 percent of the patients requiring a

blood transfusion had more than 180 ml blood drawn for diagnostic purposes. A non-invasive approach to monitoring those patients seems extremely useful, although we currently lack the technology to non-invasively measure pH, blood gases, electrolytes, Hb/hematocrit and glucose. However, recent developments are quite promising [15]. At present it is possible to determine the above-mentioned analytes in a 300 μl sample and it should be considered a grave mistake to draw more than 1 ml of whole blood for point-of-care monitoring of these analytes.

Turnaround time

Turnaround times of laboratory tests are relatively uncommon quality indicators in clinical chemistry. TAT must be related to three separate times: laboratory time, testing time and therapeutic time [16].

Laboratory time refers to the time required to perform a test after the specimen has arrived in the laboratory.

Testing time involves the time it takes to obtain the specimen, transport it to the lab, perform the test and derive the result.

Therapeutic time begins with the test order and extends until a therapeutic action based on the result has been taken.

We recently determined the TATs for blood glucose measurements in our central laboratory. It was self-evident that the time required to obtain and transport a specimen is a major factor (80 %) slowing down overall TAT (figure 6.2). Equivalent results have previously been described [12][17].
Significantly reduced TAT therefore is probably the most critical feature of POCT. In addition to eliminating most of the opportunities for pre-analytical errors, the omission of steps associated with collecting, transporting and processing specimens has cut average TAT for most POCT to less than 5 minutes.
In the POC setting, fast TATs are especially important for two reasons. In

critically-ill patients who are physiologically unstable, chances are that great during the testing time that the analytes being measured may have changed in a significant way at the moment the results are derived. In addition POCT may allow for more rapid treatment of patients and may decrease unwanted therapies because results are obtained immediately.

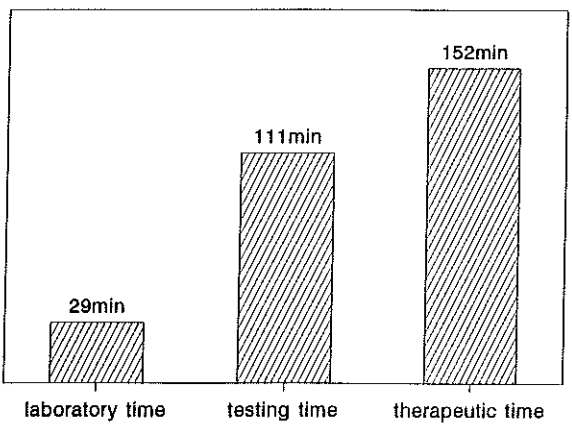


figure 6.2 TAT's glucose measurements Dijkzigt Hospital Rotterdam

A recent study [18] provides an example of such an improvement of patient care: it was found that, with bedside potassium monitoring, the occurrence of arrhythmias and the use of anti-arrhythmic agents was decreased in patients undergoing cardiac surgery. Other investigators [19] recently mentioned the failure of laboratory performance to meet clinicians expectations for acceptable TAT. There is however little agreement among clinicians about what constitutes an acceptable TAT.

Ideally standards for acceptable TAT should be based on studies that document a relationship between reporting speed and patient outcome. Unfortunately there has been little research in this direction. Therefore a group of experts, drawn from clinicians and clinical chemists, should establish appropriate TAT guidelines based on a combination of clinical requirements, clinician expectations and laboratory capabilities. If a TAT of less than 5 minutes is considered necessary for a certain analyte (e.g. in open heart surgery), a POCT analyser should be considered.

Cost Considerations

The true cost of carrying out tests, whether inside or outside the laboratory, is not easy to determine. Some techniques have been developed to assist in determining testing cost by identifying the amount of labour and supplies used to perform each test [20][21][22].

The CALC model [20] is the most appropriate one to identify which of two testing configurations has the best cost-benefit ratio. This model estimates the total annual testing costs through an addition of estimated costs in several cost categories: site preparation costs, equipment depreciation costs, equipment maintenance costs, incremental labor costs, supply and disposable costs and indirect costs, the costs that cannot be traced to a specific test or a particular piece of equipment. In virtually every situation, labor costs represent the highest percentage of total cost. It is well recognized that 60 to 70 percent of all hospital operating expenses are directly or indirectly related to labor costs [23].

For those hospitals already operating satellite labs, POCT may provide an alternative with a better cost-benefit ratio. Although TAT in satellite labs is generally acceptable, 24h per day staffing requires at least 5 fte, making satellite testing the most expensive of all testing.

Testing facilities introduced at the point-of-care do not replace the need for the corresponding testing at the central laboratory, they only supplement this testing. The main laboratory must still provide these tests 24 hours a day for all other patients and serves as a back-up facility for the POCT. Thus, the reduction in the numbers of tests at the central laboratory as a result of the availability of a POCT device only leads to minor savings in operating costs and, more importantly the costs of technical personnel. POCT is, however, usually more cost-effective than the addition of a satellite lab with its resulting additional labour.

When testing is moved to the point-of-care most of the expenses incurred in the central laboratory remain essentially unchanged, making the cost of processing tests at the point-of-care appear to be higher. As noted before, however, there may be significant cost savings for other aspects of patient care. This requires shifting of the focus from the lab to the hospital as a whole when regarding expenditures. This combines cost-benefit analysis with cost-effectiveness analysis,

that is measurement of outcome in monetary terms and natural units such as quality of care and the number of lives saved.
It has been estimated that costs generated by intensive care beds are 200 to 400 percent higher than those of other hospital beds [24][25], and that critical care services impact 10 to 20 percent [24,25] of overall hospital expenditures for a disproportionately small percentage of patients. The length of stay in those "high-rent" critical care areas can be shortened considerably with decreased TAT. Zaloga compared the average time to wean an uncomplicated postoperative patient from a ventilator, average ICU and hospital stay, in a group treated with an bedside oximeter and those monitored with blood gases in an ICU stat laboratory [17]. Bedside monitoring improved patient care, decreased weaning time, ICU stay, total hospital stay and therefore total hospital cost. Glucose, ionized calcium potassium, pH, pCO_2 and pO_2 measurements may also improve care and decrease costs when tested at the point-of-care. Finally, POCT may not only decrease the number of stat orders to the lab, but may also decrease the total number of tests ordered in the hospital. If a clinician knows that he must wait one hour for a stat result, he will continuously order stat after stat just in case a patient's condition might change. But if he or she knows that the results are available in 5 minutes, as in POCT, he or she will only order tests when it is really necessary.

Responsibility

Another important issue is to define responsibility for point-of-care testing. Many laboratorians have reasoned that not every clinical chemical activity outside the central laboratory can be supervised by the clinical chemistry department. Another view is that everything that can be defined as clinical chemistry within the institution should be performed under the supervision of the clinical chemistry department. The clinical chemist and the clinicians involved should come to an agreement on the question of responsibility. The clinical chemistry department should hold the responsibility, which extends to the following aspects:

1. The choice and evaluation of the necessary methods and equipment, keeping in mind when tests should be done and with what specifications. Special attention should be paid to the comparability of the results with the laboratory instruments and to aspects such as linearity, measurement interval, interferences, etc.

2. The implementation of an appropriate quality control system, including external proficiency testing, and apt to assure that the quality goals formulated above are reached in practice. Directives of ISO 9000 or EN 45000 should be applied to point-of-care testing as well as to central laboratory activities.

3. Supervision of the analytical process. If nursing staff is carrying out the testing, will there be direct supervision by the laboratory staff or is there an indirect line of control through the head of the clinical department?

4. Education and training of the testing personnel which should include instructions on preparation of the patient, sample handling, basic knowledge of the measuring principle, reference values and significance of deviating results, maintenance of equipment and supplies, and the practice and significance of the quality control procedures.

5. Budget control which, of course, depends strongly on the local situation.

An additional point to be negotiated is who should perform the tests, the laboratory or nursing staff. The key issue here is motivation. Nurses are generally not that eager to take on responsibilities beyond their primary task of nursing care. However, if they are motivated and an agreement on instruction, training, maintenance and quality control procedures can be reached, this solution could be advantageous for patient care. According to many publications however, experience has shown that the performance of point-of-care instruments is generally better in the hands of professional laboratory personnel. They have more experience in matters such as pre-analytical factors, instrument control and data registration than do non-technical operators.

Clinical requirements in POCT

Most patients in critical care are in an unstable medical and physiological state. They require hemodynamic and respiratory monitoring to check the effectiveness of oxygenation, tissue perfusion and organ function. An extensive range of physiological monitors exists, from both invasive and non-invasive monitoring of heart rate, blood pressure, central venous pressure, oxygen saturation, end-tidal CO_2 and other parameters. In addition to fluid and electrolyte balance, laboratory support includes analyses for cardiac and respiratory function, for hepatic and renal function, and tests aimed at early detection of complications arising from the admitting pathology; from immobilisation, such as thrombosis and embolism; from surgical complications, bleeding, transplant rejection; cardiac arrest, sepsis and other events that can occur without warning, at any time of the day or night. Only a selected group of tests, those that require a TAT of 5 minutes or less, are candidates for POCT (figure 6.3). Short TAT's for these analytes have an impact on the clinical efficiency of POC areas, by allowing faster response to the medical needs of patients.

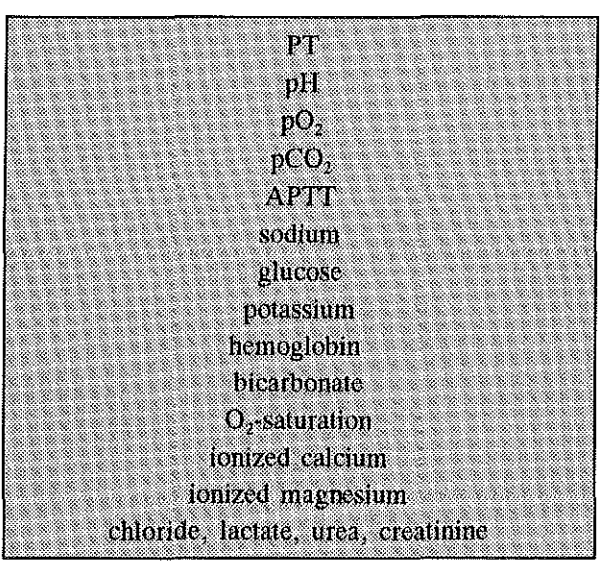


figure 6.3 *Point-of-Care test profile*

Technology and system concepts

Measurement technology by itself is not the principal driving force in the evolution of POCT analyzer concepts. Several technologies that contribute to system development are related to design and manufacturing. Computer assisted design techniques contribute to ergonomics, and to ease of manufacture. Microelectronics, with low electrical power requirements despite extensive software and memory, leads to portable miniaturisation. From advances in materials science, new plastics and polymers are used for structural components in instruments, reagent containers, and reagent matrix supports. System design, choice of components, and manufacturing techniques in a modern full-quality organisation, lead to the reliability needed in the final instrument.

Reagents can be stabilized in cassettes, by techniques such as lyophilisation, tabletting, adsorption on polymers, or even as liquid components in separated compartments to avoid chemical degradation. When reconstituted automatically in the system, these provide stable performance even after prolonged storage.

The basic idea however, is not new. Other dry chemistry approaches have been in the central laboratory for over a decade.

What is new for many POCT designs is in the selection of new plastics, in moulding and sealing techniques to avoid leakage, in the neutralisation of surface charge in reagent containers by techniques such as radiofrequency plasma treatment, and other manufacturing advances of which users are rarely aware. Manufacturing technology, rather than analytical technology, can contribute to reagent stability, to system reliability, and to reproducible analytical performance.

Variation in operator technique and reagent manipulation errors can be designed out by automation and by designs that use physical forces such as diffusion and capillary attraction, on supports of controlled pore-size, surface charge and contact angle that determine transport and reaction times. Antibody and ligand coatings can be laid down that selectively bind or separate analytes, interferents, red cells etc. Such fluidic circuits can be moulded in plastics, or printed in hydrophobic materials into a porous support, and are especially suited to immunoassay [26] and enzymatic reactions.

table 6.1 *Point-of-care analyzers and test menus*

Instrument/ *field*	Manufacturer	Test profile
pH, blood gas, electrolytes		
238	Ciba Corning	pH, pCO_2, pO_2
Gem Premier	Mallinckrodt Sensor Systems	pH, pCO_2, pO_2, Sodium, Potassium, ionized Calcium, Hematocrit*
IRMA	Diametrics Medical Inc	pH, pCO_2, pO_2
i-STAT	i-STAT Corp	pH, pCO_2, pO_2, ionized Calcium
StatPal II	PPG Industries	pH, pCO_2, pO_2
General Clinical Chemistry		
HemoCue	HemoCue AB	Glucose
i-STAT	i-STAT Corp	Sodium, Potassium, Chloride, Urea, Glucose, Haematocrit*
Piccolo	Abaxis Inc	Alkaline phosphatase, α-Amylase, Aspartate aminotransferase, total Bilirubin, Cholesterol, C-reactive protein, Creatinine, Glucose, Potassium, Total protein, Urea, Uric acid, Haematocrit*
Hemostasis and Thrombosis		
BioTrack 512	Ciba Corning	PT, APTT.
TAS	Cardiovascular Diagnostics	PT, APTT, Heparin Management Test (Activated clotting time), Streptokinase antibodies, Lysis Onset Time

* indirect measurement (conductivity)

Volumetric separation can be achieved through entrapment in a porous matrix, or by centrifugal techniques, that make it unnecessary to accurately dispense whole blood samples. Electrochemical sensors also minimise reagent manipulation, and contribute to instrument simplicity by avoiding optical components.

Factory calibration simplifies operating procedures, and leads to greater comparability between testing sites. Bar codes read by hand-wand, or directly in the instrument, carry lot-specific reagent information that provides the correct calibration curve to which the reaction should be compared. When the manufacturers' "master lot" of calibrator has been referenced to an international standard, it is more likely that the point of care result will be directly comparable to the laboratory result, also traceable to that standard.

Finally, genetic engineering, particularly recombinant DNA allows the production of recombinant materials that can be expected to improve instrument comparability.

POC analyzers

Some examples of dedicated POC analyzers of today are presented in table 6.1.

References

1. Wiemer K. Pathology measurements closer to the patient? J Clin Pathol 1980; 332: 857-863.
2. Marks V, Alberti KGMM, eds. Clinical biochemistry nearer the patient. Edinburgh: Churchill Livingstone, 1985: 240 pp.(Alberti KGMM, Galen RS, Marks V, eds. Contemporary issues in clinical chemistry; vol 2.)
3. Marks V, Alberti KGMM, eds. Clinical biochemistry nearer the patient II. London: Baillière Tindall, 1986: 192 pp.
4. Ashby JP, ed. The patient and decentralized testing. Proceedings of the ECCLS conference on decentralized clinical testing. Copenhagen 1986. Lancaster: MTP Press Limited, 1987: 128 pp.
5. 13th annual Arnold O Beckman Conference on clinical chemistry. Critical care medicine: Technology and patient management. Clin Chem 36 (special issue 8B); 1990: 1533-1622.
6. Point-of-care testing: A laboratory guide to meeting changing demands. MLO 1991 (special issue 9S) 23: 1-51.

7. Point-of-care testing: A laboratory guide to meeting changing demands II. MLO 1992 (special issue 9S) 24: 1-41.
8. Total quality management. Point-of-care testing III. MLO 1993 (special issue 9S) 25: 1-52.
9. ECCLS Standing Action Committee. Good practice in decentralized analytical clinical measurement. Scand J Clin Lab Invest 1992; 209: supplement.
10. Chernow B ed. Bedside diagnostic testing. Chest 1990; 97 (supplement): 183S-214S.
11. Emergency Care Research Institute. Risk analysis: Cardiopulmonary perfusion equipment. J Extracorporeal Technol 1987; 19: 238.
12. Salem M, Chernow B, burke R, Stacey JA, Slogoff M, Wood S. Bedside diagnostic testing. Its accuracy, rapidity and utility in blood concervation. JAMA 1991; 266: 382-389.
13. Fraser GC, Hyltoft Petersen P, Ricos C, Haeckel R. Proposed quality specifications for the imprecision and inaccuracy of analytical systems for clinical chemistry. Eur J Clin Chem Clin Biochem 1992; 30: 311-317.
14. Smoller BR, Kruskall MS. Phlebotomy for diagnostic laboratory tests in adults. N Eng J Med 1986; 314: 1233-1235.
15. 24th annual Oak Ridge conference on advanced analytical concepts for the clinical laboratory. Clin Chem 1992; 38: 1595-1926.
16. Zaloga GP. Monitoring versus testing technologies: Present and future. MLO 1991: 23: 20-31.
17. Zaloga GP. Evaluation of bedside testing options for the critical care unit. Chest 1990; 97 (supplement): 185S-190S.
18. Strickland RA, Hill TR, Zaloga GP. Bedside analysis of arterial blood gases and electrolytes during and after cardiac surgery. J Clin Anesth 1989; 1: 248-252.
19. Hilborne LH, Oye RK, McArdle JE, Repinski JA, Rodgerson DO. Evaluation of stat and routine turnaround times as a component of laboratory quality. Am J Clin Pathol 1989; 92: 613-618.
20. Statland BE, Brzys K. Evaluating STAT testing alternatives by calculating anual laboratory costs. Chest 1990; 97 (suplement): 198S-203S.
21. National Committee for Clinical Laboratory Standards. Cost accounting in the clinical laboratory. NCCLS document GP11-P 1990; vol 10: no 13.
22. Conn RB, Aller RD, Lundberg GD. Identifying costs of medical care. JAMA 1985; 253: 1587.
23. Roberts RD. Preparation of the budget and the flexible operating plan. Topics in Health Care Financing 1979; 5: 155.
24. Birnbaum ML. Cost containment in critical care. Crit Care Med 1986; 14: 1068-1077.
25. Bams JL, Miranda DR. Outcome and costs of intensive care. Intensive Care Med 1985; 11: 234-241.
26. Thorpe GH. Disposable integrated immunoassay services. Proceedings of the ABC national meeting 1993, Birmingham. p29-30, Association of Clinical Biochemists, UK.

CHAPTER 7

LABORATORY EVALUATION OF AND POINT-OF-CARE TESTING WITH THE GEM PREMIER PH, BLOOD GAS AND ELECTROLYTE ANALYZER

Submitted for publication

Summary

The subject of this chapter is the evaluation of a Gem Premier multichannel microchemistry instrument in the clinical setting for which it is intended, both in terms of comparative performance against clinical chemistry laboratory instruments and its potential utility in decreasing turnaround time and pre-analytical error. This evaluation at the point-of-care was preceded by a laboratory evaluation to provide for imprecision, inaccuracy and method comparison data under laboratory conditions.

The Gem Premier system produces results that show a significant correlation with results of established laboratory systems. Imprecision data are acceptable for within-cartridge imprecision. The between-cartridge imprecision exceeds the criteria for maximum allowable imprecision based on one or more generally accepted strategies. This is partly compensated however, by the removal of sources of error in the pre-analytical phase. The inaccuracy of means of assay results is acceptable if judged according to reference method values.

Combined with the ease of use, data registration facilities, system lock-out features for QC or calibration failure and minimal maintenance this makes the Gem Premier suitable for use in a point-of-care setting where the availability of test results with TAT's < 5 minutes is considered necessary.

Introduction

Compact pH, blood gas and electrolyte analyzers with turnaround times (TAT's) within minutes are now available and can be used at the point-of-care (POC). The Gem Premier represents a new generation of such POC monitors. It utilizes polarography, potentiometry and conductance for the measurement of pH, blood gases, sodium, potassium, ionized calcium and hematocrit (Ht). The evaluation of this and other point-of-care instruments requires testing both in the laboratory and at the point-of-care. In this chapter a two-centre evaluation of the analytical aspects of the Gem Premier is presented, together with the experiences that have been obtained with this instrument during a study in a neonatal intensive care unit (ICU). As such, the analyzer was taken to the ICU and measurements were performed by laboratory staff directly after sampling. In this setting the features of the analyzer are judged and the impact of point-of-care testing (POCT) on TAT, precision and accuracy is investigated and presented.

Materials and methods

Two laboratories participated simultaneously in the evaluation. Lab 1 refers to the Department of Clinical Chemistry, University Hospital Rotterdam, Sophia Children's Hospital, the Netherlands. Lab 2 refers to the Department of Clinical Chemistry and Hematology, Center of Obstetrics and Gynecology, University of Bonn, Germany. All analyzers are operated according to the manufacturer's standard operating and maintenance procedures and without any special adjustments carried out by the manufacturer. No adjustments of slope and or intercept were used.

Gem Premier

The Gem Premier (Mallinckrodt Sensor Systems, Ann Arbor, Michigan, USA) is a (trans)portable pH, blood gas and electrolyte system designed for the measurement of pH, pCO_2, pO_2, sodium, potassium, ionized calcium, and Ht. In addition to these measured parameters, base excess, oxygen saturation, total

carbon dioxide, and bicarbonate are derived and printed.
The heart of the Gem Premier is a disposable cartridge which contains all of the electrodes and calibration solutions [1]. This means no gas tanks or reagent preparation, no electrode maintenance or tubing changes, so no routine maintenance. System setup consists of inserting the cartridge into the instrument. The next 60 minutes, the electrodes are rehydrated and the sensor block is automatically brought to a constant temperature and than calibrated.
Each cartridge is good for either 300 test or 7 days, whichever comes first. In the near future cartridges that can be used for 14 days will be available as well. Two-point calibrations have been programmed into the instrument to occur every 60 minutes for the first 20 hours, and every 120 minutes thereafter. One-point calibrations are automatically performed every 2 minutes for the first 12 hours, and every 6 minutes thereafter. The sample size is 200 μl and results are available 90 seconds from sample introduction. Because the system is rinsed for 2 minutes after each sample, the throughput is limited to 15 samples per hour.
Quality assurance is maintained on the system by several automated and user-programmable control functions e.g. mandatory QC and system lock-out for QC or calibration failure.

Comparison instruments
Lab 1 : IL BGE pH, blood gas and electrolyte analyzer and IL 237 tonometer, Instrumentation Laboratory, Milan, Italy. OSM 3 Cooximeter, Radiometer, Copenhagen, Denmark. H1 hematology analyzer (Ht measurements), Technicon, Tarrytown, USA.
Lab 2 : CCD 288 pH, bloodgas and electrolyte analyzer, Ciba Corning Diagnostics, Medfield, USA. IL 237 tonometer, Instrumentation Laboratory, Milan, Italy. CC 800 hematology analyzer (Ht measurements), TOA Medical Electronics Company, Kobe, Japan.

Quality control materials
Euro-Trol D blood gas and electrolyte controls, level 1 (L1, lot AD-2131-2), level 2 (L2, lot AD-2124-2) and level 3 (L3, lot AD-2133-2), were obtained from Euro-Trol, Wageningen, the Netherlands.

Within-day and within-cartridge imprecision

On each of 3 days, 10 sequential samples of L1, L2 and L3 and fresh, heparinized human whole blood samples (tonometered for pCO_2 and pO_2 measurements) were analyzed with one cartridge on the Gem Premier and with the comparison instruments. The second of duplicate measurements was used for imprecision calculations.

Between-day and between-cartridge imprecision

For 7 days L1, L2 and L3 were analyzed in duplicate with one cartridge on the Gem Premier and with the comparison instruments. This procedure was repeated twice on all instruments, each time by using a different cartridge for Gem Premier measurements.

Inaccuracy

Following the ECCLS guidelines [2] the accuracy should be studied by comparison with reference method values. The reference method of pH [3] was not available in the evaluation laboratories, for that reason the mean values from the between cartridge imprecision measurements for L1, L2 and L3 were compared with the pH reference method [2] values listed by Euro-Trol. The reference material for pO_2 and pCO_2 was obtained by tonometry. Reference method values for sodium were obtained by measurements according to the candidate reference method presented in chapter 3, using two different cartridges.

Tonometry

Following the IFCC recommendation on tonometry of blood [4], tonometry tests using fresh, heparinized human whole blood were performed on 2 days with 3 certified medical gas mixtures. For each gas mixture, the mean values of 10 consecutive measurements with 1 cartridge are used for inaccuracy calculations.

This procedure was repeated on a second day, using a different cartridge. The presented data represent the mean inaccuracy over two days. The temperature in the equilibration chamber (37 °C) was checked by an electrical thermo-probe, the barometric pressure was measured with a mercury barometer.

Method comparison in the central laboratory

Approximately 100 heparinized anaerobically handled whole blood samples were analyzed over at least 5 operating days with the Gem Premier and the comparative instruments. Patient samples were selected at random from those

routinely received in the central laboratory. Among these were cord blood samples and capillary, arterial or mixed venous blood samples from ICU's (gynecology and obstetrics, neonatology, pediatrics and surgery). All samples were analyzed in singlet by laboratory staff and all testing was completed within 10 minutes of specimen receipt.

Method comparison at the point-of-care

During 10 working days, the Gem Premier was taken to the neonatal ICU in the Sophia Children's Hospital at the morning blood collection round. If a pH, blood gas and or electrolyte analysis was required, two samples (capillary or arterial blood) were collected. One sample was analyzed immediately on the Gem Premier by laboratory staff, the other sample was transported to the central laboratory (hand delivery) and analyzed with IL BGE according to standard stat procedures.

For each sample, the time the sample was collected and the time the results were available to the ICU staff was registered for both central laboratory testing and point-of care testing.

Data analysis

Method comparison data were evaluated after outlier rejection according to the ECCLS EP-9T protocol [5]. The line of best fit was calculated and a 95% confidence limit was calculated for slope and intercept according to *Passing* and *Bablok* [6]. Differences between means of series of assay results were investigated with the *Student's* paired t-test (significant difference if $p < 0.05$). The strength of the relationship between methods was determined also using *Pearson's* product correlation coefficients (r). Data are presented graphically using residual plots in which the differences between Gem Premier and comparison instrument are plotted versus the comparison method assay results.

Turnaround time (TAT) is defined and calculated as time between sample collection and availability of the results at the ICU. The mean, median and 90th percentile turnaround time for all measurements at the point-of-care method comparison study are calculated.

Significant differences between TAT results were investigated with the *Wilcoxon* signed rank test (abnormal distribution, significant difference if $p < 0.05$).

Acceptability
The Gem Premier is judged primarily on the extent of concordance with the requirements for ideal POCT instruments (figure 6.1) and on general criteria for acceptable imprecision and inaccuracy based on within-subject biological variation (wsbv) [7]. However, no such data are available for pH, pCO_2 and pO_2, and acceptable levels of imprecision an inaccuracy have not yet been established for POCT. Therefore, three additional strategies for setting quality specifications for acceptable imprecision and inaccuracy have been added: (a) The *state-of-art* performance (SoA) derived from the SKZL proficiency testing programs from 1992, (b) The opinions of more than 30 clinicians (OC) working in the ICU of the Center of Surgery at the Bonn University Hospital and in the neonatal and pediatric ICU's at the Rotterdam University Hospital, and (c) The CLIA analytical quality requirements [8][9]. In all additional strategies the maximum allowable imprecision is defined as one half the total analytical variation, and the maximum allowable inaccuracy is defined as twice the maximum allowable imprecision [7].

Results and discussion

The purpose of this investigation was to evaluate a multichannel microchemistry instrument in the clinical setting for which it is intended, both in terms of comparative performance against clinical chemistry laboratory instruments and its potential utility in decreasing TAT and pre-analytical error. This evaluation at the POC was preceded by an evaluation at the clinical chemistry laboratory to provide for imprecision, inaccuracy and method comparison data under laboratory conditions.
Imprecision The results of the imprecision measurements are presented in table 7.1. The Gem Premier within-cartridge imprecision was < 5% for all analytes and comparable to comparison instruments performances. Within-cartridge imprecision was < 10% with exception of ionised calcium measurements in the low concentration range. The between-day imprecision for the comparison instruments was < 4% for all analytes and therefore considered significantly

table 7.1 *Imprecision of the Gem Premier and the comparison methods*

Analyte	Con-trol	assigned value	Gem Premier		IL BGE		CCD 288	
			within cartridge	between cartridge	within day	between day	within day	between day
			lab1/lab2	lab1-lab2	lab 1	lab 1	lab 2	lab 2
			CV %	CV %	CV %	CV %	CV %	CV %
pH	L1	7.193	0.11/0.12	0.12/0.20	0.04	0.11	0.04	0.04
	L2	7.400	0.09/0.11	0.10/0.16	0.03	0.03	0.04	0.06
	L3	7.598	0.09/0.07	0.12/0.08	0.05	0.04	0.03	0.04
	Bl	7.37	0.11/0.09		0.07		0.07	
pCO_2	L1	8.04	3.3/2.6	3.4/3.8	1.46	3.23	1.18	1.25
(kPa)	L2	5.27	2.2/1.8	1.8/3.3	1.01	1.26	1.18	1.31
	L3	3.20	2.1/1.0	4.7/5.3	1.28	2.13	0.98	1.04
	Bl	6.25	2.1/1.5		1.4		1.5	
pO_2	L1	4.59	4.9/3.4	8.4/9.5	3.53	3.32	3.52	7.29
(kPa)	L2	11.10	3.6/2.0	4.1/8.0	3.11	2.96	2.68	3.07
	L3	17.86	2.7/1.6	4.5/9.0	2.91	2.70	2.63	2.93
	Bl	6.60	1.3/1.5		1.25		1.08	
Sodium	L1	119.9	0.63/0.72	0.49/2.21	0.63	0.68	0.62	0.74
(mmol/l)	L2	140.6	0.31/0.85	0.42/1.10	0.55	0.75	0.46	0.53
	L3	160.3	0.38/0.46	0.67/1.64	0.26	0.65	0.38	0.36
	Bl	141.4	0.60/0.54		0.46		0.38	
Potassium	L1	3.01	0.40/0.56	0.49/1.24	1.00	1.52	0.65	0.64
(mmol/l)	L2	4.48	0.39/0.52	0.79/1.10	0.90	1.12	0.38	0.42
	L3	6.24	0.33/0.34	0.90/1.32	0.90	1.35	0.38	0.37
	Bl	4.34	0.89/0.95		1.02		0.55	
iCalcium	L1	1.72	1.41/1.01	1.93/1.99	0.97	1.02	1.21	1.31
(mmol/l)	L2	1.17	0.97/2.09	6.48/4.15	1.03	1.23	1.37	1.41
	L3	0.74	1.30/2.68	11.8/7.77	1.47	1.45	2.02	2.14
	Bl	1.08	1.60/2.20		1.16		1.30	

Bl = human whole blood

better than Gem Premier between-cartridge imprecision. This could be expected because the between-cartridge imprecision in fact being a between-analyzer imprecion adds additional variance to the imprecision measurements. Because of the short lifetime of the cartridges (7 days) however, in daily practice between-cartridge imprecision corresponds to and therefore should be judged as between-day imprecision according to generally accepted procedures. The between-cartridge imprecision is most pronounced for all pO_2 and ionised calcium measurements and for pCO_2 measurements in the low concentration range (L3). The between-cartridge imprecision for pO_2 might be slightly better in daily practice, because the within-cartridge imprecision for human whole blood (Bl) showed to be markedly better than the within-cartridge imprecision for the QC materials L1, L2 and L3 (see table 7.1).

The quality specifications for imprecision, against which the acceptability of all analytical systems should be judged, are presented in table 7.2. As can be seen from the results in table 7.1, none of the instruments tested can fulfil the criteria based on wsbv for sodium measurements, but all instruments perform well within the limits based on wsbv for potassium measurements. These criteria are known to be too demanding for sodium measurements with present-day technology and therefore judgement on SoA criteria is justified as interim procedure for both imprecision and inaccuracy measurements [7].

table 7.2 *Maximum allowable imprecision based on different strategies*

Strategy	pH	pCO_2	pO_2	Sodium	Potassium	iCacium
	Maximum allowable imprecision (CV %)					
Biological variation	-	-	-	0.30	2.40	-
State-of-art performance	0.11	2.93	5.29	0.92	1.15	-
Opinion Clinicians	0.20	5.79	5.00	0.78	2.30	4.10
CLIA 88 regulations	0.27	5.99	7.94	1.42	5.75	-

For all other analytes, no primary specifications (based on wsbv) are available and other strategies based on SoA, OC and CLIA regulations were taken into account. According to these criteria both comparison instruments perform well

within the limits for all analytes under almost all circumstances. This is totally different for Gem Premier between-cartridge imprecision data. Except for potassium measurements, the maximum allowable imprecision exceeds the criteria based on one or more different strategies.

table 7.3 *Results of the inaccuracy study with the Gem Premier*

Analyte	Reference method	Reference method value (R)		Measured value (M)		Accuracy (M/R * 100 %)	
		lab 1	lab 2	lab 1	lab 2	lab 1	lab 2
pH	Potentiometry	7.193	7.193	7.149	7.159	99.4	99.5
	[3]	7.400	7.400	7.391	7.396	99.9	99.9
		7.598	7.598	7.609	7.620	100.1	100.3
pCO_2	Tonometry	4.63	2.78	4.81	3.23	103.9	116.2
(kPa)	[4]	6.53	5.71	6.36	5.99	97.4	104.9
		9.48	8.57	9.36	8.94	98.7	104.3
pO_2	Tonometry	6.62	5.65	6.82	6.53	103.0	115.6
(kPa)	[4]	11.27	13.22	11.48	14.30	101.9	108.2
		18.57	38.04	19.44	40.64	104.7	106.8
Sodium	Gravimetry	118.5		119.0		100.4	
(mmol/l)	[10]	142.7		143.1		100.3	
		161.4		162.0		100.4	

Inaccuracy In table 7.3 the results of the inaccuracy measurements are shown. For the analytes tested, Gem Premier inaccuracy was < 5% for all analytes with exception of pCO_2 and pO_2 measurements in lab 2. Based on the criteria for maximum allowable inaccuracy (table 7.4) there is a small negative bias for pH measurements in the low range (SoA and OC based criteria) and a small positive bias for sodium judged on the (too stringent) wsbv specifications. All other results are well within the limits based on each of the different strategies, except pCO_2 and pO_2 measurements in lab 2 in the low concentration range. These measurements are accurate in lab 1, again pointing at, and probably caused by the rather large between-cartridge imprecision for these analytes.

Method comparison The method comparison statistics are presented in table 7.5, and method comparison data are plotted graphically as residual plots in figures 7.1 to 7.8. Inter-instrument analyte comparisons (at the central laboratory and at the POC) were highly and significantly correlated for all analytes. The correlation coefficient however, was closer to unity in the comparison of each analyte in the central laboratory compared with the evaluation at the POC. Contamination by room air during storage at low temperature [11][12] [13], the presence of air bubbles [14][15] and the possibility of cell metabolism during storage [14][16] are all pre-analytical errors associated with pH, blood gas and electrolyte measurements. These errors are only relevant for central lab testing and therefore could be a possible explanation for the discrepancy between correlation data for central lab and POCT [17].

table 7.4 Maximum allowable inaccuracy based on different strategies

Strategy	pH	pCO_2	pO_2	Sodium	Potassium	iCacium
	Maximum allowable inaccuracy (%)					
Biological variation	-	-	-	0.20	1.60	-
State-of-art performance	0.22	5.86	10.58	1.84	2.30	-
Opinion Clinicians	0.40	11.58	10.00	1.56	4.60	8.20
CLIA 88 regulations	0.54	11.98	15.88	2.84	11.50	-

As can be seen from figure 7.1, Gem Premier pH values are well below comparison method values at low pH (pH < 7.15). Based on the results for the inaccuracy measurements, the Gem Premier results can be considered inaccurate at pH values < 7.15. During method comparison at the POC no pH values below 7.20 were seen (figure 7.1) which resulted in acceptable regression data (no significant differences in slope, intercept and means of assay results).

The method comparison data for pCO_2 measurements in lab 2 show a positive bias for pCO_2 > 7 kPa. Initially comparable results were obtained in lab 1 as well. On account of these results, the manufacturer changed the pCO_2 algorithm

table 7.5 *Method comparison study (x=comparison method, y=Gem Premier)*

		y = ax + b						
Analyte	Site	a	b	r	n	Mean x	Mean y	significances
pH	lab 1	1.111	- 0.836	0.9926	112	7.392	7.379	a,b,r,m
	lab 2	1.132	- 0.973	0.9964	105	7.293	7.277	a,b,r,m
	POC	1.118	- 0.868	0.8952	122	7.355	7.356	r
pCO_2	lab 1	1.044	- 0.289	0.9900	112	5.39	5.39	r
(kPa)	lab 2	1.193	- 0.598	0.9910	105	6.26	6.92	a,b,r,m
	POC	1.133	- 0.570	0.9219	122	6.34	6.58	a,b,r,m
pO_2	lab 1	1.125	- 0.775	0.9791	112	11.30	12.04	a,b,r,m
(kPa)	lab 2	1.007	- 0.180	0.9941	105	11.60	11.58	b,r
	POC	0.947	- 0.113	0.9333	122	7.70	7.13	r,m
Sodium	lab 1	1.066	- 8.470	0.9373	100	138.6	138.9	r
(mmol/l)	lab 2	0.936	9.268	0.9648	108	141.5	141.8	a,b,r
	POC	1.026	- 1.590	0.9355	91	137.2	138.8	r,m
Potassium	lab 1	0.955	0.168	0.9969	100	4.66	4.61	a,b,r,m
(mmol/l)	lab 2	0.909	0.419	0.9980	108	5.12	5.05	a,b,r,m
	POC	1.000	0.000	0.9823	91	4.89	4.92	r
iCalcium	lab 1	1.091	- 0.060	0.9010	100	1.17	1.17	r,m
(mmol/l)	lab 2	1.083	- 0.158	0.9912	108	1.31	1.27	a,b,r,m
	POC	0.800	0.230	0.8602	91	1.17	1.17	a,b,r
Ht	lab 1	1.118	- 0.025	0.9643	100	0.34	0.35	a,b,r,m
	lab 2	1.086	- 0.038	0.9460	95	0.34	0.33	b,r,m
	POC	1.173	- 0.018	0.8213	122	0.30	0.34	a,b,r,m
sO_2	lab 1	1.000	0.010	0.9476	100	0.87	0.87	r
	POC	1.500	- 0.535	0.6799	122	0.88	0.81	a,b,r,m

a: significant difference from slope = 1.0, b: significant difference from intercept = 0.0 both according to Passing and Bablok; r = significant correlation between x and y (Pearson's correlation, $p < 0.05$); m = significant difference between mean x and mean y (paired t-test, $p < 0.001$).

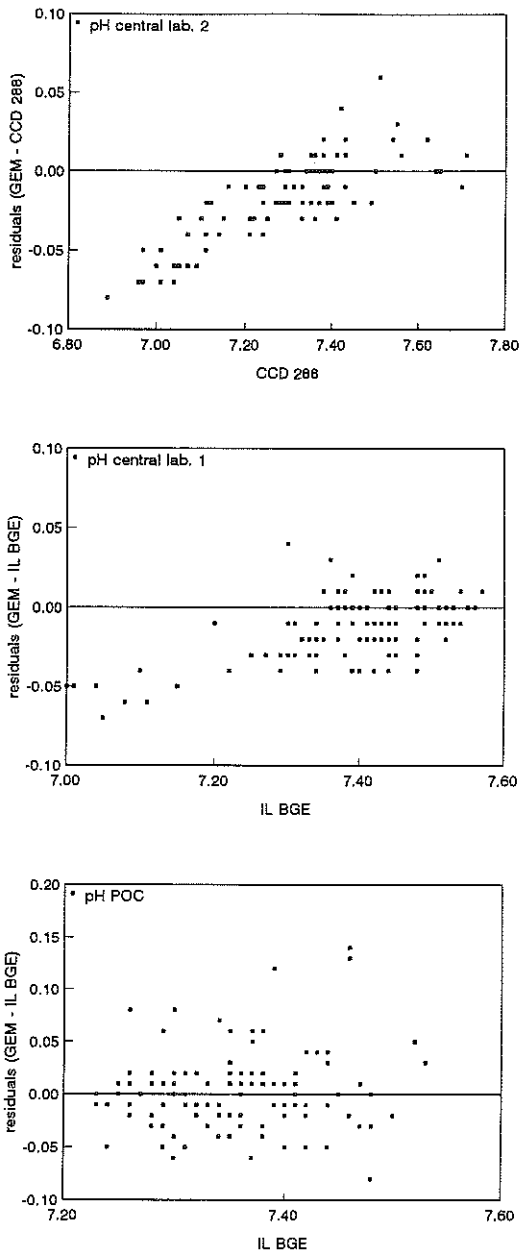


figure 7.1 Method comparison: residual plots pH measurements

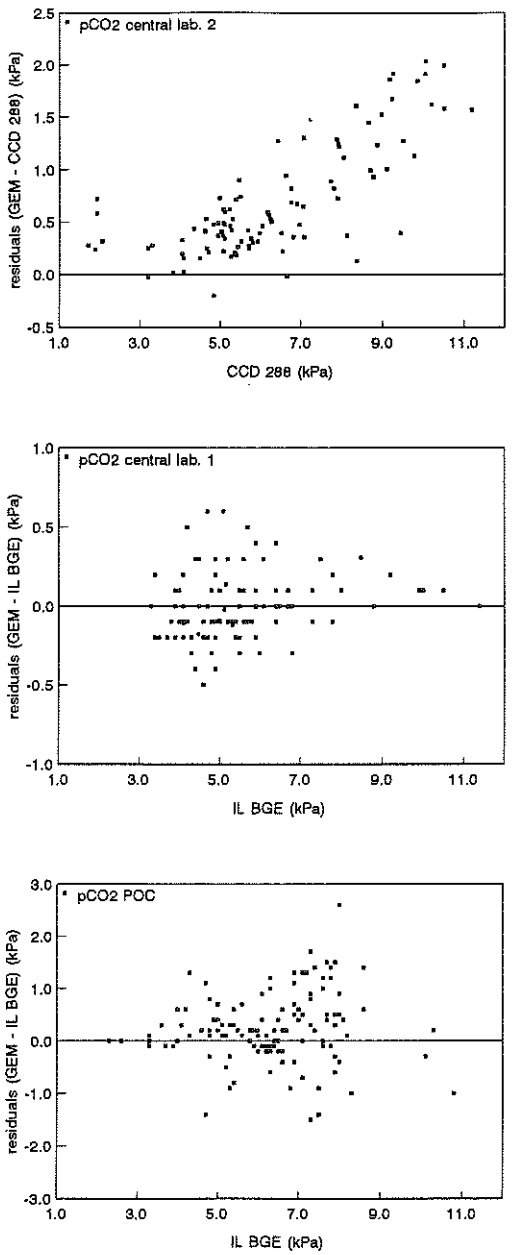


figure 7.2 Method comparison: residual plots pCO_2 measurements

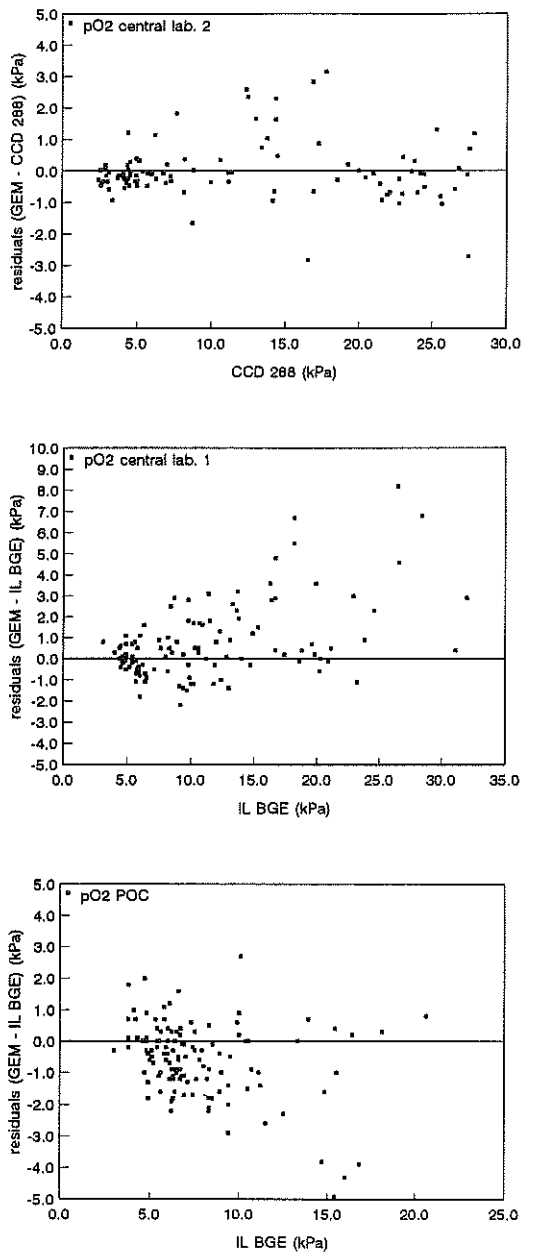


figure 7.3 *Method comparison: residual plots pO_2 measurements*

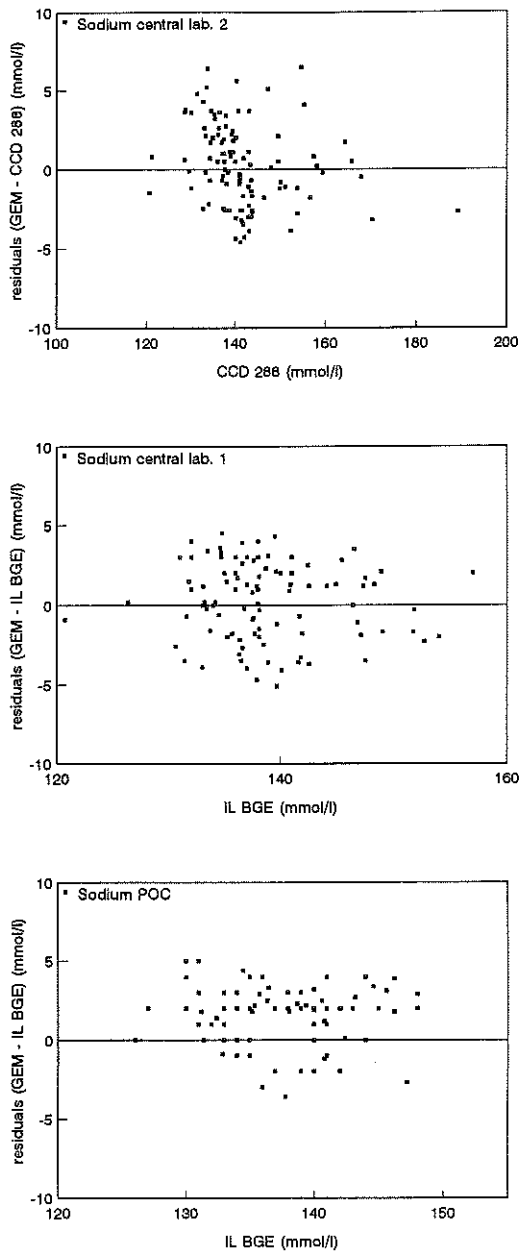


figure 7.4 Method comparison: residual plots sodium measurements

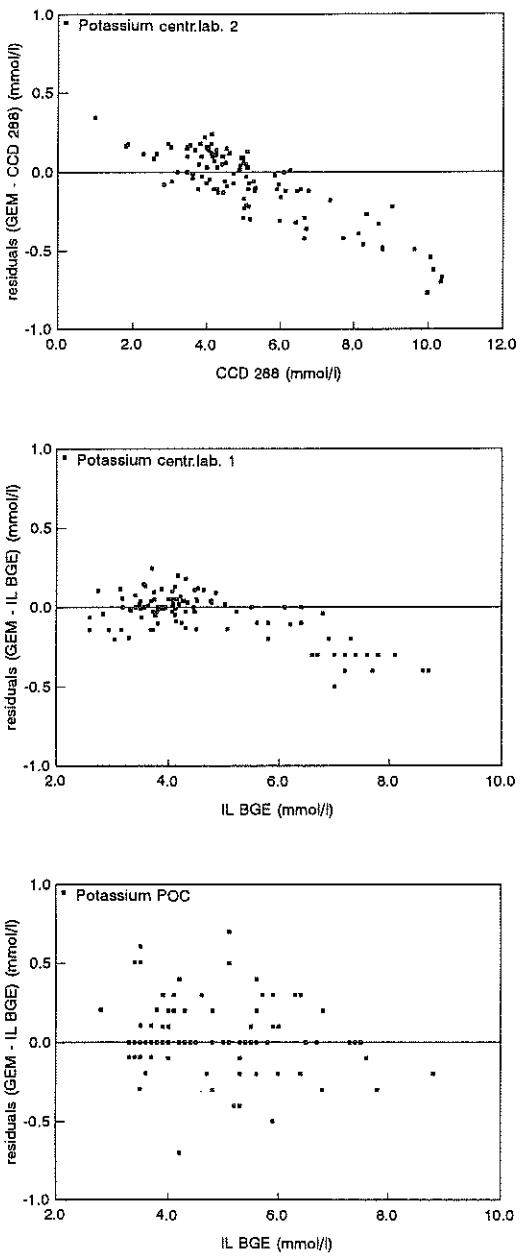


figure 7.5 Method comparison: residual plots potassium measurements

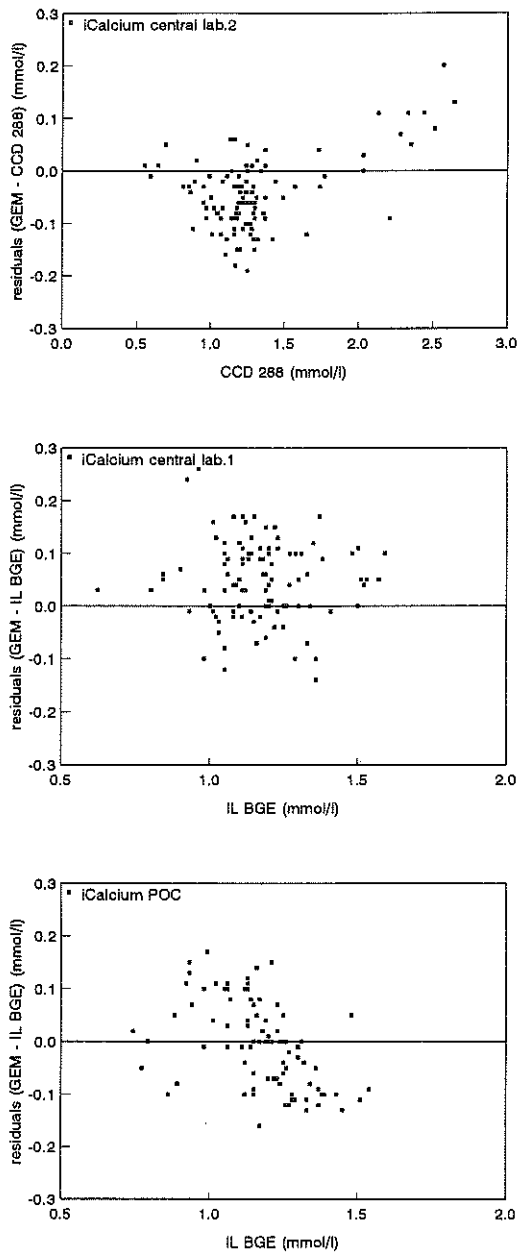


figure 7.6 *Method comparison: residual plots ionized calcium measurements*

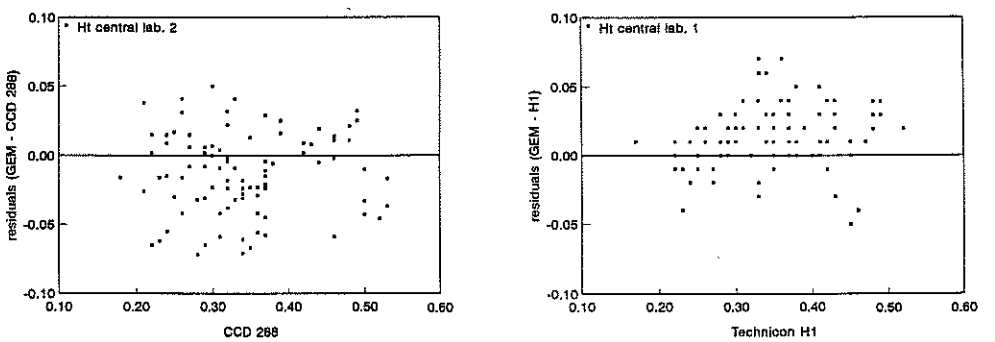


figure 7.7 *Method comparison: residual plots for hematocrit (Ht)*

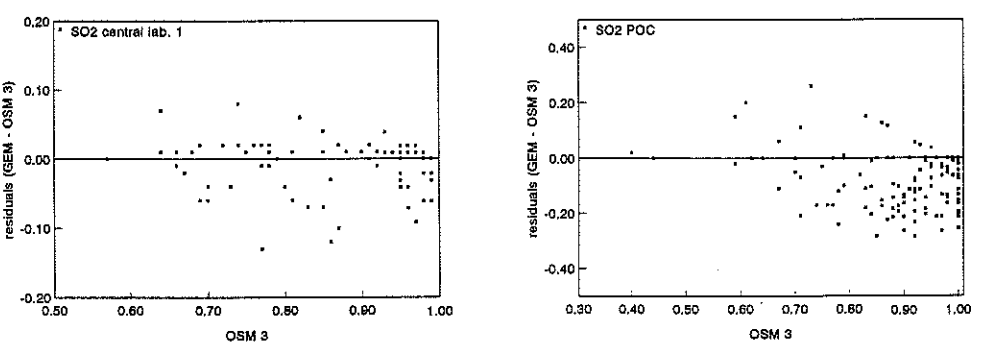


figure 7.8 *Method comparison: residual plots for oxygen saturation (sO_2)*

in a system software update. The method comparison data for lab 1, as presented in figure 7.2, are based on this new version of the software and show acceptable data for pCO_2 measurements over the entire measurement range.
Figure 7.3 illustrates the method comparison data for pO_2 measurements. These show acceptable data for pO_2 < 15 kPa. For pO_2 > 15 kPa, results are too high in the central lab1 method comparison and too low in the comparison at the POC, compared to the results of the central laboratory analyzer. In general, possible explanations for the different residuals for pO_2 > 15 kPa, is the influence of storage conditions, among which cell metabolism and diffusion of environmental air (pO_2 = 20 kPa) through the barrel of the syringe and into the blood are the most important. The influence of storage conditions on pO_2 under lab 1 testing conditions are presented in table 7.6. It is quite clear that at RT cell metabolism predominates (resulting in a reduction of pO_2) and that at 4°C the diffusion of pO_2 into the syringe compensates cell metabolism at normal pO_2 values. This compensation cannot be the only explanation for the differences between central lab and POCT. It has to be concluded that the between-cartridge variability also adds to the observed differences because different cartridges have been used in central laboratory and POCT (n = 5).

table 7.6 *Influence of sample storage or transport delay on test results*

	Mean difference from freshly drawn blood after 30 min storage	
	RT	4°C
pH	- 0.018	- 0.006
pCO_2	+ 0.10 kPa	+ 0.04 kPa
pO_2 (11.3 kPa)	- 5.3 %	1.5 %
pO_2 (18.6 kPa)	-21.3 %	1.8 %

The electrolyte results obtained from method comparison data are shown in figures 7.4 to 7.6. As was stated earlier [18], the scatter of sodium values was not as good as expected, but considered an acceptable trade-off for the more rapid electrolyte assessment necessary in critical situations. Correlation of the Gem Premier potassium analysis was excellent (figure 7.5), resulting in the ideal regression equation ($y = x$) in POCT. Ionised calcium measurements (figure 7.6) show a significant difference from comparison method values at all three testing sites. These differences cannot be judged, because no reference method for ionised calcium is available which could provide for the inevitable accuracy basis. The results on Ht measurements (figure 7.7) justify the opinion that indirectly obtained Ht measurements (based on conductivity) should only be used as a raw indication and not for clinical decision making.

The Gem Premier derived oxygen saturation (sO_2) is calculated, based on actual values of pH and pO_2 and the standard oxygen dissociation curve (ODC). This calculated sO_2 is compared to the oxygen saturation measured in the central laboratory (OSM 3). As can be seen from the results in figure 7.8 and table 7.5, there is a great number of samples with calculated $sO_2 > 0.10$ lower than measured sO_2, especially in the evaluation at the POC. Most of the outsiders are found at pO_2 values below 10 kPa and in capillary blood samples. The differences are more pronounced in the POC evaluation because here all samples were of neonatal origin. These may contain up to approximately 80% HbF, which causes a left shift of the ODC, resulting in a higher oxygen affinity of hemoglobin and erroneously low calculated sO_2 values. These results confirm those of a recently published multicenter study [19] showing that calculation of oxygen saturation should be used with great care for pO_2 values below 10 kPa and especially for capillary blood samples.

Turnaround time In table 7.7 and figure 7.9 results for the TAT measurements are plotted. It is perfectly clear that the median TAT for POCT was significantly shorter compared to median TAT for central laboratory testing. This significant reduction in TAT was also considered clinically relevant according to the opinions of the clinicians working in the neonatal and pediatric ICU's at the Rotterdam University Hospital.

table 7.7 *Turnaround time statistics for ICU and central laboratory testing procedures*

Turnaround time (min)	ICU	Central laboratory
Mean	6.2	38.3
Median	4.5	33.5
90th percentile	14.0	64.0
Minimum	2.0	11.0
Maximum	20.0	115.0

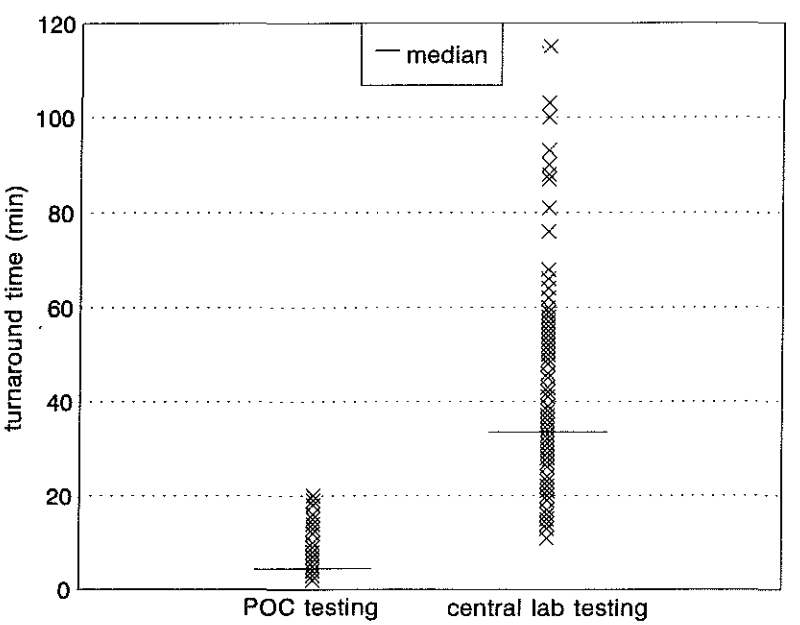


figure 7.9 Turnaround times for ICU and central laboratory testing procedures

Concordance with the requirements for ideal POCT instruments
The Gem Premier comes close to the requirements for an ideal POCT instrument presented in figure 6.1. No serious breakdown occurred during the evaluation period of more than 6 months, using different instruments and over 30 different cartridges. The only problems encountered were that two cartridges had to be exchanged because of pO_2 calibration errors. Although this can be done easily and the expelled cartridge is refunded by the manufacturer, this means 60 minutes down-time for system setup.
In the POC setting presented in this chapter, the instrument proved to be very useful and an obvious improvement in TAT as well as analytical quality could be firmly established. In particular the absence of clotting complications was appreciated. The troughput of the Gem Premier proved well suited to the described way of working.
Conclusions The Gem Premier system produces results that show a significant correlation with results of established laboratory systems. Imprecision data are acceptable for within-cartridge imprecision. The between-cartridge imprecision exceeds the criteria for maximum allowable imprecision based on one or more generally accepted strategies. This is partly compensated however, by the removal of sources of error in the pre-analytical phase. The inaccuracy of means of assay results is acceptable if judged according to reference method values.
Combined with the ease of use, data registration facilities, system lock-out features for QC or calibration failure and minimal maintenance this makes the Gem Premier suitable for use in a point-of-care setting where the availability of test results with TAT's $<$ 5 minutes is considered necessary.

References

1. Meyerhoff ME. New in vitro analytical approaches for clinical chemistry measurements in critical care. Clin Chem 1990; 36 (suppl): 1567-1572.
2. ECCLS. Guidelines for multi-center evaluation of analyzers for the measurement of blood pH, PCO_2 and pO_2. ECCLS Document 1989; 1.
3. Maas AHJ, Weisberg HF, Burnett RW, Müller-Plathe O, Wimberly PD, Zijlstra WG, Durst RA, Siggaard-Andersen O. Reference method (1986) for pH measurement in blood. J Clin Chem Clin Biochem 1987; 25: 281-289.

4. Burnett RW, Covington AK, Maas AHJ, Müller-Plathe O, Weisberg HF, Wimberly PD, Zijlstra WG, Siggaard-Andersen O, Durst RA. IFCC method (1988) for tonometry of blood: reference materials for pCO_2 and pO_2. J Clin Chem Clin Biochem 1989; 27: 403-408.
5. Method comparison and bias estimation using patient samples. NCCLS Document EP9-T 1993; vol 13: no 4.
6. Passing H, Bablok W. A new biological procedure for testing the equality of measurements from two different analytical methods. J Clin Chem Clin Biochem 1983; 21: 709-720.
7. Fraser GC, Hyltoft Petersen P, Ricos C, Haeckel R. Proposed quality specifications for the imprecision and inaccuracy of analytical systems for clinical chemistry. Eur J Clin Chem Clin Biochem 1992; 30: 311-317.
8. Final uniform proficiency testing for Clinical Laboratory Act of 1967 and Medicare laboratories. U.S. Department of Health and Human Services, Fed Reg 9538, March 14; 1990: 9538-9610.
9. Ehrmeyer SS, Laessig RH, Leinweber JE, Oryall JJ. 1990 Medicare/CLIA final rules for proficiency testing: minimum intralaboratory performance characteristics (CV and bias) needed to pass. Clin Chem 1990; 36: 1736-1740.
10. Van Suijlen JDE, Berrevoets CA, Leijnse B. A candidate reference method for coupled sodium-water determination in human serum. J Clin Chem Clin Biochem 1990; 28: 817-824.
11. Harsten A, Berg B, Inerot S, Muth L. Importance of correct handling of samples for the results of blood gas analysis. Acta Anaesthesiol Scand 1988; 32: 365-368.
12. Mahoney JJ, Harvey Ja, Wong RJ, Van Kessel AL. Changes in oxygen measurements when whole blood is stored in iced plastic or glass syringes. Clin Chem 1991; 37: 1244-1248.
13. Schmidt C, Müller-Plathe O. Stability of pO_2, pCO_2 and pH in heparinized whole blood samples: influence of storage temperature with regard to leukocyte count and syringe material. Eur J Clin Chem Clin Biochem 1992; 30: 767-773.
14. Mueller RG, Lange GE, Beam JM. Bubbles in samples for blood gas determinations. A potential source of error. Am J Clin Pathol 1976; 65: 242-249.
15. Biswas CK, Ramos JM, Agroyannis B, Kerr DNS. Blood gas analysis: effect of air bubbles in syringe and delay in estimation. Br Med J 1982; 284: 923-927.
16. Hess CE, Nichols AB, Hunt WB, Suratt PM. Pseudohypoxemia secondary to leukemia and thrombocytosis. N Engl J Med 1979; 301: 361-363.
17. Salem M, Chernow B, Burke R, Stacey JA, Slogoff M, Sood S. Bedside diagnostic testing. It's accuracy, rapidity, and utility in blood conservation. JAMA 1991; 266: 382-389.
18. Jacobs E, Nowakowski M, Colman N. Performance of Gem Premier blood gas/electrolyte analyzer evaluated. Clin Chem 1993; 39: 1890-1893.
19. Kokholm G. Simultaneous measurements of blood pH, pCO_2, pO_2 and concentrations of hemoglobin and its derivates - A multicenter study. Scand J Clin Lab Invest 1990; 203 (suppl.): 75-86.

CHAPTER 8

THE QUALITY OF TEST REQUESTING PATTERNS IN DIAGNOSING COBALAMIN DEFICIENCY

Submitted for publication

Summary

The clinical accuracy (sensitivity and specificity) of the results of the tests requested is the most important aspect to be considered when evaluating the quality of test requesting patterns. This is demonstrated by an example derived from the diagnosis of cobalamin (vitamin B12) deficiency. The assay of serum cobalamin is the only diagnostic test for vitamin B12 deficiency that is widely available. This test however, lacks both specificity and sensitivity. The availability of the accurate and precise measurements of the serum concentrations of methylmalonic acid and total homocysteine have provided for the necessary new tools in diagnosing cobalamin deficiency. The results in this chapter and those of others indicate that (1) hematological data cannot be used as a selection criterium for serum cobalamin requests, (2) 1 of 5 patients with low normal cobalamin has metabolic cobalamin deficiency and needs treatment and (3) measurements of serum methylmalonic acid and total homocysteine improve the quality of test requests in diagnosing and distinguishing cobalamin and folate deficiency. To improve the quality of test requesting patterns in the diagnosis of cobalamin deficiency, in sera of all patients with low or low-normal cobalamin values, all patients with unexplained hematologic abnormalities and all patients with unexplained neuropsychiatric abnormalities, methylmalonic acid and total homocysteine should be measured.

Introduction

There have been many attempts to influence test requesting patterns in clinical chemistry, with special emphasis on the efforts to control and modify the amount of inappropriate requests [1-9].

One of the difficulties has always been to reach agreement on what is inappropriate, due to the emphasis on reducing the number of requests for financial reasons only, rather than maximising their informative value. To my opinion the clinical accuracy (sensitivity and specificity) of the tests requested is the most important aspect to be considered when evaluating the quality of test requesting patterns. This is demonstrated in this chapter by an example derived from the diagnosis of cobalamin (Cbl; vitamin B12) deficiency.

The classical view on Cbl deficiency is that it presents with hematologic abnormalities (anemia, macrocytosis and hypersegmented neutrophils), biochemical abnormalities (increased serum lactate dehydrogenase (LDH) and or total bilirubin (Tbil)), and neuropsychiatric abnormalities. In these cases the hematological picture is taken as an important diagnostic feature [10]. It is well recognized that Cbl deficiency may sometimes present as a neurologic disorder without anemia or macrocytosis, but such cases are considered rather exceptional. A recent study however, showed that neuropsychiatric disorders due to Cbl deficiency occur commonly (in 28 percent of cases) in the absence of anemia and or elevated mean cell volume (MCV) [11]. Moreover it was demonstrated that Cbl deficiency exists in a large number of patients with only subtle biochemical changes, and no hematological and or neurological abnormalities [12]. This means that if, according to the classical view, Cbl deficiency is considered unlikely in the absence of anemia and or elevated MCV, LDH or Tbil, the correct diagnosis will not be made in a significant number of patients.

The assay of serum Cbl is the only diagnostic test for vitamin B12 deficiency that is widely available. This test however, has two serious disadvantages. Firstly, it is apparent that the serum Cbl concentration has a low specificity because a high proportion of patients with low values do not appear to be deficient in vitamin B12 [13][14]. Secondly, the assumption that the sensitivity of the assay is

(close to) 100% [15] has shown to be incorrect, because over 5 % of patients with clinically confirmed Cbl deficiency have normal serum Cbl concentrations [16]. Correct diagnosis of Cbl deficiency is however important because delayed appropriate treatment of an existing deficiency may lead to irreversible neurological damage and inappropriate treatment of a non-existing deficiency is, although not strictly harmful to the patient, an uncomfortable burden and a waste of medical services.

The availability of the accurate and precise measurements of the serum concentrations of methylmalonic acid (Mma) and total homocysteine (Hcy) have provided for the necessary new tools in diagnosing Cbl deficiency [17][18]. The serum levels of these two metabolites are increased in patients with Cbl deficiency, while in patients with folate (Fol) deficiency Hcy alone is elevated. Mma and Hcy are increasingly regarded as hallmarks for the diagnosis of clinical Cbl deficiency because their high (combined) sensitivity and specificity and their return to normal values after correct treatment [19][20]. The combined use of serum Cbl, Fol, Mma and Hcy currently provides the maximal utility in diagnosing Cbl deficiency: with this panel the diagnosis of a vitamin B12 deficiency can be made or excluded with high clinical accuracy in almost every patient [19-20].

In this chapter the difference between traditional and current view on Cbl deficiency and the consequences for test requesting patterns in diagnosing vitamin B12 deficiency are presented.

Materials and methods

During a 2-month period (july-september 1994) there were 624 samples submitted for assay of serum Cbl and or Fol at the Department of Clinical Chemistry of the University Hospital Rotterdam. The serum level of Cbl was less than or equal to 150 pmol/l (low-normal and low values) in a total of 49 samples. These were matched to 49 age and sex related samples with vitamin B12 > 150 pmol/l (normal and high values). In these 98 samples, Mma and Hcy were measured, together with creatinine to rule out high metabolite concentrations caused by renal dysfunction.

Vitamin B12 and folate were determined simultaneously by a competitive protein binding radioassay (SimulTRAC-S vitamin B12[^{57}Co]/Folate[^{125}I], Becton Dickinson, Etten-Leur, the Netherlands). Reference values serum Cbl, 120-640 pmol/l and serum Fol, 7.1 - 23.0 nmol/l.
Total homocysteine was measured after reduction and precolumn derivatization with SBD-F on a reversed-phase HPLC column, using fluorescence detection [21][22]. Reference values Hcy, 7.0 - 22.0 μmol/l.
Methylmalonic acid was measured at the department of Clinical Chemistry (Joern Schneede M.D.), University of Bergen, Norway, using capillary electrophoresis. This new procedure has not been published yet but is comparable in precision, accuracy and absolute values to the HPLC method published recently [23].
Interpretation of results; < 0.25 μmol/l : functional Cbl deficiency unlikely; 0.25 - 0.35 μmol/l: borderline, check renal function; > 0.35 μmol/l: functional Cbl deficiency, if renal dysfunction is excluded (serum creatinine < 110 μmol/l).

Results and discussion

In table 8.1 all results of Cbl, Fol, Mma and Hcy measurements in sera from 49 patients with low or low-normal vitamin B12 (Cbl < 150 pmol/l) levels and results of the same measurements in samples from the age and sex matched control samples (normal and high vitamin B12, Cbl > 150 pmol/l) are presented. None of the patients with borderline or elevated Mma or Hcy values had renal function impairment (results not shown).
Only 9 of the 49 patients (18 %) with low or low-normal serum Cbl values (patient no. 1,3,5,18,20,25,29,39,41) showed a substantial elevation of serum Mma values. All of these 9 patients had elevated Hcy values as well, but only 2 presented with hematological abnormalities. In a recent study, Savage and colleagues [20] showed a sensitivity of 99.8% for combined Mma and Hcy measurements in the detection of Cbl deficiency. Moreover, they concluded that normal levels of both analytes (Mma and Hcy) rule out clinically significant Cbl deficiency with virtual certainty. Based on these results and conclusions, all 9 patients mentioned can be considered Cbl deficient.

table 8.1 *Results for cobalamin (Cbl), folate (Fol), methylmalonic acid (Mma) and total homocysteine (Hcy) measurements in patient serum.*

Low and low-normal vitamin B12 values					*Normal and high vitamin B12 values*				
no	Cbl pmol/l	Fol nmol/l	Mma μmol/l	Hcy μmol/l	no	Cbl pmol/l	Fol nmol/l	Mma μmol/l	Hcy μmol/l
1a	110	7.8	0.43	50.3	1	328	8.0	0.21	11.9
2	136	3.9	0.17	19.8	2	458	6.0	0.20	16.5
3a	59	7.6	1.55	39.3	3	464	-	0.06	7.5
4	133	9.0	0.20	9.6	4	234	-	0.11	7.1
5a	141	7.2	1.13	30.7	5	397	13.9	0.21	10.5
6b	141	12.2	0.28	13.8	6	389	5.5	0.10	-
7c	92	2.0	0.20	52.4	7	225	8.4	0.12	15.2
8d	119	9.8	0.16	13.2	8	311	11.9	0.14	12.4
9	148	12.0	0.24	15.1	9	281	43.1	0.17	7.7
10	141	9.0	0.08	11.3	10	225	41.7	0.13	6.9
11	120	-	0.09	16.0	11	420	14.4	0.14	11.3
12	130	8.1	0.17	12.5	12	338	10.9	0.09	9.9
13	127	8.5	0.13	10.6	13c	188	2.9	0.06	82.2
14d	105	9.3	0.06	16.2	14	299	5.7	0.15	18.3
15	132	12.1	0.15	28.0	15	319	8.4	0.09	19.0
16b	103	5.8	0.25	15.8	16c	194	5.5	0.18	36.4
17b	118	15.9	0.25	13.9	17	229	9.8	0.22	14.3
18a	112	9.7	0.36	95.7	18	750	8.3	0.22	22.1
19	145	11.4	0.20	14.0	19	176	7.8	0.15	-
20a	136	15.7	0.57	30.5	20	576	11.3	0.14	28.0
21	124	6.2	0.12	17.1	21c	171	4.9	0.19	27.0
22	117	-	0.20	14.4	22	527	3.9	0.11	20.0
23	97	38.6	0.12	12.9	23	271	10.7	0.17	6.2
24b	130	-	0.30	21.6	24	977	-	0.10	12.4
25a	134	-	0.43	37.4	25	197	4.7	0.06	11.1
26	148	5.4	0.19	29.0	26	682	5.3	0.08	20.2
27c	48	5.0	0.09	52.6	27	236	12.5	0.14	10.7
28	129	5.3	0.13	12.9	28	249	9.4	0.12	10.8
29a	96	>45	1.03	41.2	29	289	7.9	0.09	18.6
30c	96	5.9	0.18	31.0	30c	272	4.5	0.15	35.5
31	116	-	0.17	9.8	31	311	4.3	0.07	-
32	111	11.1	0.11	15.4	32	316	6.8	0.09	26.7
33c	107	2.0	0.23	51.5	33c	330	5.8	0.09	31.5
34	144	4.9	0.13	21.0	34	268	7.5	0.25	13.1
35	129	5.6	0.09	9.9	35	215	8.4	0.23	7.5
36	142	>45	0.04	10.6	36	340	6.0	0.18	19.2
37	140	4.9	0.17	19.3	37	380	16.6	0.07	9.9
38	141	6.0	0.24	16.8	38	209	6.1	0.23	15.2
39a	123	8.3	0.36	46.8	39	310	9.1	0.09	-
40c	80	2.8	0.10	41.1	40	325	8.8	0.08	7.9
41a	64	4.2	1.05	42.8	41	319	6.5	0.12	19.2
42c	35	2.0	0.14	88.2	42	260	-	0.11	12.5
43	141	15.5	0.15	10.8	43	187	9.8	0.12	11.6
44b	119	7.6	0.26	19.2	44	550	>45	0.05	11.5
45	150	11.5	0.22	7.2	45	233	10.7	0.12	10.8
46	129	2.6	0.15	62.6	46	660	-	0.07	7.4
47	141	9.5	0.15	16.9	47	447	7.4	0.07	13.2
48	137	11.7	0.12	10.1	48	775	12.3	0.14	10.8
49d	92	8.7	0.19	17.9	49	198	10.4	0.10	17.0

a=functional Cbl deficiency; b=borderline Mma deficiency;
c=functional Fol deficiency; d=falsely low Cbl concentrations.

This diagnosis would probably be missed or delayed in 4 of these 9 patients (44%) if only traditional parameters would have been available, because in these samples (no. 5,20,25,39) Cbl values were low normal (range 123 - 141 pmol/l) and hematological and biochemical parameters were normal or missing (not ordered). In 5 additional patients (10 %) with low or low-normal Cbl values the Mma results were "borderline" (Mma values between 0.25 - 0.35 μmol/l) while Hcy values were within the reference range and no hematological and or biochemical abnormalities were seen (patient no. 6,16,17,24,44). Because elevation of Mma is typically an early event, often preceding the depression in serum Cbl or elevation of Hcy [16], these patients can be considered developing a Cbl deficiency. This could have been seen from Cbl measurements in 3 of 5 "borderline" Mma samples.

In the low and low-normal Cbl group 6 patients presented with both low Cbl and Fol concentrations (patient no. 7,27,30,33,40,42) of which 4 had severe anemia. It is remarkable that in these patients Hcy values are elevated and Mma values are normal. This suggests that the Fol deficiency is the primary cause and Cbl dependent Mma metabolism is unimpaired despite the low serum Cbl. The low serum Cbl level can be readily explained by assuming that the low Fol concentration causes an impairment of the ileal absorption of Cbl. One may conclude that in the absence of competition for Cbl by methylfolate- and Cbl-dependent methionine synthetase even the limited amount of available Cbl is sufficient for maintenance of normal Mma metabolism.

Additionally, in the low and low-normal Cbl group 28 patients had low-normal Cbl values (Cbl between 120 and 150 pmol/l). Considering the results of the test panel as presented in table 8.1, 6 (or 21%) of these patients (no. 5,6,20,24,25,39) have metabolic Cbl deficiency and need treatment, the results for 2 patients (no. 26 and 46) are doubtful, 1 patient (no. 46) is Fol deficient and all other patients need no further treatment.

Finally in 3 samples (patient no. 8, 14 and 49) of the low Cbl group, Cbl concentrations could be considered inaccurate (falsely low) because all additional parameters (Fol, Mma and Hcy) are normal and Cbl deficiency can be ruled out.

In the matched control group no Mma elevations are seen, even if Fol and or Hcy are elevated due to Fol deficiency, with exception of patient no. 34 with

borderline Cbl deficiency. This patient was excluded however, because of renal function impairment (serum creatinine 239 μmol/l).

These results and those of others [11,12,14,16,19,20] indicate that (1) hematological data cannot be used as a selection criterium for serum Cbl requests, (2) one in five patients with low-normal Cbl has metabolic Cbl deficiency and needs treatment and (3) measurements of serum Mma and Hcy improve the quality of test requests in diagnosing and distinguishing Cbl and Fol deficiency.

To improve the quality of test requesting patterns in diagnosis of Cbl deficiency, in sera from all patients with low or low-normal Cbl values, from all patients with unexplained hematologic abnormalities and from all patients with unexplained neuropsychiatric abnormalities Mma and Hcy should be measured. Further studies are required to answer the question whether Mma and Hcy should be considered as subsidiary investigations after Cbl and Fol values are obtained (according to the guidelines of the British Committee for Standards in Haematology [24]) or should be the essential primary tests in the investigation of Cbl deficiency.

References

1. Fraser CG, Woodford FP. Strategies to modify the test-requesting patterns of clinicians. Ann Clin Biochem 1987; 24: 223-231.
2. Young DW. Improving laboratory usage: a review. Postgrad Med J 1988; 64: 283-289.
3. Finn AF, Valenstein PN, Burke MD. Alteration of physicians orders by nonphysicians. JAMA 1988; 259: 2549-2552.
4. Davidoff F, Goodspeed R, Clive J. Changing test ordering behavior. Med Care 1989; 27: 45-58.
5. Bareford D, Hayling A. Inappropriate use of laboratory services: long term combined approach to modify request patterns. BMJ 1990; 301: 1305-1307.
6. Tierney W, Miller ME, McDonald CJ. The effect on test ordering of informing physicians on the charges for outpatient diagnostic tests. N Engl J Med 1990; 322: 1499-1504.
7. Speicher CE. Decision-oriented test request forms. A system for implementing practice parameters in laboratory medicine. Clin Lab Med 1991; 11: 255-265.
8. Peters M, Broughton PMG. The role of expert systems in improving the test requesting patterns of clinicians. Ann Clin Biochem 1993; 30: 52-59.
9. Studnicki J, Bradham DD, Marshburn J, Foulis PR, Straumfjord JV. A feedback system for reducing excessive laboratory tests. Arch Pathol Lab Med 1993; 117: 35-39.

10. McRae TD, Freedman FL. Why vitamin B12 deficiency should be managed aggressively. Geriatrics 1989; 44: 70-79.
11. Lindenbaum J, Healton EB, Savage DG, Brust JCM, Garret TJ, Podell ER, Marcell PD, Stabler SP, Allen RH. Neuropsychiatric disorders caused by cobalamin deficiency in the absence of anamia or macrocytosis. N Eng J Med 1988; 318: 1720-1728.
12. Stabler SP, Allen RH, Savage DG, Lindenbaum J. Clinical spectrum of cobalamin deficiency. Blood 1990; 76: 871-881.
13. Schilling RF, Fairbanks VF, Miller R, Schmitt K, Smith MJ. Improved vitamin B12 assays: a report on two commercial kits. Clin Chem 1983; 29: 582-587.
14. Allen RH, Stabler SP, Savage DG, Lindenbaum J. Diagnosis of cobalamin deficiency I: Usefulness of serum methylmalonic acid and total homocysteine concentrations. Am J Hematol 1990; 34: 90-98.
15. Mollin DL, Anderson BB, Burman JF. The serum vitamin B12 level: its assay and significance. Clin Haematol 1976; 5: 521.
16. Lindenbaum J, Savage DG, Stabler SP, Allen RH. Diagnosis of cobalamin deficiency II: relative sensitivities of serum cobalamin, methylmalonic acid, and total homocysteine concentrations. Am J Hematol 1990; 34: 99-107.
17. Stabler SP, Marcell PD, Podell ER, Allen RH, Lindenbaum J. Assay of methylmalonic acid in the serum of patients with cobalamin deficieny using capillary gas chromatography-mass spectrometry. J Clin Invest 1986; 77: 1606-1612.
18. Stabler SP, Marcell PD, Podell ER, Allen RH, Savage DG, Lindenbaum J. Elevation of total homocysteine in the serum of patients with cobalamin or folate deficiency detected by capillary gas chromatography-mass spectrometry. J Clin Invest 1988; 81: 466-474.
19. Ueland PM, Refsum H, Stabler SP, Malinow MR, Andersson A, Allen RH. Total homocysteine in plasma and serum: methods and clinical applications. Clin Chem 1993; 39: 1764-1779.
20. Savage DG, Lindenbaum J, Stabler SP, Allen RH. Sensitivity of serum methylmalonic acid and total homocysteine determinations for diagnosing cobalamin and folate deficiencies. Am J Med 1994; 96: 239-246.
21. Araki A, Sako Y. Determination of free and total homocysteine in human plasma by high-performance liquid chromatography with fluorescence detection. J Chromatogr 1987; 422: 43-52.
22. Ubbink JB, Vermaak WJH, Bissbort S. Rapid high-performance liquid chromatographic assay for total homocysteine levels in human serum. J Chromatogr 1991; 565: 441-446.
23. Schneede J, Ueland PM. Automated assay of methylmalonic acid in serum and urine by derivatization with 1-pyrenyldiazomethane, liquid chromatography, and fluorescence detection. Clin Chem 1993; 39: 392-399.
24. Amos RJ, Dawson DW, Fish DI, Leeming RJ, Linnell JC. BCSH guidelines on the investigation and diagnosis of cobalamin and folate deficiencies. Clin Lab Haematol 1994; 16: 101-115.

CHAPTER 9

CONCLUSIONS

For the user of clinical chemical information, only a few performance characteristics are of interest: the possible errors in the information (precision and accuracy), the clinical sensitivity and specificity of the laboratory tests, the time between the submission of the samples and the results to become available (turnaround time) and the costs associated with the analysis. A quality clinical chemistry laboratory test should therefore be defined as "the most appropriate laboratory test, correctly performed at the lowest cost, and reported within an optimal time frame".

Specifications have to be set for the performance characteristics mentioned in order to have quantitative and objective criteria for evaluating laboratory performance. Considering precision and accuracy, consensus is growing that specifications should be related to within-subject biological variation or, if that is too stringent, to achievable analytical performance characteristics (that is the state-of-art).

Total imprecision should be less than one-half of the average within-subject biological variation or less than the state of art achieved by the best 0.20 fractile of laboratories, whichever is the less stringent. The second approach may be used when data on biological variation do not exist.

Inaccuracy should be less than one-quarter of the group (within- plus between-subject) biological variation or less than one-sixteenth of the reference interval, when data on group biological variation do not exist or less than twice the ideal imprecision, if the above specifications are too demanding.

The study of accuracy in clinical chemistry is a difficult problem especially if all information on composition of reagents and analytical procedures is manufacturer dependent. In this thesis serum ferritin determination has been chosen as a good example for such a black box assay. It has been shown that compatibility between assay results from different kits is lacking. After implementation of a reference standard the inter-assay variability decreased, but still significant differences in the absolute values were found. Proper standardisation of ferritin assays is a major problem because an accuracy base is missing. Use of the mean values from different assays as a basis for accuracy assumes that the results from all methods will be randomly distributed around the true value. Regarding the results presented, this seems very unlikely. The reference system concept (reference

standard and reference method) therefore has to be used to provide a basis for accuracy.

This concept is also illustrated with the case of the serum sodium determination. There is a need for all sodium measurements (both direct and indirect) to be performed on a serum-water concentration basis. This thesis provides the necessary accuracy base, i.e. a candidate reference method for coupled sodium-water determination. The method can be used in the calibration procedure of field methods and to study the accuracy of consensus values in proficiency testing.

The use of a large number of serum specimens from patients for calibration of (new) analyzers in clinical chemistry closely approaches the goal of calibration, i.e. direct comparability of the results of a (new) method with those obtained with the comparison method in the appropriate test material. This leads to the question which comparative value is correct, i.e. which comparison methods yield good estimates of the "true value". Reference methods should be used for this purpose because these methods "have shown to have negligible inaccuracy in comparison with their imprecision". If these reference methods are absent, routine methods have to be used. In that case, calibration resulting in an adjustment of new data in accordance with cumulative patient means is the approach of choice. The use of the multi-point calibration procedure presented in this thesis (based on calibration according to the instructions of the manufacturer and NCCLS EP-9T) greatly improves the intra-laboratory comparability. If the procedure is based on reference method data, inter-laboratory comparability can be improved as well, and unbiased results are produced.

The proper use of turnaround time (TAT) as a quality aspect in clinical chemistry starts with continuous gathering of stat test TATs in the laboratory. This should be followed by establishment of pragmatic criteria for acceptability of TAT and action upon unacceptable results. Acceptable and pragmatic in this is a reflection of staffing, instrumentation and procedures on TAT limits for each test group. This is realized by defining TATmax as the 90th percentile of grouped test TAT results over a representative time period, in order to indicate the actual *state-of-art* of laboratory TAT performance. The third point is the implementation of a system for fully automated continuous TAT monitoring. The system presented in this thesis significantly reduced the number of tests exceeding TATmax.

Clinical chemists should start using TAT as a quality aspect of laboratory procedures. This means that they should start monitoring and publicizing their TAT for all stat testing in their institution and that in a dialogue with the clinicians mutual definitions and goals for TAT should be realized. With continuous monitoring and identification of factors influencing TAT or causing unacceptable TAT delay, TAT and therefore total quality improvement can be expected and the percentage displeased clinicians will certainly be reduced.
If a mutual agreement between clinical chemist and clinicians has been achieved on TATmax under certain well defined conditions, e.g. pH, blood gas and electrolyte measurements in heart surgery, the clinical chemist has to investigate whether this goal can be met in the central laboratory or that point-of-care testing (POCT) should be considered. Only a selected group of tests, those that require a TAT of 5 minutes or less, are candidates for POCT. Significantly reduced TAT is probably the most dominant feature of POCT. In addition to eliminating most of the opportunities for pre-analytical errors, the omission of steps associated with collecting, transporting and processing specimens has cut average TAT for most POCT to less than 5 minutes. Testing facilities introduced at the point-of-care do not replace the need for the corresponding testing at the central laboratory, they only supplement this testing. The main laboratory must still provide these tests 24 hours a day for all other patients and serves as a back-up facility for the POCT. Thus, the reduction in the numbers of tests at the central laboratory as a result of the availability of a POCT device only leads to minor savings in operating costs and, more important, in the costs of technical personnel. When testing is moved to the point-of-care most of the expenses incurred in the central laboratory remain essentially unchanged, making the costs of processing tests at the point-of-care appear to be higher. As noted before, however, there may be significant cost savings for other aspects of patient care. This requires shifting of the focus from the laboratory to the hospital as a whole when regarding expenditures. This combines cost-benefit analysis with cost-effectiveness analysis, that is measurement of outcome in economic terms and natural units such as quality of care and number of lives saved. In conclusion, the two most important criteria for the implementation of POCT ought to be increase of the quality of patient care and decrease of hospital costs.

Quality aspects of analytical results are to be applied to all testing regardless of where the tests are performed and by whom. So quality assurance procedures for point-of-care instruments should be identical to those in central laboratories in order to ensure the highest quality blood test results and to meet current or future regulations and/or standards. It is unrealistic to believe that conventional quality assurance procedures will be applied routinely in POCT and, if they are, that the operator would act appropriately upon the results. The quality control of performance must therefore be engineered into the system to make it as fool-proof as possible. A very important issue in point-of-care testing is the registration of measurement data from patient samples as well as from quality control material. The small, sometimes hand-held, point-of-care equipment should contain data storage facilities so as to collect measurement data for several patients together with patient identification information and registration of date and time of day. Other aspects of registration not to be forgotten are the data of the evaluation study of the point-of-care instrument and the education and training status of the testing personnel. In my opinion, the clinical chemistry department should hold the responsibility, which extends to (1) the choice and evaluation of the necessary methods and equipment, (2) the implementation of an appropriate quality control system, (3) supervision of the analytical process, (4) education and training of the testing personnel and (5) budget control.

Compact pH, blood gas and electrolyte analyzers with TATs within minutes are now available and can be used at the point-of-care (POC). The Gem Premier represents a new generation of such POC monitors. It utilizes polarography, potentiometry and conductance for the measurement of pH, blood gases, sodium, potassium, ionized calcium and hematocrit. This system produces results that show a significant correlation with results of established laboratory systems. Inter-instrument analyte comparisons (at the central laboratory and at the POC) were highly and significantly correlated for all analytes. The correlation coefficient however, was closer to unity in the comparison of each analyte in the central laboratory compared with the evaluation at the POC. Pre-analytical errors associated with pH, blood gas and electrolyte measurements are only relevant for central lab testing and therefore could be a possible explanation for the discrepancy between correlation data for central lab and POCT. Imprecision data

are acceptable for within-cartridge imprecision. The between-cartridge imprecision exceeded the criteria for maximum allowable imprecision based on one or more generally accepted strategies. This is partly compensated however, by the removal of sources of error in the pre-analytical phase. The inaccuracy of means of assay results was acceptable if judged according to reference method values. It is perfectly clear that the median TAT for POCT with the Gem Premier was significantly shorter compared to median TAT for central laboratory testing for the same analytes. Combined with the ease of use, data registration facilities, system lock-out features for QC or calibration failure and minimal maintenance this makes the Gem Premier suitable for use in a POC setting where the availability of test results with TAT's < 5 minutes is considered necessary.

To my opinion the clinical accuracy (sensitivity and specificity) of the results of the tests requested is the most important aspect to be considered when evaluating the quality of test requesting patterns. This is demonstrated by an example derived from the diagnosis of cobalamin (Cbl; vitamin B12) deficiency. The assay of serum Cbl is the only diagnostic test for vitamin B12 deficiency that is widely available. This test however, lacks both specificity and sensitivity. The availability of the accurate and precise measurements of the serum concentrations of methylmalonic acid (Mma) and total homocysteine (Hcy) have provided for the necessary new tools in diagnosing Cbl deficiency. The results in this thesis and those of others indicate that (1) hematological data cannot be used as a selection criterium for serum Cbl requests, (2) 1 of 5 patients with low normal Cbl has metabolic Cbl deficiency and needs treatment and (3) measurements of serum Mma and Hcy improve the quality of test requests in diagnosing and distinguishing Cbl and folate deficiency. To improve the quality of test requesting patterns in the diagnosis of cobalamin deficiency, in sera of all patients with low or low-normal Cbl values, all patients with unexplained hematologic abnormalities and all patients with unexplained neuropsychiatric abnormalities, Mma and Hcy should be measured.

SUMMARY

SAMENVATTING

The studies described in this thesis were undertaken to (1) illustrate the importance and the proper use of the traditional quality aspects precision and accuracy and to (2) focus on quality of laboratory service as a relatively new component of quality and its consequences for laboratory performance, with special emphasis on point-of-care testing.

The importance of accuracy based reference systems (reference preparation and reference method) in clinical chemistry is emphasized, using the measurement of serum ferritin as an example. In chapter 2 is demonstrated that the results obtained by different serum ferritin immunoassays were not comparable to each other, even after introduction of a reference preparation for the calibration of serum ferritin immunoassays. Using this NBSB 80/602 human liver reference standard for the calibration of ferritin immunoassays reduced the inter-assay variability, but represented only the first important step towards standardisation of ferritin immunoassays, because results from different assays still showed significant differences versus each other. To diminish inter-assay variability and to increase the accuracy of serum ferritin determinations, implementation of a reference method (based on detection of ferritin subunits after denaturation and separation by HPLC) in combination with the NBSB 80/602 standard, should be the second important step towards accurate standardisation of ferritin immunoassays.

The way in which the results of measurements are reported can affect the accuracy of test results. For the determination of sodium a definitive as well as a reference method have been developed. However, both methods report the assay results as molarity i.e. sodium concentration per litre serum or plasma. Therefore assay results are sensitive to the serum volume occupied by proteins and lipids. This sensitivity may lead to misinterpretation of electrolyte status if serum samples with abnormal protein and or lipid status are concerned. In chapter 3 a candidate reference method for coupled sodium-water determination based on ion-exchange sodium separation from the serum matrix followed by gravimetry as Na_2SO_4 and serum water determination by means of microwave evaporation is described. The candidate reference method can be considered a reference method because (1) the reference method values did not differ significantly from the definitive method values, (2) there is no known source for interferences or bias,

and (3) misinterpretation due to abnormal protein (or lipid) status is excluded results being presented as sodium molality (concentration per litre plasma water) or sodium molarity after correction for abnormal water contents.

In chapter 4 is shown that the calibration of two identical multitest clinical chemistry analyzers according to the instructions of the manufacturer resulted in significant differences in assay results of the two analyzers in twelve of the twenty-two methods. To remove these differences an additional multi-point procedure for method comparison and bias estimation was introduced. The use of this multi-point calibration procedure (based on calibration according to the instructions of the manufacturer and NCCLS EP-9T) greatly improved the intra-laboratory comparability if routine methods were used as comparison methods. When the procedure was based on reference method data, inter-laboratory comparability could be improved as well, and unbiased results were produced with both analyzers. I conclude that the multi-point calibration procedure therefore should be used on a regular basis for all measurement systems in clinical chemistry to order to (re)assure accurate assay results.

With regard to the quality of laboratory service, the specimen turnaround time (TAT) and the quality of test requesting patterns are considered most important. In chapter 5 a system of computer programs for fully automated monitoring of TAT of all stat test requests is presented. The program-related procedures took less than 30 minutes per day. The system was used primarily to establish pragmatic criteria for acceptability of TAT and to act upon unacceptable TAT. This resulted in a significant reduction of the number of tests exceeding the maximum allowable TAT. Secondary, the system was used to evaluate the impact of monitoring and organizational structures on TAT and to examine the need for decentralized testing.

If it is not possible to produce information within the desired TAT in the central laboratory, decentralized testing (point-of-care testing, POCT) should be considered. In my opinion (presented in chapter 6) point-of-care technology offers new possibilities for the clinical chemistry profession, provided the clinical chemists and their staff are prepared to change, to bring the service of the laboratory to the patient and to take responsibility for maintenance of the laboratory level of quality at the point-of-care (POC). Moreover, POCT, if used

for specific tests in areas where rapid TATs are critical for clinical and or financial reasons, can be a cost-effective extension of central laboratory services.
In chapter 7 all quality aspects presented previously pass in review in the evaluation of the Gem Premier, the first representative of a pH, blood gas and electrolyte analyzers especially developed for POCT. The purpose was to evaluate this multichannel microchemistry instrument in the clinical setting for which it is intended, both in terms of comparative performance against clinical chemistry laboratory instruments and its potential utility in decreasing TAT and pre-analytical error. The Gem Premier system produced results which show a significant correlation with results of established laboratory systems. Imprecision data were acceptable for within-cartridge imprecision. The between-cartridge imprecision exceeded the criteria for maximum allowable imprecision based on one or more generally accepted strategies. This was partly compensated however, by the removal of sources of error in the pre-analytical phase. The inaccuracy of means of assay results were acceptable if judged according to reference method values. Combined with the quick availability of results, ease of use and minimal maintenance, these qualities make the Gem Premier suitable for use in a POC setting where the availability of test results with TAT's < 5 minutes is considered necessary.
In chapter 8 the second important component of laboratory service, the efficiency of test requesting patterns, is demonstrated, using the differential diagnosis of cobalamin deficiency as an example. The results in this chapter and those of others indicate that measurements of serum methylmalonic acid and total homocysteine improve the quality of test requests in diagnosing and distinguishing cobalamin and folate deficiency. To improve the quality of test requesting patterns in diagnosing cobalamin deficiencies, in sera of all patients with low or low-normal cobalamin values, all patients with unexplained hematologic abnormalities and all patients with unexplained neuropsychiatric abnormalities methylmalonic acid and total homocysteine should be measured. Further studies are required to answer the question whether methylmalonic acid and total homocysteine should be considered as subsidiary investigations after cobalamine and folate values are obtained, or should be the essential primary tests in the investigation of cobalamin deficiency.

Het in dit proefschrift beschreven onderzoek werd uitgevoerd (1) om het belang van het gebruik van de traditionele klinisch-chemische kwaliteitsaspecten precisie en juistheid te illustreren en (2) om de kwaliteit van de dienstverlening vanuit het klinisch-chemisch laboratorium nader te definiëren en als extra kwaliteitsaspect te implementeren met speciale aandacht voor decentraal laboratoriumonderzoek.
Het belang van het gebruik van op juistheid gebaseerde referentiesystemen wordt benadrukt in hoofdstuk 2, waarbij de meting van serumferritine als voorbeeld werd gebruikt. Uit hoofdstuk 2 blijkt dat de resultaten van verschillende ferritinebepalingen significant van elkaar verschilden, ook na kalibratie van de methoden met het internationale referentiepreparaat NBSB 80/602. Die kalibratie verminderde weliswaar de inter-methodevariabiliteit, maar de methoden bleven veelal significant van elkaar verschillen. Om de inter-methodevariabiliteit verder te verminderen en tevens de juistheid van de ferritinebepalingen te verhogen, dient implementatie van een referentiemethode voor ferritine (bijvoorbeeld gebaseerd op de bepaling van ferritinesubunits na denaturatie en scheiding middels HPLC en gekalibreerd met het genoemde referentiepreparaat) als tweede belangrijke stap in de richting van een juiste standaardisering van ferritinebepalingen te worden gezien.
De manier waarop meetresultaten worden weergegeven kan de juistheid van het resultaat beïnvloeden. Voor de bepaling van natrium is zowel een definitieve als een referentiemethode ontwikkeld, maar beide rapporteren de meetresultaten als molariteit, d.w.z. natriumconcentratie per liter serum of plasma. Daardoor zijn de meetresultaten afhankelijk van het watergehalte van het serum of plasma, wat leidt tot mis-interpretatie van de meetuitkomsten indien het natriumgehalte in monsters met een afwijkend eiwit- en of lipidengehalte (en dus een afwijkend watergehalte) wordt bepaald. In hoofdstuk 3 wordt een kandidaat-referentiemethode beschreven die dit probleem oplost. De methode is gebaseerd op een scheiding van natrium van de matrix op basis van ionen-wisselings chromatografie gevolgd door een gravimetrische natriumbepaling en een serum waterbepaling met behulp van een magnetron. Deze methode kan als een referentiemethode worden beschouwd, daar de resulten niet significant van die van de definitieve methode bleken te verschillen, er geen bekende bron voor interferenties en onjuistheid bekend is en mis-interpretatie, veroorzaakt door een

afwijkend eiwit en of lipidengehalte uitgesloten is door weergave van de resultaten als natrium-molaliteit (concentratie per liter plasma water) of als natrium-molariteit na correctie voor een abnormaal watergehalte.
In hoofdstuk 4 wordt aangetoond dat de resultaten van 12 van de geteste 22 methoden op twee identieke analyse-instrumenten significant van elkaar verschilden indien de beide analyse-instrumenten gekalibreerd werden volgens de voorschriften van de betreffende fabrikant. Deze verschillen konden worden weggenomen door gebruik te maken van een meer-punts kalibratieprocedure die de kalibratie volgens voorschriften van de fabrikant uitbreidt met een op een methodevergelijking gebaseerde juistheidscorrectie volgens het NCCLS EP-9T protocol. Indien deze procedure werd gebaseerd op een vergelijking met routinemethoden werd daarmee de intra-laboratorium vergelijkbaarheid verhoogd en kon het uitgangspunt, een gelijkblijvend patiëntengemiddelde, worden gerealiseerd. Indien beschikbaar werden echter bij voorkeur referentiemethoden gebruikt, daar dan ook de inter-laboratorium vergelijkbaarheid en de juistheid verbeterd konden worden. Het is noodzakelijk om de meer-punts kalibratieprocedure periodiek te herhalen om de juistheid van de betreffende meetresultaten te kunnen handhaven.
Voor wat de dienstverlenende aspecten betreft, kan gesteld worden dat de doorlooptijden van laboratoriumonderzoek en de kwaliteit van de aanvragen voor laboratoriumonderzoek het meest belangrijk blijken te zijn. In hoofdstuk 5 wordt een volledig geautomatiseerd systeem voor de monitoring van de doorlooptijden van al het cito aangevraagde onderzoek beschreven. De benodigde procedures namen gemiddeld niet meer dan 30 minuten per dag in beslag. Het systeem werd primair gebruikt om pragmatische criteria voor de maximaal toelaatbare TAT (TATmax) op te stellen en te handelen bij overschrijding van de TATmax. Dit resulteerde in een significante reductie van het aantal testen dat de TATmax overschreed. Secundair werd het systeem gebruikt om de invloed van monitoring en organisatiestructuur op TAT te evalueren en om de noodzaak voor decentraal laboratoriumonderzoek (point-of-care testing, POCT) te onderzoeken.
Indien het niet mogelijk blijkt om in het centrale laboratorium resultaten te produceren binnen de gewenste TAT, dan dient POCT te worden overwogen. Naar mijn mening (weergegeven in hoofdstuk 6) dient zich met het beschikbaar

komen van point-of-care technologie een nieuw belangrijk aandachtsgebied aan voor de klinische chemie. Indien de klinisch chemicus daarop anticipeert en door actief beleid decentraal een laboratoriumkwaliteit garandeert, kan POCT een waardevolle aanvulling zijn van centraal laboratoriumonderzoek. Dit geldt evenwel voor een relatief klein aantal testen waarvoor onder speciale omstandigheden zeer korte doorlooptijden om klinische, organisatorische of financiële redenen gewenst zijn.

In hoofdstuk 7 passeren alle voorgaande kwaliteitsaspecten de revue bij de evaluatie van de Gem Premier, de eerste vertegenwoordiger van een pH-, bloedgas- en elektrolyten-analyser, speciaal ontwikkeld voor POCT. Deze analyser werd zowel op het laboratorium als op een intensive care geëvalueerd, met speciale aandacht voor de potentiële mogelijkheid tot reductie van TAT en pre-analytische foutenbronnen, zijnde de twee belangrijkste wapenfeiten van POCT. De Gem Premier produceerde resultaten die een significante correlatie met overeenkomstige methoden op het centrale laboratorium vertoonden. De binnen-cassette imprecisie is acceptabel, de tussen-cassette imprecisie overschrijdt echter de maximum toelaatbare imprecisie. Dit wordt echter (deels) gecompenseerd door de afwezigheid van de pre-analytische component in de totale imprecisie. De onjuistheid van de methoden is acceptabel gezien de resultaten van de vergelijkingen met de betreffende referentiemethoden. Dit, in kombinatie met de zeer korte doorlooptijd, het gebruiksgemak en het niet noodzakelijk zijn van onderhoudswerkzaamheden, maakt dat de Gem Premier geschikt is voor POCT in een omgeving waar TATs van < 5 minuten noodzakelijk geacht worden.

In hoofdstuk 8 wordt het belang van de tweede belangrijke komponent van de dienstverlening, de kwaliteit van de aanvraag, gedemonstreerd aan de hand van het onderzoek naar een cobalamine (Cbl) deficiëntie. Mijn resultaten en die van anderen toonden aan dat bepaling van methylmalonzuur (Mma) en totaal homocysteïne (Hcy) de kwaliteit van de aanvragen voor onderzoek naar diagnose van een Cbl-deficiëntie en het onderscheid met een folaatdeficiëntie verhogen. Verder onderzoek dient uit te wijzen of Mma en Hcy als aanvullende bepalingen op de bepaling van Cbl en folaat moeten worden gezien, of dat beide bepalingen primair bij verdenking op een Cbl-deficiëntie moeten worden aangevraagd.

Articles:

1 **van Suijlen JDE, Leijnse B.** Serumferritine. Tijdschrift NVKC 1989; 14: 12-18.

2 **van Suijlen JDE, Leijnse B.** Laboratoriumcertificatie. Analyse 1989; 7: 167-170.

3 **van Suijlen JDE, van Noord PCH, Leijnse B.** Accuracy of serum ferritin determinations in tissue preparations and human serum. J Clin Chem Clin Biochem 1990; 28: 43-48.

4 **van Suijlen JDE, Berrevoets CA, Leijnse B.** A candidate reference method for coupled sodium-water determination in human serum. J Clin Chem Clin Biochem 1990; 28: 817-824.

5 **van Suijlen JDE, Blijenberg BG, Boerma GJM, Leijnse B.** Calibration of Technicon Chem 1 multitest analysers. Eur J Clin Chem Clin Biochem 1991; 29: 205-208.

6 **van Eijk HG, Cleton-Soeterman MI, van Suijlen JDE.** IJzer in het lichaam. Forum Diagnosticum 1992; 2: 3-6.

7 **van Suijlen JDE, Struijk LP, Blijenberg B.** Evaluatie van de ABL 505 en de IL BG3. Tijdschrift NVKC 1992; 17: 269-273.

8 **van Suijlen JDE, Lindemans J.** Aspects qualitatifs et coût-efficacité des examins biologiques effectués auprès du patient: avantages et inconvénients par rapport au laboratoire central. Option Bio 1993; 105 (Suppl): 6-8.
Also published as:
van Suijlen JDE, Lindemans J. Qualitative aspekte und kostenwirksamkeit des point-of-care-testing im vergleich zum zentrallaboratorium: gemeinsamkeiten und unterschiede. Management & Krankenhaus 1993; 10: Supplement.

9 **Lindemans J, van Suijlen JDE.** Le rôle du laboratoire central dans les examins biologiques effectués auprès du patient. Option Bio 1993 (Suppl): 4-5.
Also published as:
Lindemans J, van Suijlen JDE. Die rolle des zentrallaboratoriums in der point-of-care-analytik. Management & Krankenhaus 1993; 10: Supplement.

10 **Blijenberg BG, Giskes E, van Suijlen JDE.** The performance of the Chem 1 revisited: An evaluation after 3 years. Eur J Clin Chem Clin Biochem 1993; 31: 787-791.

11 **van Suijlen JDE, de Lijster de Raadt J, van Toorenenbergen A.** Evaluatie van het ALK Magic Lite SQ allergie systeem. Tijdschrift NVKC 1995; Accepted for publication.

12 **van Suijlen JDE, van der TAS A, Zwang L, Slim P, Lindemans J.** Turnaround time as a quality aspect of clinical chemistry laboratory testing. Submitted for publication.

13 **van Suijlen JDE, Schlebusch H, Lindemans J.** Laboratory evaluation of and point-of-care testing with the Gem Premier pH, blood gas and electrolyte analyser. Submitted for publication.

14 **van Suijlen JDE, Brouwer HJ, ter Heerdt PGJ, Schneede J, Lindemans J.** The quality of test requesting patterns in diagnosing cobalamin deficiency. Submitted for publication.

Abstracts:

1 **van Suijlen JDE, Leijnse B.** Variability of serumferritin determinations. European Iron Club Meeting. September 1988. Frankfurt am Main, Germany.

2 **van Suijlen JDE, Leijnse B.** Development and use of reference methods. FEDERA 1990. Leiden, the Netherlands.

3 **van Suijlen JDE, van Vliet HHDM.** Bepaling van heparine afhankelijke antilichamen in patiënten met een heparine geïnduceerde trombocytopenie. Voorjaarsledenbijeenkomst NVKC. April 1992. Utrecht, Nederland. (Tijdschrift NVKC 1992; 17: 82).

4 **Starreveld JS, van Dijk JP, Kroos MJ, van Suijlen JDE, van Eijk HG.** The effect of iron on ferritin in human cytotrophoblasts. International meeting on Porphyrin Metabolism and Iron Metabolism. Mei 1992. Papendal, Nederland. (Neth J Med 1993; 42: A53).

5 **Starreveld JS, Kroos MJ, van Suijlen JDE, van Dijk JP.** Trophoblastic ferritin contents in relation to iron supplementation of the culture medium. XIIth Rochester Trophoblast Conference. Oktober 1992. Rochester (New York), USA.

6 **van Suijlen JDE, van der Tas A, Zwang L, Slim P, Lindemans J.** Doorlooptijden als extra kwaliteitsindikator van cito laboratoriumonderzoek. Wetenschappelijke vergadering van het 46e NVKC congres. April 1994. Veldhoven, Nederland. (Tijdschrift NVKC 1994; 19: 128).

7 **van Suijlen JDE, de Lijster de Raadt J, van Toorenenbergen A.** Bepaling van specifiek IgE met het Magic Lite SQ systeem en het Pharmacia CAP RIA systeem; correlatie met de intracutane huidtest. Wetenschappelijke vergadering van het 46e NVKC congres. April 1994. Veldhoven, Nederland. (Tijdschrift NVKC 1994; 19: 130).

Other activities:

1 **van Suijlen JDE, Leijnse B.** Vademecum. Termen, definities, normen, richtlijnen en algemene informatie bij het CCKL symposium "Labcertificatie in de gezondheidszorg ?" Maart 1989. Utrecht, Nederland.

2 **Gouget B, van Suijlen JDE, Perryman DW.** Organizing committee. **van Suijlen JDE.** Quality aspects and cost-effectiveness in point-of-care testing: Elements common to and differrent from central lab testing (oral presentation). Workshop Managing point-of-care-testing: a new challenge for the clinical chemist. Eurolab '93, 10th European Congress of Clinical Chemistry. April 1993. Nice, France.

3 **van Suijlen JDE.** Point-of-Care testing (lezing). PAOKC-cursus klinische chemie: Klinische chemie en intensive care geneeskunde. Mei 1994. Rotterdam, Nederland.

Jeroen Dirk Evert van Suijlen werd op 18 juli 1962 te Zeist geboren. Het VWO diploma werd in 1981 behaald aan de Gemeentelijke Scholengemeenschap te Doetinchem. Aansluitend hierop begon hij aan de studie Farmacie aan de Rijksuniversiteit Utrecht waar in 1987 het doctoraalexamen werd behaald. In de doctoraalfase werd het bijvak klinische chemie gevolgd op het Klinisch Chemisch Laboratorium van het St. Antonius Ziekenhuis te Nieuwegein (prof. dr. J.B.J. Soons). Per 1 september 1987 werd hij aangesteld als onderzoeksmedewerker op het Instituut Chemische Pathologie van de Erasmus Universiteit te Rotterdam (prof. dr. B. Leijnse en prof. dr. H.G. van Eijk). In dat dienstverband werd begonnen met het onderzoek dat uiteindelijk geresulteerd heeft in dit proefschrift. Per 1 oktober 1988 werd hij, eveneens als onderzoeksmedewerker, aangesteld bij de afdeling Klinische Chemie (dr. J. Lindemans) van het Academisch Ziekenhuis Dijkzigt / Sophia te Rotterdam. Op 1 november 1990 is hij aldaar begonnen met de opleiding tot erkend klinisch chemicus. Deze opleiding werd op 1 november 1994 voltooid.